Multiple Linear Regression (MLR) Model: A Tool for Water Quality Interpretation

Ogbozige, F. J[1]*., Toko, M. A[2] and Arawo, C. C[3]

[1]Department of Civil Engineering, Federal University Otuoke, Nigeria (*engr.ogbozige@gmail.com).

[2]Department of Water Resources and Environmental Engineering, Ahmadu Bello University, Zaria, Nigeria.

[3]Department of Industrial & Production Engineering, University of Ibadan, Nigeria.

ABSTRACT

The lack of standard water analysis equipment as well as inadequate trained personnel especially in the developing countries has discouraged many researchers in such countries to execute water quality researches. Hence, this paper presents developed mathematical relationship among some physicochemical parameters in order to aid the determination of the concentrations of certain parameters with the use of minimal equipment. This was achieved by weekly analyzing 7 physicochemical parameters of two sources of potable water (tap water and borehole water) stored in different containers for a period of 6 weeks using standard methods. The storage containers used were black plastic tank, blue plastic tank, green plastic tank, coated steel metal tank, uncoated steel metal tank and clay pot. The parameters examined were turbidity, electrical conductivity (EC), pH, alkalinity, chloride ion (Cl⁻), dissolved oxygen (DO) and total hardness. Results showed that the relationship between electrical conductivity (EC), alkalinity (Alk), total hardness (TH) and chloride ion (Cl⁻) is given as; $EC = -224.8066493 + 6.244028022(Alk) + 0.28204735(TH) + 0.000518108(Cl^-)$. A programing language was written on the models using Visual Basic.Net (VB.Net) version 2018.

Keywords: Water, Physicochemical, Parameters, Function, Equation.

1. INTRODUCTION

Due to the presence of impurities in the global water resources, only a very few proportion of the world's available water is potable or fit for human consumption (Andrew, 2014; Nala and Jagals, 2013; and Ochekpe, 2011). Knowing the concentrations of the physicochemical and bacteriological parameters of a given sample of water is very important because, comparing the known concentrations with the permissible limits set by regulatory bodies will help in deciding whether such water is potable or not. However, the exact concentration of water quality parameters can only be known by thorough analysis of the water with standard equipment as well as reagents, which could be expensive, time consuming and risky. Hence, it is necessary to develop a means of reducing the time usually spent in analyzing water quality as well as the cost incurred. This could be achieved by building an equation (mathematical model) out of previous studies that will act as a representation of reality without interfering directly with the quality of the result.

A mathematical model will actually reduce the risk, cost and also the time spent in analyzing the quality of water samples. A well written programing language of the model will further reduce the time spent when installed in a computer.

The multiple linear regression (MLR) model has been reported to be very useful in modelling factors influencing hydrological parameters such as infiltration and evapotranspiration rates (Patle et al., 2018; Saito et al., 2016; Helder et al., 2016; Rashidi et al., 2014; Shirgure, 2012). Many researchers have used MLR model successfully to determine factors influencing groundwater flow (e.g. Yan et al., 2018; Zomlot et al., 2015; Sahoo and Jha, 2013) and for checking water quality (Mustafa and Fuaad, 2019; Salam et al., 2018; Magda and Marco, 2018; Chen and Liu, 2015; Ahamad et al., 2015; Narayanan et al., 2015). However, the efficacy of this model for potable water quality is not yet studied. Hence, this paper presents the results of the investigation carried out to check the possibility of using MLR model in understanding the relationship between the quality parameters of potable water.

2. METHODOLOGY

Tap water obtained from the water treatment plant of the Ahmadu Bello University water works, Zaria (11° 8'17.43"N, 7°39'29.43"E) were filled into disinfected tap-fitted tanks made from coated steel metal, uncoated steel metal, black plastic, green plastic, blue plastic and clay pot having volume of 25L each. Another set of these storage tanks were used in storing borehole water obtained at No. 1 Basawa Road, Zaria (11° 9'34.96"N, 7°38'59.86"E). Water samples were collected from the storage containers into properly labeled disinfected sample bottles through the taps fitted in the storage containers on a weekly basis (7 days interval). The water filled in the above mentioned containers were stored for a period of six weeks (42 days). This storage period and sampling frequency is in agreement with the reports of Ogbozige et al. (2018a); Michael (2015); Henry et al. (2014); and Akulari et al. (2014). The water quality parameters that were monitored includes turbidity, electrical conductivity (EC), pH, alkalinity, chloride ion (Cl⁻), dissolved oxygen (DO) and total hardness.

