Assessment of water quality and identification of pollutants flowing into river musa using principal component analysis

M M Maina1* and A Yunusa2

1* Department of Agricultural and Environmental Engineering, Bayero University, Kano
2National Cereals Research Institute Badeggi, Bida, P.M.B.8, Niger State

Email: mainam@buk.edu.ng

Abstract. Available evidence shows that rivers in any water shed area acts as a carrier of various pollutants and affect water quality for various purposes. The main aim of this study was to assess and identify pollutants flowing into River Musa in Bida for irrigation purpose. Canadian Council Water Quality index (CCWQI) and Principal Component Analysis (PCA) were used. Water quality data were collected at different five stations along the river in 2018 during rainy and dry seasons. Ten water quality parameters (temperature, PH, EC, Mg, Na, Iron, Mn, Calcium, Potassium and SAR) were determined in five stations for assessing annual water quality index. Annual average water quality index at five locations are (74.47,72.85, 64.69,47.02 and 51.56). The results showed that three location are marginal and the remaining of the two are fair. The PCA was used to identify three major pollutants flowing to the river to be industrial, municipal and erosion. The WQI of the river is marginal and the implication is that the river is thus poised threat to high yielding results.

1. Introduction
River water quality depends upon its contents of living organisms, organic matters contents and the minerals contents. Good irrigation water quality has the potential to give a better yield of crops under good soil condition and good water management culture. Rivers in any water shed area acts as a carrier of various pollutants and thereby affecting the quality of water for different purposes. The economic value of every water decreases due to the presence of pollutants which affects human health and other aquatic forms of life [1]. Evidence reveals that human beings benefit immensely from water for drinking, agriculture and even for transportation and still they are the major contributors of water pollution through industrial effluents, agricultural wastes dumping, municipal wastes dumping etc. Hence, the assessment of various river pollutions and gathering of reliable information on it quality for different purpose is essential. Knowledge of the spatial variation and source apportionment of water quality parameters can provide detailed understanding of the environmental condition and help policy makers to establish measures for sustainable water management.

In Nigeria today, surface waters quality deterioration is not a new phenomenon. It was reported that there was increasing of industrial effluents in Ologe Lagoon [2]. According to Adamu and Ado [3] the dry season water quality of River Jakara in Kano is almost made up of municipal and industrial pollutions. All these affect the water of these places for various purposes.

The agricultural major river pollutants composed of phosphorus, nitrate and pesticides. Higher concentration of nitrates and phosphorus beyond or below established standard in surface waters reduce
their ability to support plant [4]. In agriculture, quality of water attached to its effects on soils, crops and management practices necessary to compensate problems linked to water quality. Various problems of soil degradation can be connected to irrigation water quality variation [5].

Identifying sources of river pollution should be seen as a means of water quality assessment. In Niger state, River Musa in Bida, is the source of surface water to Badeggi irrigation scheme. The Badeggi Rice irrigation scheme is located 12 km East of Bida town in the River Gbako plains, which among tributaries of River Niger that passes through the state. The scheme supplied water by means of gravity from head works on River Musa to irrigate about 800 ha of land. Government intention for initiating the scheme in 1978 was that double cropping of rice would be encouraged. Since then the yield of major crops are low owing to poor water control, low soil fertility and toxicity [6].

The main focus of this paper is to assess the water quality and identification of pollutants sources flowing into River musa, Bida using Canadian Council Water Quality Index and Principal Components Analysis (PCA).

2. Methodology

2.1 Study Area

River Musa is the major river of economic purposes that flows in Bida town, Niger state. This river flows in south-east, which covers an area of about 17.5 km$^2$ and lies 60 00’ -60 15’E; 90 2’ -90 5’ N. Bida has annual rainfall ranges from 1000 mm to 1200 mm with a marked rainy season from April to October. Average annual air temperature is 27°C. The river is use for economic and agricultural importance in the town. It served the people of Bida and its environs for different activities (irrigation, fishing and household usage), the map of the study area is shown in figure 1.

