

FACTOR ANALYSIS IN HYDROCHEMICAL SURVEY

BASSEY EDEM EPHRAIM

(Received 11 November 2002; Revision accepted 13 June 2003).

ABSTRACT

When subjected to factor analysis, the analytical data of water samples of the Calabar Flank area was transformed into factor score sets that are considered most consistent with known geological and environmental processes, and appropriate to the limits of analytical precision. Factor 1 reflects geochemical interactions of transported foreign materials with those occurring within the study area. Factor 2 reflects weathering & decompositions of rock materials occurring within the study area. Factor 3 reflects the oxidation of sulphide minerals occurring within the study area, and Factor 4 clearly indicates contributions from atmospheric precipitation. The present contribution demonstrates that factor analysis statistical reduction technique compresses bulk geochemical data into factor score groups that correspond to distinguishable controlling environmental factors.

KEY WORDS: Environmental Factors, Hydrochemistry, Multivariate Statistics, Factor Analysis, Calabar Flank

INTRODUCTION

Factor analysis has long been identified as a suitable multivariate statistical reduction method useful for the development or refinement of correlations among solutes and environmental factors (Hem, 1985). The technique that it employs to interpret the structures within the variance-covariance matrix of a multivariate data collection is to extract the eigenvalues and eigenvectors from the matrix of correlations of the observation. This has been applied to solve various hydrochemical problems.

Dalton and Upchurch (1978) applied R-mode factor analysis to the interpretation of mixing between sulphate and bicarbonate groundwater bodies. Matsoshchik (1975) utilized a similar technique in the study of hydrogeochemical conditions of groundwater of the Polish lowland. Dawdy and Feith (1967) also used this approach to study controls of groundwater composition in the Mojave River valley near Victorville, California, while Reeder et al (1972) applied it to aid in the assignment of sources to the solute load of the Mackenzie River system in Canada.

The beauty of factor analysis lies in the fact that it compresses bulk geochemical data into

factor score sets that correspond to distinguishable controlling environmental factors. This is most beneficial as no water reserve can be effectively developed, managed or safeguarded without prior knowledge of the controlling environmental factors.

R - mode factor analysis has a well-developed mathematical background, which permits mathematical simplification of the original data to yield a matrix of correlation coefficient. It involves relatively simple and clear step-by-step calculation with a certain system of the original data as a function of the obtained result (Matsoshchik, 1975).

The present study demonstrates the effectiveness and suitability of R-mode factor analysis statistical reduction method in hydrochemical survey.

METHOD OF STUDY

Fieldwork

Water samples were collected from rivers and streams within the floodplains of the Calabar Flank area (fig 1) and the recommendations of Beckman and Ryall (1976) were strictly adhered to during such samplings and preservations. Various laboratory techniques were later applied to analyze these samples (see Ephraim, 1998).