The water turbidity was assessed using a turbidimeter (HACH 2100N) made by HANNA LTD, England while pH was determined using a multi-parameter photometer (HI8320) made by HANNA LTD, England. A pocket-sized conductivity meter (TDS & EC hold, ±2%) made by Griffin Company, USA was used in determining electrical conductivity. On the other hand,

alkalinity and total hardness were analyzed by Acid-Base and EDTA titration methods respectively, while dissolved oxygen was analyzed by Winkler's method. Chloride ion was analyzed by argentometric method in the presence of potassium chromate indicator. All parameters were analyzed in line with the standard methods (APHA, 2012).

The relationship between the water quality parameters was determined through the regression tool for statistical analysis by employing the method of Multiple Linear Regression (MLR) to fit a relationship for the model. The justification for adopting MLR was based on past literatures related to water quality modelling such as Mustafa and Fuaad (2019); Salam et al. (2018); Magda and Marco (2018); Chen and Liu (2015); Ahamad et al. (2015); and Narayanan et al. (2015). The general MLR model is shown in equation (1) as reported in Rajib (2018).

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + e \tag{1}$$

Where, Y is the dependent variable, $X_1, X_2,\ldots, X_k$ are independent variables, $\beta_0, \beta_1, \beta_2,\ldots, \beta_k$ are the unknown parameters and e is the error associated with estimate of Y.

For n observations on $k+1$ variables, equation (1) transforms as shown in equation (2).

$$Y_i = \beta_0 + \sum_{j=1}^{k} \beta_j X_{i,j} + e_i \tag{2}$$

Where Y_i is the ith observation of the dependent variable and $X_{i,j}$ is the ith observation of the jth independent variable.

The concept of sum of squares of errors was applied on equation (2) to yield equation (3). Thereafter, the associated error e was minimized by partial differentiation of equation (3) with respect to β_0, β_1, $\beta_2,\ldots, \beta_k$, and equating the values to zero. The resulted equations were transformed into matrix form as shown in equation (4)

$$\sum_{i=1}^{n} e_i{}^2 = \sum_{i=1}^{n}\left(Y_i - \beta_0 - \sum_{j=1}^{p} \beta_j X_{i,j}\right)^2 \tag{3}$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} \begin{bmatrix} 1 & X_{1,1} & X_{1,2} & X_{1,3} & \cdots & X_{1,k} \\ 1 & X_{2,1} & X_{2,2} & X_{2,3} & \cdots & X_{2,k} \\ 1 & X_{3,1} & X_{3,2} & X_{3,3} & \cdots & X_{3,k} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{n.1} & X_{n,2} & X_{n,3} & \cdots & X_{n,k} \end{bmatrix} \tag{4}$$

Equation (4) was solved simultaneously and the obtained values of $\beta_0, \beta_1, \beta_2,\ldots, \beta_k$ were substituted in the general MLR model given in equation (1).

The flowchart for the model was drawn using CorelDraw X7 while the programming language for the model was written through Visual Basic.Net (VB.Net) version 2018. It was finally validated using JMP version 14.0.

3. RESULTS AND DISCUSSION

The laboratory results of the 7 physicochemical parameters monitored during the six (6) weeks retention period for both sources of water are shown in table 1.

Table 1. Variation in physicochemical parameters of water during storage period.