2.2 Sampling and Analytical Technique

Samples were collected from five locations and were analysed for Electrical Conductivity (EC), pH, Dissolved Oxygen (DO), temperature, Sodium (Na), Potassium (K), Calcium (Ca), Magnesium (Mg), Sodium Adsorption Ratio (SAR), Ammonia (NH4), chloride (Cl$^-$), Fe and Mn. The electrical Conductivity (EC), PH, Dissolved Oxygen, turbidity and temperature were analysed using conductivity meter, pH meter, Dissolved Oxygen meter, Thermometer and Turbidity meter respectively. Total Dissolved Solid, Ammonia and Chloride were determined using American Public Health Association [7] standard recommended. Calcium, Sodium, Potassium and Magnesium were measured by [8]. Iron and Manganese were determined with the aid of Atomic Absorption Spectrometer (AAS). Iron and Manganese were evaluated by digesting water samples using concentrated of nitric acid HNO3 and concentration of Iron (Fe), and Manganese (M) with S series atomic absorption spectrophotometer (AAS).

2.3 Determination of Water Quality Index and Principal Component Analysis (PCA)

The WQI determination using CCME consist of Scope (F1), Frequency (F2) and Amplitude (F3) factors [9]. They were calculated for five locations.

Scope are those water quality parameters which there objectives were tested during the time period of index determination.

$$F1 = \left( \frac{\text{Number of failed variables}}{\text{Total number of variables}} \right) \times 100$$

Frequency (F2) is the percentage of individual tests that do not meet the standards (failed tests):

$$F2 = \left( \frac{\text{Number of failed test}}{\text{Total number of test}} \right) \times 100$$

3. Amplitude (F3)

This measures the amplitude. It was determined in three steps:
Step i: Evaluation of Excursion in which failed tests greater than objective involves the use of formula below

$$\text{Excursions} = \left( \frac{\text{Failed Test Value}}{\text{Objective}} \right) - 1$$  \hspace{1cm} (3i)

Step ii: Calculation of Excursion in which failed tests less than objective was calculated as below

$$\text{Excursions} = \left( \frac{\text{Objective}}{\text{Failed Test Value}} \right) - 1$$  \hspace{1cm} (3ii)

Step iii: Determination of Normalized Sum of Excursions (nse)

$$\text{Nse} = \frac{\sum_{k=1}^{n} (\text{Excursions})}{\text{Total number of tests}}$$  \hspace{1cm} (4)

$$F3 = \left( \frac{nse}{0.01nse+0.01} \right)$$  \hspace{1cm} (5)

The CWQI was finally evaluated as:

$$\text{WQI} = 100 - \sqrt{F1^2 + F2^2 + F3^2}$$  \hspace{1cm} (6)

Table 2. shows the Water Quality ranking System

| Rank       | Water Quality ranking system |
|------------|------------------------------|
| Poor       | 0-44                         |
| Marginal   | 45-64.9                      |
| Fair       | 65-79.9                      |
| Good       | 80-94.9                      |
| Excellent  | 95 – 100                     |

Source: CCMEWQI Categorizations Schema [10]

According to [11] description of water ranking system for different objectives are interpreted as follows;

Excellent: Water quality satisfies all criteria for different purpose.
Good: Water quality rarely disagree criteria for different purpose.
Fair: Water quality sometimes disagree criteria for different purpose.
Marginal: Water quality often violates criteria for different purpose.
Poor: Water quality do not satisfies any criteria for different purpose.

The Principal Component Analysis (PCA) was used in the determination of various pollutants sources. Twelve water variables namely; temperature, Total Dissolve Solids (TDS), PH, Magnesium, Sodium, Calcium, Potassium, Sodium adsorption ratio (SAR), Ammonia (NH4), chloride (Cl-) Fe and Mn were used in PCA to know major sources of water pollution flowing into river. The followings procedures were used in PCA on data.
2.4 Computation of Factor Descriptive Statistics

This was determined by entering the selected variables and then transferred into the right box on the software. Before PCA on data set, measure of sampling adequacy (KMO) and Bartlett’s test of sphericity were performed on data set. This is to ascertain the use of PCA on data according to [12]. The KMO was used to measure the adequacy of samples in order to point out the strength of relationship between the water variables. According to Shrestha and Kazama [13] high values of 0.5 and above are considerably required and low values below it were considered unrequired. Bartlett’s test of sphericity was used to check the status of correlation matrix. If the matrix determined is identity type, it means that variables are not related and therefore, PCA cannot use on data. MSA was used to evaluate each variable through evaluation of anti-image matrices and KMO for all variables. This was achieved as shown in figure 2.
2.5 Computation of Communalities
This was evaluated by communalities estimation table (estimation of the portion of variability in each variable that was shared with other, in factor analysis). According to Peter [14] the communalities greater than 0.2 are acceptable.