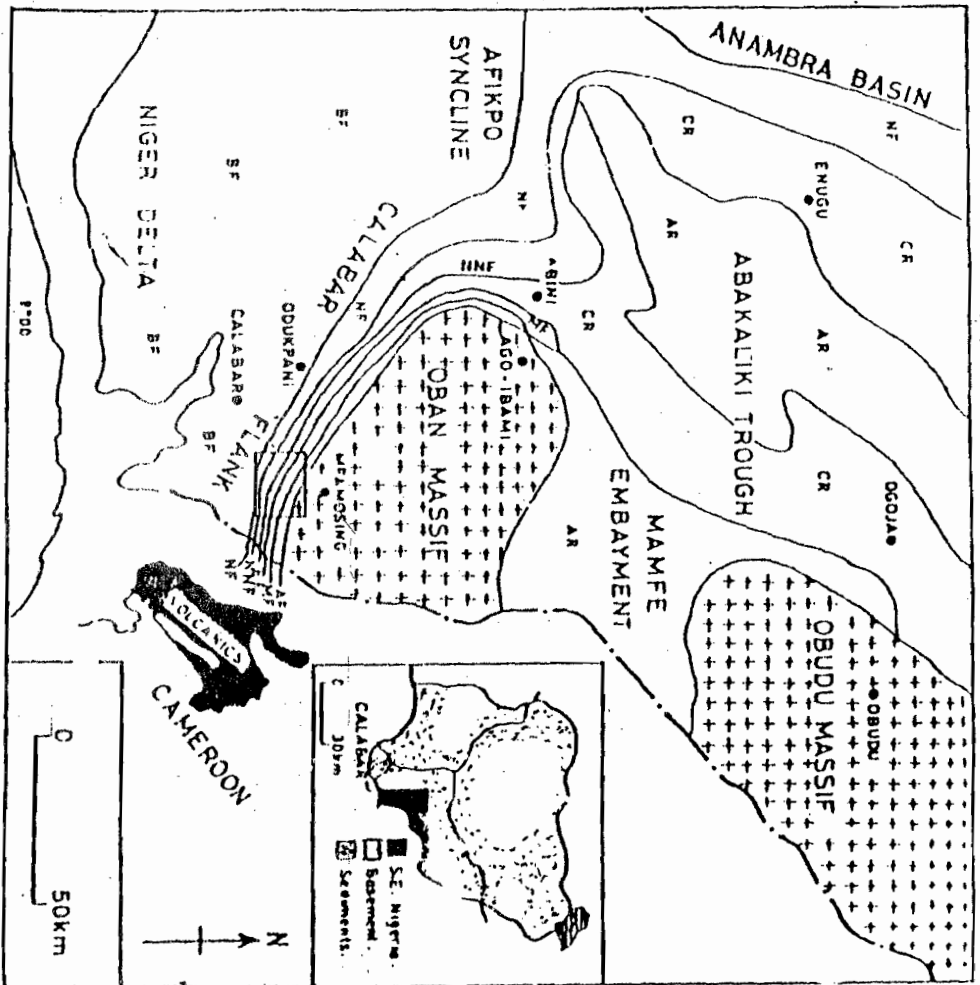


FIGURE 1: Geological sketch map of southeastern Nigeria showing the position of the Calabar Flank and location of study (compiled from Peters 1982 and Ekwueme 1995).

- LEGEND
- CR Cross River Group
 - AR ASU River Group
 - BF Benin Formation
 - NF Nkporo shale Formation
 - NMF New Netim Marl Formation
 - EF Ekenkpon shale Formation
 - MF Miamosing Limestone Formation
 - AF Awi Formation
 - BC Basement Complex
 - CV Cameroon Volcanics
 - Location of Major Towns
 - Study Area

Data Analysis

R – mode factor analysis was applied to the water analysis result and all computations were performed on an IBM computer at the University of Ibadan by means of SPSS software of Nie et al (1975).

Much attention was paid to the parameters chosen for the analysis, since careful data selection is a pre-requisite for successful factor analysis operation (Reyment and Joreskogg, 1993; Klován, 1975; Malinowski; 1991; Garrett and Nichol, 1969). Data adjudged not to be self-consistent and those believed to offer no new information in the analysis were excluded from the matrix used. The parameters selected as most reliable for the analysis were:

Conductivity, TDS, temperature, Dissolved oxygen, Na⁺, K⁺, Ca²⁺, Mg²⁺, Cl⁻, HCO₃⁻, CO₃²⁻, SO₄²⁻, SiO₂, NH₄⁺ and NO₃⁻ (see table 1). The matrix of the data employed contained 15 rows corresponding to the numbers of variables (parameters) and 15 columns corresponding to the number of observations (sample locations).

Prior to the principal component analysis (PCA), the selected data were pretreated by Kaiser normalization and after the PCA; the extracted eigenvectors were subjected to varimax rotation. Various factor models were computed, and after careful inspection of each of these, a 4 – factor model that accounted for 82.7% of the data variance was considered most consistent with known geological and environmental processes and appropriate to the limits of analytical precision

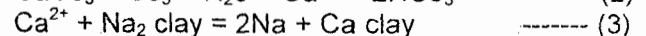
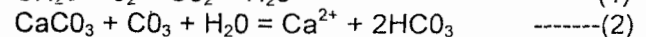
RESULTS AND DISCUSSIONS