Retention Time (Week)	Type of Water and Storage Container	Turbidity (NTU)	EC (µS/cm)	pH	Alkalinity (mg/l CaCO₃)	Cl⁻ (mg/l)	DO (mg/l)	Total Hardness (mg/l CaCO₃)
0	Tap water before storage	2.000	118.99	6.6	54.01	21.04	1.5	27.61
	Borehole water before storage	1.091	707.02	7.0	126.13	160.01	1.3	138.07
1	BKP_t	2.08	91.01	7.0	46.00	12.53	2.0	23.04
	USM_t	16.07	97.98	7.4	47.99	15.00	2.0	24.00
	GRP_t	2.072	83.95	7.5	43.96	9.95	2.1	21.95
	CSM_t	1.191	105.04	7.0	51.00	16.25	2.1	24.36
	BLP_t	2.210	90.96	7.1	45.95	12.51	2.0	23.01
	CLP_t	1.053	126.04	7.0	57.50	23.44	2.7	24.98
	BKP_b	0.214	637.60	7.2	130.90	157.51	2.1	197.03
	USM_b	6.080	626.90	8.5	129.95	154.96	2.4	186.95
	GRP_b	0.209	616.05	7.3	129.01	152.58	2.4	177.50
	CSM_b	0.799	637.60	7.4	130.92	157.49	2.1	197.04
	BLP_b	0.211	644.03	7.4	131.04	160.00	1.9	205.02
	CLP_b	0.219	658.08	7.1	131.99	163.66	2.6	222.48
2	BKP_t	1.960	108.49	8.1	52.03	18.33	1.6	24.51
	USM_t	12.801	108.46	8.3	51.96	17.50	1.0	24.49
	GRP_t	1.731	118.98	8.2	56.00	19.98	1.4	24.95
	CSM_t	1.710	119.01	8.1	55.97	19.95	1.3	25.04
	BLP_t	1.904	125.96	8.2	57.54	22.46	1.1	25.00
	CLP_t	2.001	136.46	7.8	60.02	24.97	1.4	25.06
	BKP_b	0.197	734.96	7.1	135.96	185.04	1.8	399.97
	USM_b	8.310	665.01	8.7	132.01	164.97	1.1	232.52
	GRP_b	0.122	728.00	7.6	136.04	182.45	1.4	374.96
	CSM_b	0.142	720.95	7.2	135.02	179.92	1.3	345.02
	BLP_b	0.148	735.02	7.8	136.01	184.95	1.3	399.95
	CLP_b	0.127	664.97	7.4	132.00	165.03	1.4	232.48
3	BKP_t	1.830	136.45	7.7	59.98	24.95	1.6	25.05
	USM_t	9.490	108.39	7.8	52.01	17.55	1.4	24.45
	GRP_t	1.560	87.53	8.0	45.03	11.25	2.4	22.21
	CSM_t	1.410	136.51	7.6	61.06	24.99	1.7	25.09
	BLP_t	1.650	97.95	7.6	49.04	15.04	1.3	23.96
	CLP_t	0.342	118.89	7.2	55.96	19.94	1.4	24.98
	BKP_b	0.303	706.97	7.0	134.02	177.46	1.3	305.04
	USM_b	1.290	615.99	8.1	129.94	152.48	1.3	177.52
	GRP_b	0.317	717.53	7.2	135.03	179.96	1.3	333.04
	CSM_b	0.806	706.95	6.9	133.95	177.48	1.5	305.01

	BLP_b	0.110	734.99	7.6	135.99	185.03	1.5	400.00
	CLP_b	1.561	745.47	7.2	137.99	187.51	1.3	400.20
4	BKP_t	1.795	112.04	7.5	52.97	19.00	1.3	24.56
	USM_t	11.460	136.48	6.8	59.97	24.98	1.4	25.10
	GRP_t	1.880	122.54	7.7	56.21	21.25	1.7	24.95
	CSM_t	1.427	87.46	7.7	44.96	11.26	1.7	22.18
	BLP_t	1.920	98.02	7.6	48.02	14.98	1.3	23.95
	CLP_t	1.131	108.55	7.3	52.00	17.51	1.4	24.52
	BKP_b	1.957	714.04	6.9	135.01	178.75	1.6	325.04
	USM_b	8.011	714.02	7.8	134.95	178.79	0.8	324.95
	GRP_b	0.216	735.01	7.3	136.03	184.95	1.2	399.96
	CSM_b	0.523	713.98	6.8	135.00	178.74	1.3	325.00
	BLP_b	0.490	735.04	7.6	136.04	185.95	2.2	400.10
	CLP_b	0.635	749.03	7.1	137.00	188.75	1.6	439.80
5	BKP_t	1.955	105.04	6.0	51.97	16.24	1.1	24.44
	USM_t	13.447	111.97	6.2	56.03	18.74	1.2	24.61
	GRP_t	1.815	112.02	6.4	53.01	18.72	1.7	24.58
	CSM_t	1.300	87.52	6.7	45.04	11.25	1.4	22.16
	BLP_t	1.930	91.00	6.7	46.00	12.54	1.4	23.04
	CLP_t	0.696	122.52	6.8	56.14	21.26	2.1	25.00
	BKP_b	0.259	637.50	6.3	130.94	157.54	1.3	196.96
	USM_b	9.820	629.94	6.6	130.00	154.98	1.1	186.99
	GRP_b	0.263	616.08	6.7	128.95	152.53	1.6	177.45
	CSM_b	0.803	637.71	6.6	131.34	157.49	1.4	197.00
	BLP_b	0.661	644.02	6.6	130.97	159.95	1.6	204.98
	CLP_b	0.890	658.05	6.5	132.01	163.68	2.1	222.46
6	BKP_t	1.915	108.52	6.0	52.03	18.09	1.6	24.49
	USM_t	11.861	108.54	5.8	51.95	17.53	0.9	24.53
	GRP_t	1.808	119.00	6.7	56.04	19.95	1.8	24.95
	CSM_t	1.479	119.04	7.4	56.00	21.00	1.4	25.04
	BLP_t	1.926	126.00	6.6	57.52	22.49	1.7	24.95
	CLP_t	1.276	136.48	6.0	59.97	24.97	1.5	25.14
	BKP_b	0.206	735.03	6.4	136.00	184.95	1.4	400.12
	USM_b	9.450	665.00	6.1	131.95	165.02	2.1	232.51
	GRP_b	0.166	728.01	6.3	136.00	183.98	1.3	375.00
	CSM_b	0.471	721.00	6.5	134.99	179.95	1.1	345.03
	BLP_b	0.180	735.03	6.6	136.04	184.94	1.6	185.04
	CLP_b	0.173	665.02	6.2	132.01	166.01	1.4	232.45

