Received: 30-11-2013
Accepted: 25-12-2013

ISSN: 2321-4902
Volume 1 Issue 4

Online Available at www.chemijournal.com

International Journal of Chemical Studies

Chemometric Characterisation of Surface Water Quality in Uruan, Nigeria

Imaobong Inyang Udousoro[1*], Ndoto Etim Ikpeme[2]

1. Department of Chemistry, Faculty of Science, University of Uyo, AkwaIbom State, Nigeria.
2. Department of Chemistry, Faculty of Science, University of Uyo, AkwaIbom State, Nigeria.

*[Email: imaobong2i@yahoo.com. Tel: +234 8028416551]

Chemometric techniques were applied to surface water datasets collected form Uruan, Nigeria. Water quality index (WQI) and pollution index (PI) were employed to evaluate the pollution status of each station and parameter respectively. WQI ranged from 8 (Excellent for drinking) to 641(Unsuitable for drinking); 13 stations recorded excellent and 2 stations were unfit for drinking. PI ranged from 0.003 (Class 1, No Pollution) to 11.91 (Class 5, Seriously Polluted). The presence of total dissolved solids, Mn and Fe (Class 5) compromised water quality for domestic use. Five principal sources accounted for 81.329% of total variation which may be explained by natural mineralization, wash load of top soil from agricultural soils, mineral organic component, anthropogenic input from industrial activities and fossil fuel exploitation. Sampling stations were classified into six clusters; stations of excellent quality (Group 6) represented reference sites for future management of surface water resource.

Keyword: Surface water quality, Chemometrics, Nigeria.

1. Introduction

Surface water (freshwater streams, lakes, rivers) is a principal natural resource. In recent times, one of the most important environmental issues is water contamination[1-4] due to natural (changes in precipitation inputs, soil erosion, weathering processes, ground water flow) and anthropogenic (urban, agricultural and industrial activities, increasing exploitation of water resources) activities. The level of undesirable chemicals in surface water environment is ever increasing and has caused major global concern[5]. Municipal and industrial wastewater discharges are constant pollution sources while surface runoff-a seasonal phenomenon, is largely affected by climate[1, 6, 7]. With the increasing concentration of chemicals in water sources, the quantity and quality of surface water resources available for propagation of aquatic life, drinking, industrial, irrigational and recreational purposes is affected by the day resulting in disease and poverty especially in the developing world[2]. For protection of public health as well as preservation of quality of these water sources, water quality monitoring is important. The problem with water quality data is its enormity and complexity. Most researchers still prefer the use of the univariate statistical method for explanation of water quality data, but the complexity and large variation of environmental datasets limit its usefulness. Moreover, since analytical and environmental studies have gone beyond concentration measurements, the application of geo-statistical method is recommended[6, 7]. With the rapid development of computer technology, statistics and geographic information system (GIS), it is possible to combine these techniques for processing of data of complex hydrogeochemical systems that are difficult to explain and document[6].

Chemometrics is the extraction of information from multivariate chemical data using statistics, mathematics and formal logic to design or select optimal experimental procedures, provide maximum relevant chemical data, and to obtain knowledge about chemical systems[8]. More recently, chemometrics has been extended to the use of linear algebra calculation methods, to make either quantitative or qualitative measurements of chemical data, primarily spectra. Chemometrics is not just a single tool but a range of methods including basic statistics, signal processing, factorial design, calibration, curve fitting, multivariate analysis, factor analysis, pattern recognition and neural networks. It can be applied to a wide range of problems in chemistry, chemical engineering and related fields. Various Chemometrics softwares are now available to recognize patterns in virtually any type of multidimensional analytical data[9-11].

Water quality index (WQI) and pollution index (PI) are important techniques for classifying surface water quality[3, 4, 12-14]. Recently, multivariate statistics (principal component analysis, cluster analysis, discriminate analysis, factor analysis, apportionment, correlation) a critical facet even in the earliest applications of chemometrics, are being frequently used in a wide range of problems in analytical and environmental studies[1, 15-25]. Chemometrics have been used for characterization and evaluation of surface, rain and ground water[1-3, 15, 18, 21, 27], classification (to discover patterns in complex data sets) and comparison of different samples[6, 14, 28]. It has been applied to assess temporal and spatial variations from natural and anthropogenic origins[2, 17, 19, 24, 29]; investigate variable and site correlations[8, 15, 19]; correlate chemical and sensory data in drinking water by factor analysis[8] and also used, as a multicriteria decision-making tool for identification of possible factors affecting water environments that are responsible for the variations in water quality. It is thus a valuable tool for developing appropriate strategies for effective management of water resources[1, 2, 8, 13, 16, 18-20, 27, 30, 32].

The aim of this study is to apply chemometrics techniques (WQI, PI, correlation, PCA and CA) as analytical tool to classify 15 physicochemical parameters and 31 sampling stations, effectively identify the main sources of pollution and determine the patterns in variations of surface water quality in Uruan.