The water analysis results presented in table 1 show that the water is of low mineralization with TDS values less than 1000mg. It is generally dominated by Na⁺ and HCO₃⁻, followed by CO₃²⁻, Cl⁻ and Ca²⁺, while components such as SO₄²⁻, Mg²⁺, and K⁺ are in comparison in lower proportion. The minor components (NO₃⁻, NH₄⁺, and SiO₂) are generally less than 1.0mg/l in the water. Measured values of temperature and dissolved oxygen ranges from 28.1 to 28.8°C and 3.6 to 5.4mg/l respectively, while conductivity values varies widely between 36.0 and 66.4µs/cm. Further details on the

physicochemical features of the water are in Ephraim (1998) and Elueze et al. (2001).

The rotated factor matrix is given in table 2, while that consisting of only variables with loadings greater than ± 0.55 is presented in table 3 together with the eigenvalues, total variance percentage and cumulative percentage. These variables are considered significant members of each factor.

According to these tables, factor I accounts for 37.3% of total data variance. The high loadings of Na⁺, K⁺, HCO₃⁻ and CO₃²⁻ reflect geochemical interactions of transported foreign materials with those occurring within the study area. The southward hydraulic gradient and the general relief of the area permits the transportation of waters and weathering products such as clays from the adjoining basement terrain into the drainage basin of the study area. The transported waters carry a different chemical composition for water mixing processes, while the clays bring in the much needed base exchange minerals for cation exchange operations, Foster (1950) and Krothe and Parizek (1979) agree that sodium bicarbonate waters exist in drainage basins that had or are presently experiencing cation exchange provided all of sodium carbonate, organic deposits and base exchange minerals are present beneath such drainage channels. The cation exchange operation that probably occurred in the area is explained by the following sets of reactions.



Reaction (3) drives the carbonate dissolution equilibrium to the right by absorbing Ca²⁺ and Mg²⁺ and the exchange sites thereby concentrating Na⁺ and HCO₃⁻ and probably K⁺ and CO₃²⁻. It is obvious from these equations and already known geological information that Ca²⁺, HCO₃⁻ and CO₃²⁻ are sourced from rocks of the study area, while Na⁺ and K⁺ are transported in solution or in base exchange minerals occurring in clays from the adjoining basement areas.

A system of classification of surface waters based on their TDS and Na⁺/(Na⁺ + Ca²⁺) ratio after Gibbs (1970), Feith (1971) and Faure (1991) corroborates these interpretations as the sampled water plots within the central part of the

Table 1: water Analysis Result
(All measurements in Mg/L unless otherwise stated)