Note: BKP_t = Tap water stored in a black plastic tank;

USM_t = Tap water stored in an uncoated steel metal tank;

GRP_t = Tap water stored in a green plastic tank;

CSM_t = Tap water stored in a coated steel metal tank;

BLP_t = Tap water stored in a blue plastic tank;

CLP_t = Tap water stored in a clay pot while BKP_b, USM_b, GRP_b, CSM_b, BLP_b and CLP_b represent borehole water stored in corresponding containers

3.1. Regression Analysis

Curves were plotted between the parameters of the results shown in table 1 in a linear graph, and it was observed that the electrical conductivity (EC) was a function of alkalinity (Alk), total hardness (TH) and chloride ion (Cl⁻) present in the water samples as represented in equation (5).

$$EC = f(Alk, TH, Cl^-) \tag{5}$$

In order to fit a relationship between the dependent variable (EC) and independent variables (Alk, TH and Cl⁻), equation (5) was transformed into equation (6) as;

$$EC = \beta_0 + \beta_1(Alk) + \beta_2(TH) + \beta_3(Cl^-) \tag{6}$$

Where, β_0, β_1, β_2 and β_3 are constants.

Assuming the line of best fit in equation (6) is associated with an error e, then by applying the concept of Sum of Squares of Errors (SSE) to equation (6), it resulted to equation (7) as follows.

$$SSE = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}[EC_i - \beta_0 - \beta_1(Alk_i) - \beta_2(TH_i) - \beta_3(Cl^-_i)]^2 \tag{7}$$

The error e, was minimized by differentiating equation (7) partially with respect to the constants, β_0, β_1, β_2 and β_3 and then equating to zero to produce equations (8), (9), (10) and (11) respectively. That is;

$$\frac{\partial \sum_{i=1}^{n} e_i^2}{\partial \beta_0} = 2\sum_{i=1}^{n}[EC_i - \beta_0 - \beta_1(Alk_i) - \beta_2(TH_i) - \beta_3(Cl^-_i)][-1] \quad = \quad 0 \tag{8}$$

$$\frac{\partial \sum_{i=1}^{n} e_i^2}{\partial \beta_1} = 2\sum_{i=1}^{n}[EC_i - \beta_0 - \beta_1(Alk_i) - \beta_2(TH_i) - \beta_3(Cl^-_i)][-Alk_i] = 0 \tag{9}$$

$$\frac{\partial \sum_{i=1}^{n} e_i^2}{\partial \beta_2} = 2\sum_{i=1}^{n}[EC_i - \beta_0 - \beta_1(Alk_i) - \beta_2(TH_i) - \beta_3(Cl^-_i)][-TH_i] = 0 \tag{10}$$

$$\frac{\partial \sum_{i=1}^{n} e_i^2}{\partial \beta_3} = 2\sum_{i=1}^{n}[EC_i - \beta_0 - \beta_1(Alk_i) - \beta_2(TH_i) - \beta_3(Cl^-_i)][-Cl^-_i] = 0 \tag{11}$$