2. Materials and Methods

2.1 Study Area, Sample Collection and Preservation

For this study, 31 surface water samples (Table 1) were collected from Uruan, a rural community in Akwa Ibom State in Nigeria in September 2003. The study areas are located within Uruan local government area which extends from latitude 4° 52' to 5° 08' N of the equator and longitude 7° 55' to 8° 10' E of the Greenwich Meridian. It is bounded on the East by Odukpani Local Government Area in Cross River State, on the South by Okobo Local Government Area, on the West by Nsit Atai and Ibesikpo Asutan Local Government Areas and on the North by Itu Local Government Area. (Fig. 1). Samples collected were stored in pre-cleaned (washed with mild detergent; soaked in 10% v/v HNO_3 for 24 hours and rinsed with distilled-deionised water) polypropylene bottles. Four composite samples of water were obtained from each sampling point. For metals, 1 M concentrated $HNO_3$was added to prevent loss of metals and microbial growth. All samples were preserved at 4 °C in ice coolers.

Table 1: Number of samples collected per location

Location	Sample Identity	Number collected
Adadia	3; 14; 27	3
Eman Uruan	5; 16	2
Ibiaku Uruan	8; 19	2
Ifiayong	1; 10; 21; 25; 28	5
Ikpa Uruan	7; 13; 18; 24; 31	5
Ishiet Uruan	4; 12; 15; 23; 30	5
Mbiaya Uruan	6; 9; 17; 20	4
Nwaniba	11; 22; 29	3
Use Uruan	2; 26	2
Total		31

2.2 Chemical Analysis

A total of 15 physicochemical parameters namely pH, temperature (Temp), electrical conductivity (EC), total dissolved solids (TDS), total suspended solids (TSS), biochemical oxygen demand after incubation for five days (BOD_5), total hardness (TH), alkalinity (ALK), Cl^-, NO_3^-, NO_2^-, Mn, K, Na, and Fe were analysed. Temperature, pH, electrical conductivity (EC) and BOD_5 were measured at the sites using mercury glass thermometer, JENWAY 3305 pH meter, HACH 44600-00 EC meter with JYD-IA DO meter after incubation for five days respectively. Other physicochemical parameters

and metals levels were measured in the laboratory using standard procedures (APHA, 1992). All reagents used were of analytical grade from Sigma Aldrich. The analytical quality control was ensured through procedural blank measurements, duplicate analysis of water samples and standardisation of analytical instruments.

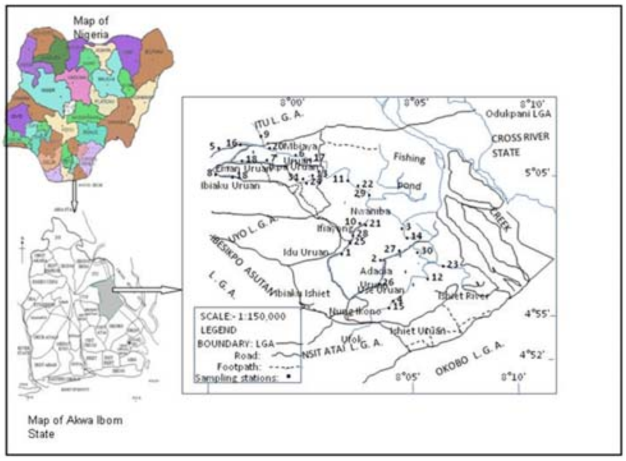

Fig 1: Location of Sampling Points in the Map of Uruan.

2.3 Statistical Analyses

Statistical analyses were performed using SPSS 19.0 software and Microsoft Office Excel 2007. The results of all measurements were subjected to the different chemometrics methods. Spearman's rank correlation coefficient is used to determine the strength of association between the physicochemical parameters[33]. Cluster analysis (CA) is used for grouping the parameters and the 31 sampling stations into similar classes[33]; CA is a statistical tool that classifies the true groups of data according to their similarities to each other. All variables were also standardized by Z-score mode before being subjected to hierarchical CA. Ward's procedure using Squared Euclidean distances is employed to identify the number of clusters[15,33,34]. Principal component analysis (PCA) is used to discover factors affecting surface water characteristics; PCA is a multivariate statistical method applied in environmental studies to explain data structures. The major objective of PCA is to find and interpret hidden complexes between dataset features. The initial principal components (PCs)

resulting from PCA are sometimes difficult to interpret; rotating the axis defined by PCA and constructing new variables is therefore necessary. Orthogonal rotation of the selected initial components with eigenvalues ≥1[35] to terminal factor solutions is achieved by Kaiser's varimax to reduce the dimensionality of the data, identify most significant variables and allow inference to the processes that control surface water chemistry in Uruan[17, 26, 33, 36]. The number of PCs is usually less than the number of the original variables. Varimax correlation coefficients (component loading) greater than 0.75, 0.75-0.50, and 0.30 are considered as strong, moderate and weak significant factor loading respectively[37]. In this study, component loading greater than or equal to 0.70 will be used for interpretation.

To assess the surface water environment quality in Uruan, pollution index (PI) for fifteen physicochemical parameters and water quality index (WQI) for each of the 31 sampling stations were computed. PI is a method of rating that shows the composite influence of individual parameter on the overall quality of water. From this rating, individual parameter that is of risk is

then identified. The rating value ranges from 0 - ≥5 and is computed[3]as:

$$PI = \sqrt{\frac{\left[\frac{C_i}{S_i}\right]^2_{max} + \left[\frac{C_i}{S_i}\right]^2_{min}}{2}} \qquad \text{Eqn. 1}$$

Where: PI = Pollution index of individual parameter;

C_i = The mean concentration as determined;

S_i = World Health Organization drinking water quality standard (WHO 2011).