Water Components	Group A (Siltstone, Sandstone & Shale)					Group B (Limestone & Marlstone)					Group C (Grit & Basement Rock)					
	Sample Number →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Ca ²⁺	2.4	4.8	2.4	3.2	3.2	3.2	15.2	12.4	12.6	6.4	4.0	4.0	4.0	2.4	2.4	3.2
Mg ²⁺	1.6	3.2	6.4	1.6	0.6	0.8	0.6	0.4	0.8	0.8	0.0	2.4	0.8	0.8	2.4	2.4
Na ⁺	16.6	20.5	8.3	16.3	14.9	19.1	18.2	7.6	9.0	27.8	26.4	25.5	17.7	17.0	22.8	
K ⁺	0.9	1.2	0.8	0.9	0.8	1.0	1.3	0.9	0.9	1.5	1.3	1.4	0.9	1.0	1.2	
Cl ⁻	9.5	21.5	10.5	10.5	9.2	11.8	13.0	10.9	10.9	20.6	10.9	21.9	8.7	12.2	15.2	
HCO ₃	18.3	18.3	18.4	18.4	15.4	18.3	30.5	15.3	18.3	27.46	27.5	21.4	18.4	18.3	24.4	
CO ₃	11.7	11.7	11.7	11.7	9.7	11.7	19.5	9.9	11.69	17.5	17.5	13.6	11.7	11.7	15.6	
SO ₄	1.4	5.6	1.9	2.1	2.4	3.4	5.2	5.3	5.3	1.62	2.0	6.4	1.9	2.9	1.3	
SiO ₂	0.91	0.77	0.69	0.92	0.71	1.16	0.82	0.81	0.77	0.33	0.94	1.06	0.9	0.72	0.95	
NO ₃	0.37	4.34	0.54	0.43	0.13	0.1	0.09	0.08	0.8	0.83	0.48	0.29	0.73	1.0	0.43	
NO ₂	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.01	0.01	0.0	0.0	0.0	0.0	0.0	0.0	
NH ₄ ⁺	0.08	0.32	0.22	0.36	0.34	0.09	1.87	0.92	0.84	0.07	0.06	0.01	0.21	1.26	0.7	
PO ₄	0.0	0.0	0.0	0.0	0.13	0.004	0.003	0.004	0.004	0.006	0.0	0.001	0.0	0.0	0.0	
Dissolved Oxygen	5.0	4.5	4.9	4.5	4.4	3.6	4.0	4.1	4.1	5.2	4.5	4.8	5.0	4.6	5.1	
TDS	62.3	86.7	60.3	64.6	56.2	69.3	103.5	62.7	69.5	103.7	89.7	96.6	82.5	67.9	86.0	
TSS	5	6	6	4	4	3	4	4	3	3	5	6	6	4	6	
Conductivity (Ms/Cm)	39.9	55.5	38.6	41.4	36.0	44.3	66.3	40.1	44.5	66.4	57.4	61.8	52.8	43.5	55.1	
Temperature (C)	28.2	28.3	28.1	28.6	28.6	28.6	28.6	28.7	28.8	-28.2	28.4	28.3	28.5	-28.5	28.3	
PH (pH Unit)	5.72	5.8	5.7	5.3	5.32	5.3	5.66	5.67	5.7	5.86	5.3	6.22	5.05	5.28	5.14	
Turbidity (FTU)	1	5	7	9	11	2	33	35	37	2	1	20	8	10	2	
Colour (Pt/(Co))	25	50	49	81	93	37	270	23	39	22	25	81	62	72	31	
Total Hardness (mg/l)	12.6	25.2	32.3	14.6	10.5	11.3	40.5	32.6	34.8	19.9	10.0	19.9	9.3	15.9	17.9	