2.6 Extraction of Eigen Values and Related Loading on Components
(PCA) was applied on interrelated large set of data variables which were reduced into smaller set of new variables known as Principal Components with the loss of small information of original variables. New variables of (PCS) were further subjected to varimax rotation option and new factor known as varimax factors produced by reducing the size of original data information to identify the important known variables for easy interpretation [15]. With the grouping of the important variables in terms of component loading, the source identification was actualized. Factor loadings values greater than 0.75 were considered as strong significant factor, factor values loadings in the range of 0.75-0.5 were considered to be moderate and factor loadings values in the range of 0.5-0.3 were considered to be weak loadings as it was reported by [13]. Component loadings above 0.6 were considered in reporting of the most significance variables as pointed out by Nazireet al. [16]. Eigen values measures amount of variance in total sample accounted for by each component. Components with eigen values greater than one were retained for the river water variation interpretation.

3. Result and Discussions

3.1 Results
Scopes (F1) Values, Frequencies (F2) Values and Amplitudes (F3) Values with their corresponding water quality index for Rainy season and Dry season. These results were presented in table 1. The water quality index values of River Musa during the rainy season at the five locations (Edokota mini bridge, Musa mini bridge, Bida Minna mini bridge, Ciriko mini bridge and Army barrack mini bridge) shown in table 1 were 78.08, 78.15, 83.89, 54.41 and 45.50. The values revealed that the water quality ranking system of the river
during the rainy season is fair at Edokota station and Musa mini bridge stations, Good at Bida/ Minna station, Marginal at Ciriko and Army barrack stations.

The water quality index values of the river during the dry season at the locations (Edokota mini bridge, Musa mini bridge, Bida/ Minna mini bridge, Ciriko mini bridge and Army barrack mini bridge) as shown in table 3.1 were 70.86, 67.55, 45.48, 39.63 and 57.63 respectively. The water quality index values during dry season were relatively small compared to wet season values except at army barrack station; but respective ranks remained the same at Edokota and Musa stations, at Bida/Minna station the rank was marginal, at Ciriko station the WQI was poor and remained same at Army barrack station.

| Station       | Rainy season Frequency | Amplitude | WQI  | Dry Season Frequency | Amplitude | WQI  |
|---------------|------------------------|-----------|------|----------------------|-----------|------|
| Edokota       | 30                     | 23        | 3.46 | 78.08                | 30        | 20   | 31.97 | 70.86 |
| Musa          | 30                     | 20        | 11.57| 78.15                | 30        | 25   | 30.56 | 67.55 |
| Bida/Minna    | 20                     | 15        | 12.40| 83.89                | 40        | 27.5 | 83    | 45.48 |
| Cirriko       | 30                     | 27.5      | 67.67| 54.14                | 50        | 42.5 | 81.4  | 39.63 |
| Army Barrack  | 30                     | 20        | 87.24| 45.50                | 50        | 35   | 40.76 | 57.63 |

3.2 Mean Annual Water Quality Index Values
The mean annual index values of the river presented in table 2 (ranged 47.02 to 74.47) showed that the river water quality for irrigation can be ranked as fair at Edokota and Musa bridges. The reason is Edokota bridge station is the upstream of the river. The value reduced lower than 65 showing the water to be ranked as marginal for the remaining stations. The decrease in value below 65 in three of the five locations is an indication of different types of pollutants flowing in to river because of various human activities occurring at those stations. The annual average water quality index at sample point five still marginal as in sample point four, being the last sample point(downstream) and the scores was little higher than sample point four. This is because at sample point four various water tributaries with various pollutants meet here before moving to downstream point. This result is in line with the findings of Mohammed and John [17] which revealed that pollutants such as the discharge of abattoir waste at the bank of river Landzun cause its poor quality. Agricultural runoff wastes from farm lands near the stations of river could be responsible for the low quality of the river.