Water quality index (WQI) may be defined as a single numerical expression of pollution, which reflects the composite influence of significant physical, chemical and biological parameters in water quality and would provide an opportunity to evaluate water conditions in relative terms[31]. Although no one WQI has found general acceptance[13], it remains an important tool for summarising large amount of water quality data for the general public and agency personnel; it communicates the quality of water to those with limited technical knowledge. It could be useful in comparing various bodies of water or to study trends in water quality at different points of a river, stream and other water sources; and could help pinpoint stream and river stretches which have altered significantly[10,11] and which, if necessary, could be investigated in greater details. WQI was computed using the Weighted Arithmetic Index. It has been reported that WQI based on additive indexes can be calculated very quickly, within minutes than the multiplicative type of indexes[13, 14, 28]. The quality rating scale for each parameter qi is computed using:

$$q_i = \left(C_i / S_i \right) \text{ x } 100 \qquad \text{Eqn. 2}$$

Where: C_i = Concentration of water sample;
S_i = World Health Organization drinking water quality Standard (WHO 2011).

The relative weight W_i was computed by a value inversely proportional to WHO recommended standard S_i of the corresponding parameter

$$W_i = 1 / S_i \qquad \text{Eqn. 3}$$

Water quality index of a particular sampling station is given by:

$$WQI = \sum_{n=1}^{i=n} q_i W_n \qquad \text{Eqn. 4}$$

Where: q_i = quality rating of the ith parameter;
W_i = weight of the ith parameter;
n =number of the parameter consider.

The overall WQI is computed as follows:

$$Overall\ WQI = \sum q_i W_i \Big/ \sum W_i \qquad \text{Eqn. 5}$$

WQI with multivariate analysis (correlation, PCA and CA) were combined for interpretation to enable better understanding of the quality of surface water in Uruan.

3. Results and Discussion

Summary of selected descriptive statistics are presented in Table 2. More than half of the parameters significantly deviated from normality having standardized skewness outside the normal range. The values of the following parameters show standardized skewness outside the expected range- EC, Fe, K, Mn, NO_2^-, pH, TDS, TH; and standardised kurtosis outside expected range-Mn, NO_2^-,NO_3^-,TDS, TH. Deviations from normality for these parameters suggest input from anthropogenic sources[4]. All parameters except pH, TDS, Mn and Fe are within WHO[38] drinking water limit. BOD_5 has a mean value of 13.63 mg/l which is rather high.

Table 2: Summary of statistics of physicochemical properties of surface water in Uruan, Nigeria

Parameter	Count	Mean	Sd	Coeff. of variation	Min.	Max.	Range	Stnd. skewness	Stnd. kurtosis	WHO[38]
ALK	31	2.28	0.55	24.21%	1.30	3.00	1.70	-0.45	-1.26	200 mg/l
BOD_5	31	13.63	2.85	20.94%	9.73	18.74	9.01	1.06	-0.82	
Cl^-	31	14.65	1.45	9.90%	12.70	17.00	4.30	0.30	-1.49	250 mg/l
EC	31	518.66	764	147.44%	27.20	2120	2092.8	**3.30**	0.56	1000 mg/l
Fe	31	1.451	1.494	102.94%	0.151	5.050	4.899	**2.81**	0.62	0.3 mg/l
K	31	0.17	0.09	51.94%	0.10	0.40	0.30	**3.07**	1.67	250 mg/l
Mn	31	0.111	0.23	207.58%	0.005	0.950	0.945	**7.77**	**12.67**	0.1 mg/l
Na	31	1.65	0.45	27.36%	0.80	2.20	1.40	-1.32	-1.07	200 mg/l
NO_2^-	31	0.04	0.07	181.42%	0.01	0.30	0.29	**8.28**	**13.97**	<3 mg/l
NO_3^-	31	0.15	0.05	32.05%	0.10	0.22	0.12	-0.05	**-2.16**	50 mg/l
pH	31	5.78	0.46	8.03%	5.20	6.80	1.60	**2.08**	0.19	6.5-8.5
TDS	31	706.34	1488	210.75%	12.60	5040	5027.4	**5.85**	**6.24**	<600 mg/l
Temp	31	27.8	0.65	2.35%	27.00	29.00	2.00	0.49	-0.65	27-28 ^oC
TH	31	15.28	5.72	37.40%	9.00	30.00	21.00	**4.02**	**2.80**	100-300 mg/l
TSS	31	0.65	0.08	12.91%	0.47	0.76	0.29	-1.12	-0.47	<10 mg/l

Sd-Standard deviation. Min. -Minimum. Max. -Maximum. Stnd. -Standard
Bold red face -Standardised skewness and kurtosis outside normal range of -2 to +2.

3.1 Assessment of surface water quality in Uruan

Pollution index (PI) for individual parameters is used to assess the quality of surface water in Uruan (Table 3).