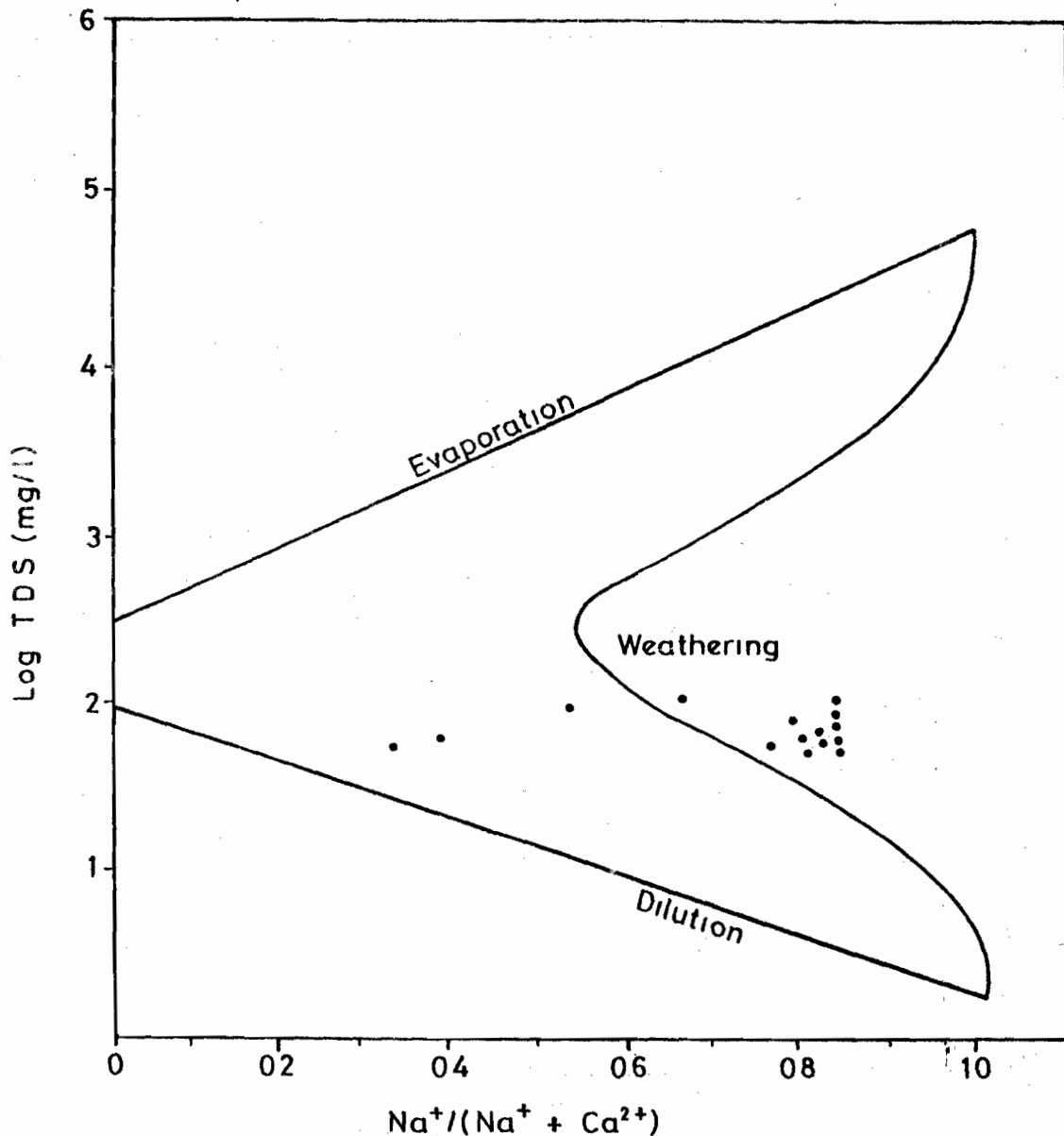
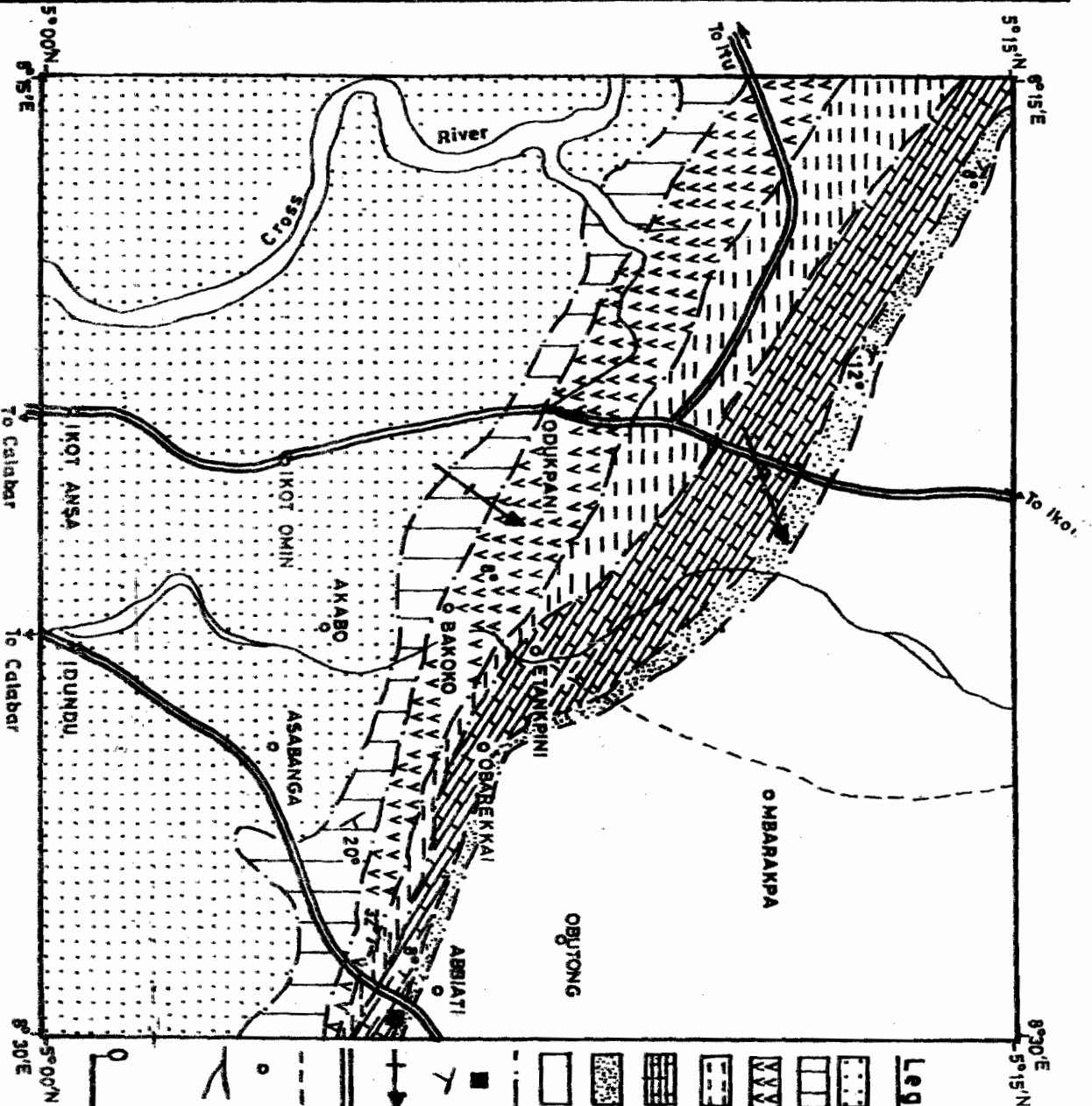