Equations (8), (9), (10) and (11) were respectively rearranged to obtain equations (12), (13), (14) and (15) as follows:

$$n\beta_0 + \beta_1\sum_{i=1}^{n}(Alk_i) + \beta_2\sum_{i=1}^{n}(TH_i) + \beta_3\sum_{i=1}^{n}(Cl^-_i) = \sum_{i=1}^{n}(EC_i) \tag{12}$$

$$\beta_0\sum_{i=1}^{n}(Alk_i) + \beta_1\sum_{i=1}^{n}(Alk_i)^2 + \beta_2\sum_{i=1}^{n}[(Alk_i)(TH_i)] + \beta_3\sum_{i=1}^{n}[(Alk_i)(Cl^-_i)] = \sum_{i=1}^{n}[(Alk_i)(EC_i)] \tag{13}$$

$$\beta_0\sum_{i=1}^{n}(TH_i) + \beta_1\sum_{i=1}^{n}[(TH_i)(Alk_i)] + \beta_2\sum_{i=1}^{n}(TH_i)^2 + \beta_3\sum_{i=1}^{n}[(TH_i)(Cl^-_i)] = \sum_{i=1}^{n}[(TH_i)(EC_i)] \tag{14}$$

$$\beta_0\sum_{i=1}^{n}(Cl^-_i) + \beta_1\sum_{i=1}^{n}[(Cl^-_i)(Alk_i)] + \beta_2\sum_{i=1}^{n}[(Cl^-_i)(TH_i)] + \beta_3\sum_{i=1}^{n}(Cl^-_i)^2 = \sum_{i=1}^{n}[(Cl^-_i)(EC_i)] \tag{15}$$

The values of the summation (i.e. $\sum$) constants in equation (12) to (15) were estimated as;

$\sum_{i=1}^{n}(Alk_i) = 6887.97$, $\quad \sum_{i=1}^{n}(TH_i) = 11302.57$, $\quad \sum_{i=1}^{n}(Cl^-_i) = 7014.83$, $\quad \sum_{i=1}^{n}(EC_i) = 29564.48$,

$\sum_{i=1}^{n}(Alk_i)^2 = 761539.3263$, $\quad \sum_{i=1}^{n}[(Alk_i)(TH_i)] = 1441780.633$, $\quad \sum_{i=1}^{n}[(Alk_i)(Cl^-_i)] = 881948.7181$,

$\sum_{i=1}^{n}[(Alk_i)(EC_i)] = 3613718.788$, $\quad \sum_{i=1}^{n}(TH_i)^2 = 3241079.143$, $\quad \sum_{i=1}^{n}[(TH_i)(Cl^-_i)] = 1834879.596$,

$\sum_{i=1}^{n}[(TH_i)(EC_i)] = 7376714.233,$ $\sum_{i=1}^{n}(Cl^-_i)^2 = 11003624.99$ and   finally, $\sum_{i=1}^{n}[(Cl^-_i)(EC_i)] =$ 4453156.08.

The sample size $n = 74$ (Table 1) hence, the numerical values of the summation constants were substituted in equation (12) to (15) in matrix form as shown in equation (16).

$$\begin{bmatrix} 74 & 6887.97 & 11302.57 & 7014.83 \\ 6887.97 & 761539.3263 & 1441780.633 & 881948.7181 \\ 11302.57 & 1441780.633 & 3241079.143 & 1834879.596 \\ 7014.83 & 881948.7181 & 1834879.596 & 11003624.99 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 29564.48 \\ 3613718.788 \\ 7376714.233 \\ 4453156.08 \end{bmatrix} \tag{16}$$

The 4×4 matrix in equation (16) was reduced to Echelon or upper triangular matrix shown in equation (17) using R_1, R_2, R_3 and R_4 to represent first, second, third and fourth row respectively.