Table 3: Pollution index of individual parameters

Parameter	PI	Class	Status
pH	1.56	2	Slightly polluted
Temp	1.04	2	Slightly polluted
EC	1.50	2	Slightly polluted
TDS	5.94	5	Seriously polluted
TSS	0.06	1	No pollution
BOD_5	2.49	3	Moderately polluted
TH	0.64	1	No pollution
ALK	0.01	1	No pollution
Cl^-	0.06	1	No pollution
NO_3^-	0.003	1	No pollution
NO_2-	0.07	1	No pollution
Mn	6.72	5	Seriously polluted
K	0.001	1	No pollution
Na	0.008	1	No pollution
Fe	11.91	5	Seriously polluted

Water quality classification based on PI and the percentage of parameters in each category is shown in Table 4. The results reveal that TSS, TH, ALK, Cl^-, NO_3^-, NO_2^-, K, and Na, belong to Class 1(no pollution) and represents 53.3% of the parameters, with no threat to surface water quality. 20% (Class 2: pH, Temp and EC) of the parameters reveal slight pollution, 6.7% (Class 3: BOD_5 only) moderate pollution, and 20% (Class 5: TDS, Mn and Fe) serious pollution. Fe, Mn in water could encourage the growth of microbials resulting in water borne diseases. No parameter is found in class 4 (strong pollution).

Table 4: Parameter classification based on pollution index [8]

Class	PI	Status	% Parameter
1	<1	No pollution	53.3
2	1-2	Slightly polluted	20.0
3	2-3	Moderately polluted	6.70
4	3-4	Strongly polluted	0.00
5	>5	Seriously polluted	20.0

WQI which is an excellent management and administrative tool in communicating water quality information is also computed for the different sampling stations and the overall WQI of surface water presented in Table 5; and Table 6 is based on the classification system proposed by Sahu and Sikdar[16]. 14 water samples (seven stream and seven river stations) have excellent water quality, 5 water samples (one stream and four river stations) good quality, 7 water samples (six stream and one river stations) poor quality, 3 water samples (three stream stations)very poor water, and 2 water samples from two stream stations are unsuitable for drinking. WQI reveals that water from the same stream or river at different stretches have varying quality depending on the location and pollution source[20].

Table 5: WQI values for each sampling stations

ID	Location	$\sum q_i W_i$	$\sum W_i$	WQI
1	Ifiayong	9726.147	15.17	641
2	Use Uruan	472.748	16.17	29
3	Adadia	2597.807	17.17	151
4	Ishiet Uruan	964.931	18.17	53
5	Eman Uruan	3469.138	19.17	181
6	Mbiaya Uruan	6664.491	20.17	330
7	Ikpa Uruan	3679.312	21.17	174
8	Ibiaku Uruan	871.510	22.17	39
9	Mbiaya Uruan	6119.377	23.17	264
10	Ifiayong	944.602	24.17	39
11	Nwainba	326.5207	25.17	13
12	Ishiet Uruan	3071.479	26.17	117
13	Ikpa Uruan	523.715	27.17	19
14	Adadia	974.931	28.17	34
15	Ishiet Uruan	3479.138	29.17	119
16	Eman Uruan	6674.491	30.17	221
17	Mbiaya Uruan	3689.312	31.17	118

18	Ikpa Uruan	881.510	32.17	27
19	Ibiaku Uruan	6129.377	33.17	185
20	Mbiaya Uruan	954.602	34.17	28
21	Ifiayong	336.521	35.17	10
22	Nwainba	3081.479	36.17	85
23	Ishiet Uruan	533.715	37.17	14
24	Ikpa Uruan	2618.807	38.17	69
25	Ifiayong	9750.147	39.17	249
26	Use Uruan	496.7483	40.17	12
27	Mbiaya Uruan	961.602	41.17	23
28	Ifiayong	343.521	42.17	8
29	Nwainba	3088.479	43.17	72
30	Ishiet Uruan	540.715	44.17	12
31	Ikpa Uruan	2625.807	45.17	58
	Overall WQI	**86592.68**	**935.39**	**93**

ID - Sample identity

Table 6: Water quality index categories[4, 16]

WQI value	WQ category	Sampling stations
<50	Excellent	2, 8, 10, 11, 13, 14, 18, 20, 21, 23, 26, 27, 28, 30,
50-100	Good	4, 22, 24, 29, 31
100-200	Poor	3, 5, 7, 12, 15, 17, 19
200-300	Very Poor	9, 16,25
>300	Unsuitable for Drinking	1, 6

3.2 Relationships between physicochemical parameters

Spearman's rank correlation coefficients, a non-parametric method is applied to the results of all physicochemical parameters analysed for surface water in Uruan since more than half of the parameters are not normally distributed (Table 2). The resulting matrix is presented in Table 7. Strong relationships are indication of common origin. Significantly positive correlations exist between pH and temperature while significant negative correlations are observed with EC, TDS, Cl^-, NO_3^-, and Mn, indicating that solubility of ions in Uruan surface water depends on pH. Temperature shows positive correlation with K (p=0.05) and indicates that temperature increases the solubility of K in water; however, significant negative correlations were obtained between temperature and EC, TDS, Cl^- and Fe. EC reveals strong positive correlations with TDS, BOD_5, TH, alkalinity, NO_3^-but strong negative correlations with K and Na. This shows that EC in Uruan surface water is controlled by TDS, BOD_5, TH, alkalinity, and NO_3^-. Negative significant correlations were observed between TSS and TH, Cl^-, NO_3^-, Mn, K and Fe. Strong positive correlations were observed between TDS and BOD5, TH, alkalinity, Cl^-, NO_3^-, and strong negative correlations with K at p=0.05. BOD_5 shows significant positive correlation with alkalinity (p=0.01) and negative correlation with K (p=0.05). Significant positive correlations were found between TH and Cl^-, Mn and Fe (p=0.01). Alkalinity reveals positive correlations with K at p=0.05. Strong correlations between Cl^- and NO_2^-, Mn, Fe indicate common pollution source. NO_3^- shows negative correlations with NO_2^- and Na (p=0.01) while NO_2^- shows strong negative relationships with Na and Fe (p=0.05).Mn shows significant positive correlations with K and Fe.