Fig. 2. Gibbs diagram of water analyses results.

Gibbs' boomerang (fig 2), an indication that the water is the rocks dominance type. This means that much of the dissolved solids in the waters have been derived from the weathering/leaching of rocks. The high loading of TDS and conductivity in factor 1 grouping is also consistent with these interpretations

Factor 2 accounts for about 23.3% of the

total data variance. The high loading of Ca^{2+} , NH_4^+ and temperature in this grouping reflects weathering and decomposition of rock materials within the study area. The occurrence of abundant deposits of limestone, marlstones and organic shale deposits together with the high temperature in the study area (fig 3) supports this interpretation since temperature is a very



Legend

- BENIN FORMATION
- NKPORO SHALE FORMATION
- NEW NETIM MARL FORMATION
- ERENKPON SHALE FORMATION
- MFAMOSING LIMESTONE FORMATION
- AWL FORMATION
- BASEMENT ROCK
- GEOLOGIC BOUNDARY
- MFAMOSING LIMESTONE QUARRY
- DIP AND STRIKE
- FOLD AXIS
- MAIN ROAD
- FOOT PATH
- LOCATION
- RIVER / STREAM



Fig.3: Geological Map of parts of the Calabar Flank, South Eastern Nigeria
 (Compiled from Ekwueme, 1985 and Ekwueme et al, 1995)

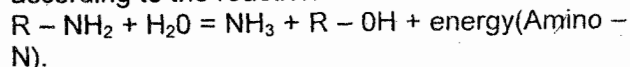
TABLE 2: R – MODE VARIMAX ROTATED FACTOR MATRIX

Variable	Factor 1	Factor 2	Factor 3	Factor 4
Conductivity	.96554	.11467	.03658	.13841
K ⁺	.95670	.02941	.00742	.21054
TDS	.94719	.1667	.4398	.18458
HCO ₃ ⁻	.89089	.12459	-.19392	-.21040
Na ⁺	.85455	-.43098	.09406	.03255
CO ₃ ²⁻	.80496	.28565	-.26801	-.5856
Ca ²⁺	.16890	.93493	.3985	-.06567
NH ₄ ⁺	.11586	.80016	.04659	-.01127
Temp	-.28747	.61290	.48185	-.38111
SO ₄ ²⁻	.00183	.23896	.86474	.19328
SiO ₂	-.05125	-.27743	.75697	-.21871
DO	.22348	-.54612	-.72105	.12001
NO ₃ ⁻	.08018	-.16123	-.00349	.81505
Cl ⁻	.45821	.29101	.07433	.69867
Mg ²⁺	-.27349	-.26848	.36830	.57559