$$\begin{bmatrix} 74 & 6887.97 & 11302.57 & 7014.83 \\ 0 & 120402.4246 & 389729.7805 & 229003.6019 \\ 0 & 389729.7808 & 1514753.622 & 763452.4731 \\ 0 & 229003.602 & 763452.4729 & 10338654.18 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 29564.48 \\ 861837.0135 \\ 2861111.468 \\ 1650591.198 \end{bmatrix} \begin{array}{l} \longrightarrow R_2 - 93.08066756R_1 \\ \longrightarrow R_3 - 152.7374324R_1 \\ \longrightarrow R_4 - 94.795R_1 \end{array}$$

$$\begin{bmatrix} 74 & 6887.97 & 11302.57 & 7014.83 \\ 0 & 120402.4246 & 389729.7805 & 229003.6019 \\ 0 & 0 & 253239.9737 & 22192.28782 \\ 0 & 0 & 22192.28803 & 9903092.771 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 29564.48 \\ 861837.0135 \\ 71437.16155 \\ 11390.15173 \end{bmatrix} \begin{array}{l} \longrightarrow R_3 - 3.236893128R_2 \\ \longrightarrow R_4 - 1.90198497R_2 \end{array}$$

$$\begin{bmatrix} 74 & 6887.97 & 11302.57 & 7014.83 \\ 0 & 120402.4246 & 389729.7805 & 229003.6019 \\ 0 & 0 & 253239.9737 & 22192.28782 \\ 0 & 0 & 0 & 9903092.771 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 29564.48 \\ 861837.0135 \\ 71437.16155 \\ 5129.868091 \end{bmatrix} \tag{17}$$
$$\longrightarrow R_4 - 0.087633432R_3$$

The constants β_0, β_1, β_2 and β_3 were determined by applying the Gauss – Jordan elimination (backward substitution) method in equation (17) to yield equation (18) to (21).

$$74\beta_0 + 6887.97\beta_1 + 11302.57\beta_2 + 7014.83\beta_3 = 29564.48 \tag{18}$$
$$120402.4246\beta_1 + 389729.7805\beta_2 + 229003.6019\beta_3 = 861837.0135 \tag{19}$$
$$253239.9737\beta_2 + 22192.28782\beta_3 = 71437.16155 \tag{20}$$
$$9901147.985\beta_3 = 5129.868091 \tag{21}$$

From (21), $\boldsymbol{\beta_3} = \frac{5129.868091}{9901147.98} = \mathbf{0.000518108}$

From (20), $\boldsymbol{\beta_2} = \frac{71437.16155 - 22192.28782(0.000518108)}{253239.9737} = \mathbf{0.28204735}$

From (19), $\boldsymbol{\beta_1} = \frac{861837.0135 - 389729.7805(0.28204735) - 229003.6019(0.000518108)}{120402.4246} = \mathbf{6.244028022.}$ Lastly,

From (18), $\boldsymbol{\beta_0} = \frac{29564.48 - 6887.97(6.244028022) - 11302.57(0.28204735) - 7014.83(0.000518108)}{74} = \mathbf{-224.8066493}$

Substituting the values of constants β_0, β_1, β_2 and β_3 into equation (6) yields equation (22).

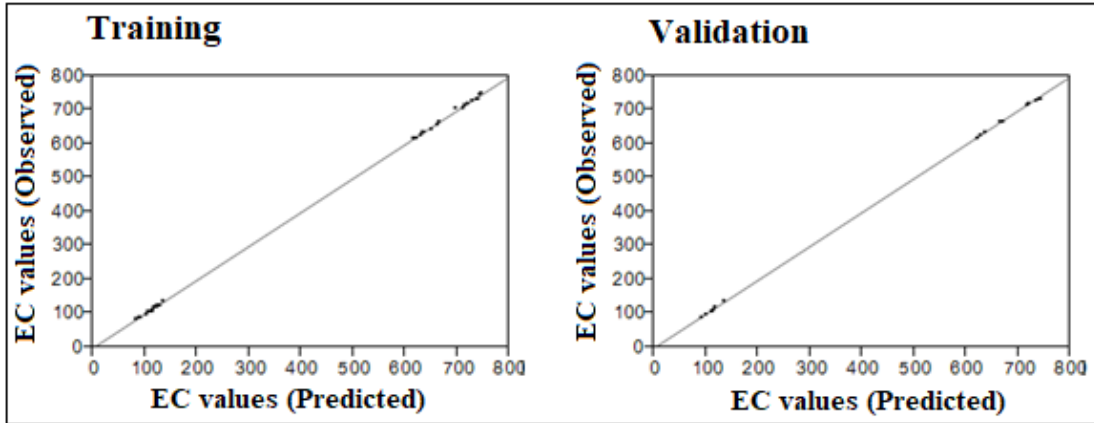$$\boldsymbol{EC = -224.8066493 + 6.244028022(Alk) + 0.28204735(TH) + 0.000518108(Cl^-)} \tag{22}$$