Table 7: Spearman's rank correlation coefficients of parameters analysed in surface water form Uruan

Parameters	pH	Temp	EC	TDS	TSS	BOD_5	TH	ALK	Cl^-	NO_3^-	NO_2^-	Mn	K	Na	Fe
pH	1.000														
Temp	$.392^{*}$	1.000													
EC	$-.532^{**}$	$-.511^{**}$	1.000												
TDS	$-.577^{**}$	$-.463^{**}$	$.961^{**}$	1.000											
TSS	.336	.061	-.264	-.185	1.000										
BOD_5	-.270	-.111	$.552^{**}$	$.631^{**}$	.240	1.000									
TH	-.071	-.354	$.376^{*}$	$.403^{*}$	$-.465^{**}$	.270	1.000								
ALK	.084	-.102	$.513^{**}$	$.518^{**}$	.287	$.539^{**}$	.189	1.000							

Cl^-	-.359*	-.567**	.418*	.450*	-.376*	.207	.618**	.248	1.000						
NO_3^-	-.368*	-.273	.621**	.495**	-.417*	-.025	-.127	-.060	-.142	1.000					
NO_2^-	.160	-.017	-.292	-.217	.061	-.046	.194	.265	.495**	-.684**	1.000				
Mn	-.431*	-.054	.132	.306	-.446*	.252	.493**	-.106	.424*	-.094	.311	1.000			
K	-.104	.450*	-.476**	-.440*	-.443*	-.389*	-.084	-.385*	-.069	-.149	.157	.407*	1.000		
Na	.267	-.053	-.406*	-.287	.288	.004	.257	-.148	.124	-.607**	.384*	.189	.119	1.000	
Fe	-.095	-.479**	.004	-.044	-.554**	-.132	.580**	-.159	.525**	-.032	.389*	.493**	.317	.216	1.000

*. Correlation is significant at the 0.05 level (2-tailed). **. Correlation is significant at the 0.01 level (2-tailed). N = 31.

Table 8: Total variance explained by principal components

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.127	27.514	27.514	4.127	27.514	27.514	3.387	22.578	22.578
2	2.840	18.931	46.445	2.840	18.931	46.445	2.820	18.801	41.379
3	2.196	14.639	61.084	2.196	14.639	61.084	2.327	15.513	56.893
4	1.694	11.295	72.379^{a}	1.694	11.295	72.379	1.873	12.488	69.381
5	1.352	9.012	81.391	1.352	9.012	81.391	1.802	12.010	81.391

Extraction method: Principal component analysis

3.3 Data structure determination and source identification: In addition to PI and WQI, multivariate analysis (correlation, CA and PCA) is performed to enable a more accurate interpretation of the quality of surface water in Uruan. In this study PCA is applied to standardized matrix of the raw data described by 15 physicochemical parameters to compare the compositional patterns between water samples at a sampling station and to identify the factors that influence each one[17].

The initial PC and its eigenvalues and percentage of variance contributed in each PC are presented in Table 8. Fig. 2 shows the Scree plot of the eigenvalues for each which helps to identify the number of PCs to retain[19]. Both Table 8 and Fig. 2 indicate that the first five-components are the most significant [29]; and they explain 81.391% of the total variance in surface water quality of Uruan- 27.514% by PC1, 18.931% by PC2, 14.639% by PC3, 11.295% by PC4 and 9,012% by PC5.

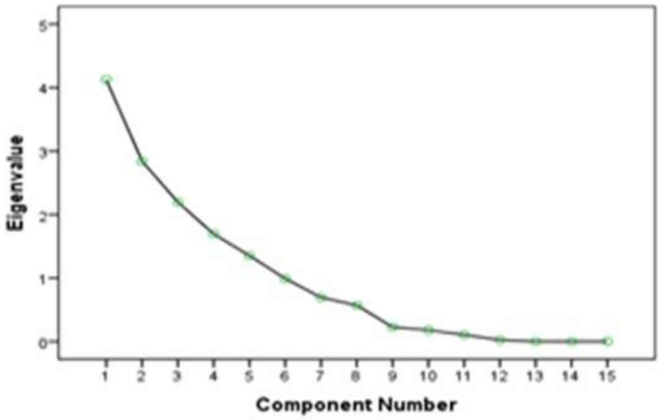

Fig 2: Scree plot showing eigenvalues and principal components.

Component loadings are presented for each parameter in the five PCs before varimax rotation along with the communalities (Table 9) and after varimax rotation with Kaiser Normalization converged in 10 iterations (Fig. 3). Also, Fig. 4 shows the component plots in rotated space.

Communalities provide an index to the efficiency of the reduced set of components and the degree of contribution of each parameter in the selected five-components. All parameters have been explained by more than 70% in this study except for Mn (67%) and alkalinity (66%).