| Station       | Average Annual WQI | Remark  |
|---------------|---------------------|---------|
| Edokota       | 74.47               | Fair    |
| Musa          | 72.85               | Fair    |
| Bida/Minna    | 64.69               | Marginal|
| Cirriko       | 47.02               | Marginal|
| Army Barrack  | 51.56               | Marginal|

The marginal water quality rating observed of the river for irrigation purpose is determined by the analytical results of the selected physicochemical parameters (Table 3 and 4) chosen for WQI calculation.
et al., [19] which reported that nitrate and chloride are discharged into the river in the rainy season. The Sar on this component was due overland flow of agricultural wastes and industrial wastes getting into the river which have elevated due to the geological influence or pollution from anthropogenic sources near the river bank. VFI also has chloride loading which is one of the ions constituents of Total Dissolved Solids (TDS) and it could be a source of animal waste from abattoir industry and fertilizer waste from farm lands which are discharged into the river in the rainy season. This result finding is agreement with Animanshaun et al., [19] which reported that nitrate and chloride are discharged into the river in the rainy season. The

3.3 Principal Components Interpretation

From table 5 three varimax factors (VFs) with three different eigen values generated after their were subjected to varimax option rotation based on Kaiser standard using principal component analysis. In Kaiser Standard, only components with eigen values greater than one are considered and SPSS follows this rule [18]. Erosion waste, Municipal waste and Industrial waste were identified as pollutants flowing into the river.

The VF1 has an eigen value of 4.177 with a strong positive loading on Na, SAR, and moderate on CI. The VF1 also has strong negative loading on Fe and Mn. The river is being compromised by industrial effluents. Salinity and Conductivity of water increases with the increasing of dissolved ions. The loading of Na and Sar on this component was due overland flow of agricultural wastes and industrial wastes getting into the river which have elevated due to the geological influence or pollution from anthropogenic sources near the river bank. VFI also has chloride loading which is one of the ions constituents of Total Dissolved Solids (TDS) and it could be a source of animal waste from abattoir industry and fertilizer waste from farm lands which are discharged into the river in the rainy season. This result finding is agreement with Animanshaun et al., [19] which reported that nitrate and chloride are discharged into the river in the rainy season. The
probable source of heavy metals (Mn and Fe) is from mechanic villages surrounding the river whose wastes are disposed into water.

The VF2 has an eigen value of 2.830. It has a strong loading on K and Calcium and has a moderate loading on N-NH4. The loading of calcium was due to dispose of lime stone into river from nearby block industries surrounding the locations. The loading of potassium in the river was due to fertilizers run off from nearby farm lands. The N-NH4 nitrogen compound was found loading on this component. The major sources of nitrogen include dead plants and animals, soils fertilizer, municipal wastes [20].

(iii) The VF3 has an eigen value of 1.713 with moderate positive loading on Mg, T and PH. Magnesium and calcium are the two significant ions present in most rocks. The Mg and Ca are the major ions responsible for water hardness [21]. Total hardness of the water occurs as a result of calcium carbonate, chloride, sulphate, and nitrate of calcium and magnesium occurring in water. The rocks surrounded the water locations could be the source of hardness. The dissolution of gypsum and limestone from rocks indicates soil erosion processes.

| Parameters | Component 1 | Component 2 | Component 3 |
|------------|-------------|-------------|-------------|
| Sodium     | 0.888       | 0.170       | 0.151       |
| SAR        | 0.874       | -0.316      | 0.114       |
| Fe         | -0.782      | 0.294       | -0.043      |
| Cl         | 0.754       | 0.024       | -0.338      |
| Mn         | -0.711      | -0.179      | 0.108       |
| K          | 0.299       | 0.943       | 0.021       |
| Ca         | 0.169       | 0.864       | 0.159       |
| N-NH4      | -0.240      | 0.667       | 0.362       |
| Mg         | -0.293      | 0.150       | 0.680       |
| T          | 0.259       | 0.262       | 0.654       |
| PH         | 0.239       | -0.565      | 0.647       |
| TDS        | -0.479      | 0.520       | 0.564       |
| Eigen Values | 4.117       | 2.830       | 1.713       |

4. Conclusions
The study revealed that sources of water pollution resulting to the marginal ranking of River Musa were verified using Principal Component Analysis (PCA). The sources of pollution of river Musa are different in kind. Erosion, municipal and industrial wastes were identified. The water quality index of the River Musa was determined to be marginal using CCME index. The identified pollutants towards water quality variation are an indication of potential problems ahead. Without the appropriate policies or adaptive strategies in place, the use of the river Musa water for irrigation farming in the study area will continue to give low yielding of farm production.