TABLE 3: R – MODE VARIMAX FACTOR MATRIX(ONLY SIGNIFICANT LOADINGS)

	Factor 1	Factor 2	Factor 3	Factor 4
Loading	Conductivity 0.97 K ⁺ 0.96 TDS 0.95 HCO ₃ ⁻ 0.89 Na ⁺ 0.85 CO ₃ ²⁺ 0.80	Ca ²⁺ 0.93 NH ₄ ⁺ 0.80 Temp. 0.61	SO ₄ ²⁻ 0.86 SiO ₂ 0.76 DO 0.72	NO ₃ ⁺ 0.82 Cl ⁻ 0.70 Mg ²⁺ 0.58
Eigenvalue	5.60	3.48	1.72	1.61
Percent of Total Variance	37.3	23.3	11.4	10.7
Cumulative Percentage	37.3	60.6	72.0	82.7

important factor in rock weathering processes. Decomposition of the organic shale deposits probably involved the breakdown of protein into amino acids with the eventual release of NH₃ according to the reaction



The NH₃ released is quickly hydrolyzed to NH₄⁺

Factor 3 accounts for about 11.4% of the total data variance. The inverse association of SO₄²⁻ with DO in the group may be taken to suggest geological and hydrological controls by the oxidation of sulphide minerals such as pyrite. Hem (1985) states that when sulphide minerals undergo weathering in contact with aerated waters, sulphur is oxidized to yield sulphate ions which go into solution in the waters. This may be the main mode of acquisition of SO₄²⁻ in the waters since abundant pyrite minerals occur in the organic depositions of the study area (Adeleye and Fayose, 1978; Reymont, 1965).

The high loading of SiO₂ in the grouping could not be immediately explained but is believed to relate with the association of the organic deposits with marlstones in the area (Nyong, 1995, Reymont, 1965).

Factor 4 accounts for the lowest data variability of 10.7%. The ionic assemblages consisting of NO₃⁻, Cl⁻, Mg²⁺ clearly indicate atmospheric controls. In coastal area, atmospheric precipitation are usually very rich in Cl⁻, Mg²⁺ and probably NO₃⁻. Indications are that most, if not all of Cl⁻, Mg²⁺ and the low NO₃⁻ concentrations are sourced from such atmospheric precipitations.

SUMMARY AND CONCLUSIONS

The application of R – mode factor analysis to the studies of water samples of the Calabar flank area has been quite beneficial as it effectively revealed 4 factors that control the distribution of various physiochemical species within the waters.

The factor analysis technique compressed bulk geochemical data into factor score groups that correspond to distinguishable controlling environmental factors. This is undoubtedly most rewarding as it makes hydrochemical surveys easier, time saving and less tedious. Factor analysis therefore extends a beguiling assistance to hydrochemical survey researchers faced with a welter of complex data and little insight into the structure of the data.