Table 9: Unrotated component matrix^a for five-factor model

Parameter	Component					Communalities*
	1	2	3	4	5	
pH	-.448	.572	-.128	-.423	.123	.739
Temp	-.687	.124	-.219	.555	.207	.885
EC	.748	-.503	.074	-.283	.033	.899
TDS	.883	.288	-.056	.177	-.095	.905
TSS	.097	.498	-.695	-.227	.085	.798
BOD_5	.650	.175	-.341	.371	.271	.781
TH	.739	.444	.233	.271	-.116	.885
ALK	.555	.161	-.270	.129	.484	.658
Cl^-	.745	.084	.466	.069	-.141	.804
NO_3^-	.060	-.876	-.030	-.069	.036	.778
NO_2^-	-.154	.359	.434	-.148	.750	.924
Mn	-.103	-.279	-.116	.731	.170	.666
K	-.480	.068	.475	.486	-.180	.729
Na	-.081	.806	.099	.073	-.472	.894
Fe	-.001	.156	.861	-.097	.297	.864

Extraction method: Principal component analysis. * Extraction.
^{a}5 components extracted.

PCI has strong positive component loading on TDS, TH and Cl^- (.824, .880 and .855 respectively) explaining 27.514% of total variance (Table 8). This can be ascribed to natural source of enrichment; also positive correlation between TDS and TH (Table 7) suggests that TDS is of mineral origin. PC2 explains 18.931% of the total variance and is characterized by strong negative loading (-.840) on NO_3^- and strong positive loading (.886) on Na. Strong loadings of NO_3^- has been attributed to washing of top soil from agricultural fields[17] and decomposed nitrogen containing organic compounds from waste dumps[29].

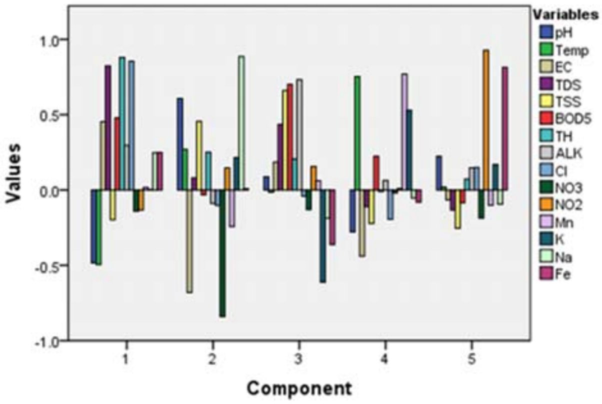

Fig 3: Plot of rotated component loadings for five-factor model

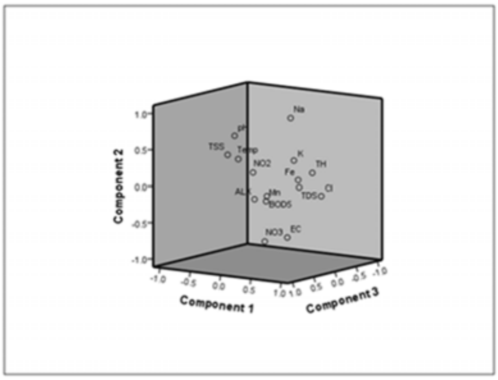

Fig 4: Component plot in rotated space

The parameters that contribute to PC3 are BOD_5 with component loading of .702 and alkalinity with component loading of .733 relating to mineral organic components of the surface water, and could come from storm runoff carrying animal and human feces and other organic biodegraded wastes. PC4 is ascribed to anthropogenic activities and has positive loadings on temperature (.753) and Mn (.770). Parameters that contribute to PC5 are NO_2^- and Fe with strong positive loadings of .928 and .815 respectively. This relate to anthropogenic activities from industrial products like leaching of Fe from metal scraps and fossil fuel contribution[27] of NO_2^- carried by rain water into surface water. The presence of Fe suggests, probably mixed anthropogenic and pedo-chemical origin in PC5.

3.4 Classification of sampling stations and investigated parameters by cluster analysis

Cluster analysis is applied to the water quality data set in order to group similar sampling stations (special variability)[18] and also allows identification of groups of associated parameters. Sampling stations were grouped into six significant clusters (Fig. 5): Group 1 contains stations 24, 31, 3, 4, and 14; Group 2 stations 22, 29, and 12; Group 3 stations 6 and 16; Group 4 stations 7, 17, 9, 19, 5, and 15; Group 5 stations 1 and 25; Group 6 stations 23, 30, 13, 21, 28, 11, 8, 18, 2, 26, 20, 27, and 10. Stations of excellent water quality (Table 6) which are clustered in group 6 (Fig. 5) represents reference sites for future monitoring program and, also provide baseline data for Uruan surface water for decision making.