References
[1] James P N 2008 Water quality for ecosystem and human health. 2nd Edition. United Nations Environmental programme monitoring system. (GEMS)
[2] Oluronntola T 1990 Battle over industrial effluents (Nigeria: Lagos) Wat. Res. J. 1 1.
[3] Adamu M and Ado A 2012 Application of Principal Component Analysis and Multiple Regression Models in Surface Water Quality Assessment. J. Environ. Earth Sci., 2 2 16 – 23.
[4] Santosh S M, Saroj K S and Bagwati B P 2015 Water pollution and Agriculture (Indian: State of Indian Agriculture ) (Electronic Materials) 3-10
[5] Krida C 2014 Assessing of irrigation water quality (Turkey: Cukurova University, Faculty of Agriculture- Adana) 367-337
[6] Olaniyan G O, Okunsanyan, B A and Ethan S 1999 Aspect of water management in the Badeggi irrigation scheme ,problems and prospects. *Irrig Res Proi for Nigeria* (Nwa Ed) 73-87.

[7] APHA 1995 Standard Methods for the Examination of Water and Wastewater. 19th Edition, American Public Health Association Inc., New York.

[8] APHA 1998 Standard Methods for the Examination of Water and Wastewater. 20th Edition, American Public Health Association, American Water Works Association and Water Environmental Federation, Washington DC.

[9] John M D 2006 Application and Tests of the Canadian Water Quality Index for Assessing Changes in Water Quality in Lakes and Rivers of Central North America, *Lake and Reservoir Management* 22 4, 308-20, DOI: 10.1080/0743814060935436.

[10] Canadian Council of Ministers of the Environment, (CCME) 2001 Canadian Water Quality Index Technical report 1.0 (Electronic Materials) p 3

[11] Canadian Council of Ministers of the Environment, (CCME) 2005 Canadian Water Quality Index Guidelines for the protection of aquatic life (Electronic Materials).

[12] Jiwen G, Guihua R, Wenjie M, Huafeng C, Shuyuan,W and Lamei C 2013 Water Quality Assessment of Gufu River in Three Gorges Reservoir (China) Using Multivariable Statistical Methods *J. Advance Food Sci. Tech*. 5 7: 908-20.

[13] Shrestha S, Kazama F 2007 Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. *Environ. Modelling & Software* 22: 464-475.

[14] Peter S 2016 Advice on Exploratory Factor Analysis (Electronic Materials) 1-9

[15] Ul-Saufie A Z, Yahya A and Ramli N A 2011 Improving Multiple Linear Regression Model Using Principal Component analysis for predicting Pm10 Concentration In SeberangPrai, Pulau Pinang Inter. *J. Env. Sci.* 2 2.

[16] Nazir M, Adem O and Suleyman M 1999 Interpretation of Water Quality Data by Principal Components Analysis. *J. Eng. Environ. Sci.* 23:19-26.

[17] Mohammed S and John J M 2012 Impact of Abattoir Effluent on River Landzuun, Bida, Nigeria. *J.Chem. Sci.* 2 (Nw 1.21) 132-136.

[18] Nasir M F, Abdul Z M, Juahir H, Hussain H, Zain S M, and Ramli 2012 Application of receptor models on water quality water data in source apportionment in Kuantan River Basin *J. Enviro. Health Sci. Eng.* 9 18: 1-2.

[19] Animanshan I M, Busari M B and Mohammed T B 2016 Receptor Modelling Application on Surface Water Quality and Source Apportionment. *Int. J. of Scie Res. In agric. Scie.* 3 1 001-10.

[20] Kitt F 2000 Water Quality and Monitoring by a river is a report card for its watershed (Alan Levere Connecticut Department of Environmental Protection)

[21] Chapman D 1996 Water Quality assessments, a guide to the use of biota, sediments and water in environmental monitoring 2nd edition, E and FN SPON, London.