REFERENCES

- Adeleye, D. R. and Fayos., E. A., 1978. Stratigraphy of the type Section of Awi formation, Odukpani area, Southeastern Nigeria. *J Min. Geol.*, 15(1): 33 - 37.
- Beckman, D. A. and Ryall, 1976. *Field geologist manual*. Parkville, Victoria: Australian inst. Min. and Metall: 195p
- Dalton, M. G. and Upchurch, S. B., 1978. Interpretation of Hydrochemical facies by factor analysis. *Groundwater*, 16 (4): 228 - 233
- Dawdy, D. R. and Feith, J. H., 1967. Applications of Factor Analysis In the study of chemistry of groundwater quality in Mojave Valley, California: *Water Resources Research*, 3: 505 - 510
- Ekwueme, B. N., 1985. The chemical composition and industrial Qualities of limestone and marls on the Calabar flank, Southeastern Nigeria. *J. Min. Geol.* 22 (1 & 2): 51 - 56
- Ekwueme, B. N., 1995. The Precambrian geology of Oban Massif, Southeastern Nigeria. In: Ekwueme, B. N., Nyong, E.E., And Petters, S.W. (eds), *Geological excursion guidebook To Oban Massif, Calabar Flank and Mamfe Embayment, Southeastern Nigeria*. Dec. - Ford Publishers Ltd, Calabar: 14 - 25
- Ekwueme, B. N., Nyong, E. E. and Petters, S. W., 1995. *Geological Excursion guidebook to Oban Massif, Calabar Flank and Mamfe Embayment, Southeastern Nigeria*. Dec. - Ford Publishers Ltd, Calabar: 14 - 25
- Elueze, A. A., Ephraim, B. E. and Nton, M. E., 2001. Hydrochemical Assessment of Surface water in parts of Southeastern Nigeria. *Mineral Wealth*, (119) 2001: 45 - 58.
- Ephraim, B. E., 1998. Water chemistry evaluations in parts of the Calabar flank, Southeastern Nigeria. M.Sc Project, University of Ibadan: 90p. (Unpubl.)
- Faure, G., 1971. *Principles and applications of inorganic Geochemistry*. Macmillan publishing Co., N. Y. 626p.
- Feith, H. I., 1971. Mechanism controlling world water chemistry. Evaporation - Crystallization process. *Science*, 172: 870 - 871
- Foster, M. D., 1950. The origin of high sodium bicarbonate in the Atlantic and Gulf Coastal Plains. *Geochim Cosmochim. Acta*, 1: 33 - 48.
- Garrett, R. G. and Nichol, I., 1969. Factor Analysis as an aid in the Interpretation of regional geochemical stream sediment Data. *Col. Sch. Mines Q.*, (64)1: 245-264.
- Gibbs, R. J., 1970. Mechanisms controlling world water chemistry. *Science*, 170: 1088 - 1090.
- Hem, J. D., 1985. Study and interpretation of the chemical Characteristics of Natural water. 3rd ed., U. S. Geological Survey water supply paper 1473.
- Klovan, J. E., 1975. R - and Q - mode factor analysis. In: McCammon, R.B. (ed.), *Concepts in Geostatistics*. Springer - Verlag, Berlin: 21- 69.
- Krothe, N. C., Parizek, R. R., 1979. An anomalous Occurrence of Sodium Bicarbonate waters in a flood plain in a carbonate Terrain. *Groundwater* 17 (6): 595 - 603
- Malinowski, E. R., 1991. *Factor analysis in chemistry*. Wiley Inter- Science publication. John Wiley and sons, Inc., N. Y.: 350p.
- Matsoshchik, A., 1975. Use of factor analysis for study of Hydrogeochemical conditions. *Moscow Univ. Geol. Bull.* 35: 40 - 44.
- Nie, N. H., Hill, C. H., Jenkins, I. G., Steinbrenner, K. and Bent, D. N., 1975. *Statistical Package for the Social Sciences*. McGraw Hill, N.Y., 2nd ed. 675p.
- Nyong, E. E., 1995. Cretaceous sediments in the Calabar Flank: In Ekwueme, B. N., Nyong, E. E., And Petters, S. W., (eds.) *Geological excursion guidebook to Oban Massif, Calabar Flank and Mamfe Embayment, Southeastern Nigeria*. Dec. - Ford Publishers Ltd, Calabar: 14 - 25
- Petters, S. W. and Ekweozor, C. M., 1982. Petroleum geology of Benue Trough and Southeastern Chad basin, Nigeria. *Amer. Assoc. Petrol. Geol. Bull.* 66: 1141 - 1149.
- Reeder, S. W., Hitchon, B and Levinson, A. A., 1972. Hydrogeochemistry of the surface waters of the Mackenzie River drainage basin, Canada. I. Factors controlling Inorganic composition, *Geochim Cosmochim. Acta.*, 36: 825 - 865.
- Reyment, R. A., 1965. Aspects of the geology of Nigeria. Ibadan Univ. Press, Ibadan, Nigeria: 145p.
- Reyment, R and Joreskogg, K. G., 1993. *Applied factor Analysis in The natural Sciences*. Cambridge Univ. Press. U.S.A.: 370p.