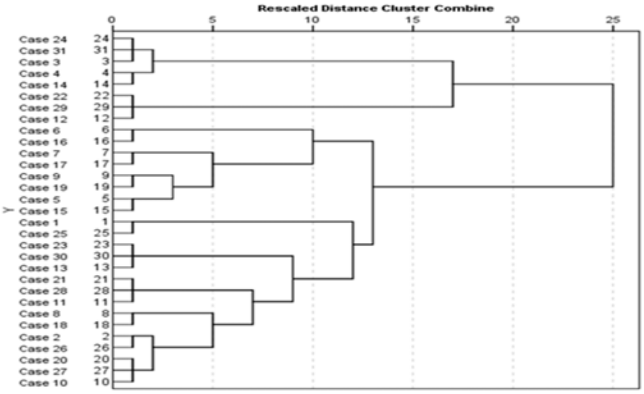

Fig 5: Dendrogram using Ward's linkage showing the clustering of sampling sites

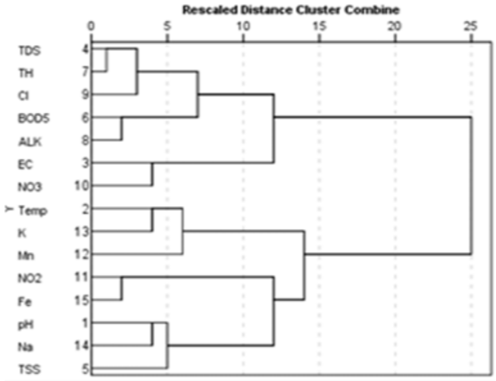

Fig 6: Dendrogram using Ward's linkage showing the clustering of parameters

Parameters investigated are grouped into five significant clusters (Fig. 6): Group 1 (**a**-TDS, TH, Cl^-, and **b**-BOD_5, ALK); Group 2 (EC and NO_3^-); Group 3 (Temp, K, and Mn); Group 4 (NO_2^- and Fe); Group 5 (pH, Na and TSS). The groups formed in the clusters represent varying pollution sources as discussed in sub-section 3.3 and shown in Fig. 3.

4. Conclusion

Chemometric treatments (WQI, PI, Spearman's rank correlations, PCA and CA) are applied to 15 parameters and 31 sampling stations from Uruan Local Government Area in Akwa Ibom State, Nigeria to enable pattern-recognition and source-identification of pollutants. The results of WQI reveal that 13 stations have excellent water quality (WQI <50) fit for drinking, 5 stations have good water quality (WQI = 50 - 100), 7 stations with poor water quality (WQI = 200-300), 3 stations with very poor water quality (WQI = 200 - 300), and water from 2 stations are unsuitable for drinking. PI ranges from 0.003-11.91. Individual parameters analysed are classified based on PI. The results show that 53.3% of the parameters belong to Class 1 (No pollution, PI <1: TSS, TH, ALK, Cl^-, NO_3^-, NO_2^-, K, Na), 20,0% in Class 2 (Slightly polluted, PI=1-2: pH, Temp, EC), 6.70% in Class 3(Moderately polluted, PI=2-3: BOD_5), 0.00% in Class 4 (Strongly polluted, PI=3-4: no parameter), and 20.0% in Class 5 (Seriously polluted, PI >5: TDS, Mn, Fe). Principal component analysis indicates that surface water composition of Uruan is affected by natural hydrochemistry and various human activities. Cluster analysis groups parameters into five clusters depicting the pollution sources identified by PCA. Sampling stations were classified into six groups according to different water quality. Group six clusters contains sampling stations with excellent WQI, therefore for quick assessment of surface water quality in Uruan, one site from each cluster could be sufficient to determine the water quality of the entire network, reducing the number of samples to be analysed saving both time and cost. Group six clusters could also be used as a baseline for future monitoring of surface water in the area to correct for deviations as a result of anthropogenic source of contamination.

5. Acknowledgement

The authors kindly acknowledge the staff of Akwa Ibom State Water Corporation, Uyo for analysing the samples.

6. References

1. Bayacioglu H. Surface water quality assessment using factor analysis. Water SA 2006; 32(1):389-394.
2. Armah FA, Obiri S, Yawson DO, Onumah EE, Yengog GT, Afrifa EKA, Odoi JO. Anthropogenic Sources and environmentally relevant concentrations of heavy metals in surface water of a mining district in Ghana: A multivariate statistical approach. J Environ Sci. Health Part A 2010; 45(13):1804-1813.
3. Amadi AN, Dan-Hassan MA, Okoye NO, Ejiofor IC, Aminu T. Studies on pollution hazards of shallow hand-dug wells in Erena and environs, North-Central Nigeria. Environ and Natural Resour Res 2013; 3(2):69-77.
4. Amadi AN, Olasehinde PI, Yisa J, Okosun EA, Nwankwola HO, Alkali YB. Geostatistical assessment of groundwater quality from coastal aquifers of Eastern Niger Delta, Nigeria. Geosciences 2012; 2(3):51-59.
5. Jarvie HP, Whitton BA, Neal C. Nitrogen and phosphorus in east coast British river: speciation sources and biological significance. Sci Total Environ 1998; (210/211):79-110.
6. Kowalkowski I, Zbytniewski R, Szpejna J, Buszewski B. Application of chemometrics in river water classification. Water Res 2006; 40:744-752.
7. Chen CH, Liu WL, Leu HG. Sustainable water quality management framework and a strategy planning system for a river basin. J Environ mgt 2006; 38:952-973.
8. Juahir H, Ekhwan TM, Zain SM, Mokhtar MB, Jalaludin Z, Jan IKM. The use of chemometrics analysis as a cost-effective tool in sustainable utilization of water resources in Langat river catchment. Am-Eurasian J Agric Environ Sci 2008; 4(2):258-265.
9. Juan DA, Tauler R. Chemometrics applied to unravel multicomponent processes and mixtures. Revisiting latest trends in multivariate resolution. Anal Chim Acta 2003; 500(1–2):195-210.
10. Illman DL, Callis JB, Kowalski BR. Process analytical chemistry: a new paradigm for analytical chemists. Am Lab 1986; 18:8-10.
11. Geladi P, Esbensen K. The start and early history of chemometrics: selected interviews Part 1. J Chemometrics 2005; 4(5):337–354.
12. Ishaku JM, Nur A, Matazu HI. Interpretation of groundwater quality in Fefore, Northeastern Nigeria. International J Earth Sci and Engr 2012; 5(3):373-382.
13. Udousoro II. Physiochemical quality assessment of Ogun river and the environmental impact on the Lagos Lagoon. Ph.D. thesis, Chemistry Department, Faculty of Science, University of Ibadan, Nigeria 1997, 386.
14. Bhargava DS. Use of water quality index for river classification and zoning of Ganga river. Environ Pollut Series B 1983; 6(1):51-67.
15. Forina M, Armanino C, Raggio V. Clustering with dendrograms on interpretation variables. Anal Chim Acta 2002; 454:13-19.
16. Sahu P, Sikdar PK. Hydrological framework of the aquifer in and around East Kolkata wetlands, West Bengal. India Environ Geol 2008; 55:823-835.
17. Vega M, Pardo R, Barrado E, Deban L. Assessment of Seasonal and polluting effects on the river by exploratory data analysis. Water Resour 1998; 32:3851-3592.
18. Librando V. Chemometric evaluation of surface water quality at regional level. Fresenius J Anal Chem 2000; 339:613-619.
19. Zhou F, Huang GH, Guo H, Zhang W, Hao Z. Spatio-temporal patterns and source apportionment of coastal water pollution in eastern Hong Kong. Water Res 2007; 41:3429-3439.
20. Lelpo P, Cassano D, Lopez A, Pappagallo G, Uricchio VF, De Napoli PA. Source apportionment of ground water pollutants in Apulian agricultural Sites using multivariate statistical analysis: case study of Foggia Province. Chem Central J. 2012; 6(2):1-12.
21. Gangopadhyah S, Gupta AD, Nachabe MH. Evaluation of ground water monitoring network by principal component analysis. Ground Water 2001; 39:181-191.
22. Quyang Y. Application of principal component and factor analysis to evaluate surface water quality monitoring network. Water res 2005; 39:2621-2635.
23. Simeonov V, Einax JW, Stanimirova I, Kraft J. Environmetric modeling and interpretation of river water monitoring data. Analytical and bio-analytical chemistry 2002; 374:898-905.

24. Panstar KM, Mujunen SP, Hatzimihalis G, Koutofides P, Mikkinen P, Wilkie PJ, Cannor MA. Multivariate data analysis of key pollutants in sewage samples: A case study. Anal chem Acta 1999; 393:181-191.
25. Singh KP, Malik A, Singh UK. Water Quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques: A case study. Anal Chem Acta 2005; 538:355-374.
26. Shrestha S, Kazama F. Assessment of surface water quality using multivariate statistical technique: A case study of the Fuji river Basin. Japan Environ Model software 2007; 22:404-475.
27. Mouli PC, Molian VS, Reddy SJ. Rain water chemistry at a regional representative urban site: Influence of terrestrial sources of ionic composition. Atmos Environ 2005; 39:999-1008.
28. Landwehr JM, Deininger RA. A comparison of several water quality indexes. J Water Pollut Control Fed 1976; 48(5):954-958.
29. Muhiyuddin WM, Ibrahim W. Komoo I, Pereira JJ. Relationship between Physical degradation and land cover Changes in the Langat Basin (in Malay), proceedings of the 1999 Langat Basin Research Symposium, LESTARI, University KebangSaan Malaysia, Bangi, Malaysia, 1999; 126-134.
30. Khalil WA, Goonetillke A, Kokot S, Carrol S. Use of Chemometrics methods and multicriteria decision-making for site selection for sustainable on-site sewage effluent disposal. Analytucal Chemistry Acta 2003; 506:41-56.
31. Joung HM, Miller WM, Mahannah CN, Guitjens JC. A generalized water quality index based on multivariate factor analysis. J Environ Qual 1979; 8(1):95-100.
32. Simeonov V, Stratis JA, Samara C, Zachariadas G, Voutsa D, Anthemidis A, Sofoniou M, Kouimtzis T. Assessment of the surface water quality in northern Greece. Water Res. 2003; 37:4119-4124.
33. Miller JN, Miller JC. Statistics and chemometrics for analytical chemistry. Edn 5, Pearson, London, 2005, 213-237.
34. Otto M. Mltivariate methods in analytical chemistry. (Eds Kellner R, Mermet JM, Otto M, Widmer HM), Wiley-VCH, Wenheim.
35. Kim JO, Mueller CW. Introduction to Factor Analysis: What it is and how to do it. Sage University Press, Newbury Park, 1987.
36. Kaiser HF. The Verimax criteria for analytical rotation in factor analysis. Psychometrika 1958; 23(3):187-200.
37. Liu CW, Lin KH, Kuo YM. Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. Sci Tot Environ 2003; 313(1-3):77-89.
38. WHO (2011a). Guideline for Drinking Water Quality, Edn 4, http://www.who.int/water_sanitation_health/publications/2011/dwq_chapters/ en/Retrieved on Monday 29 July, 2013.