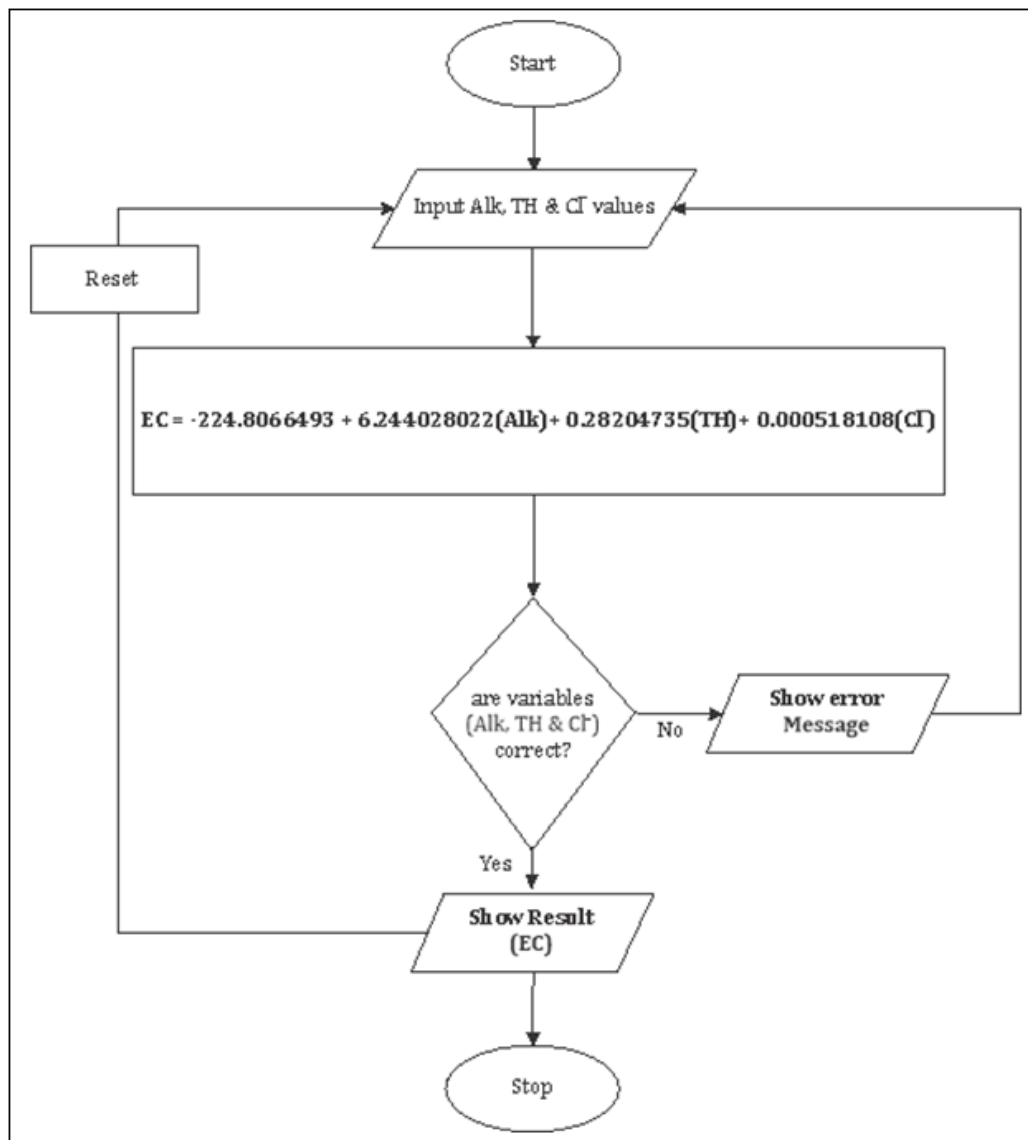Figure 1. Comparism of EC (µS/cm) between observed and predicted values.

Figure 2. Flowchart of computer program.

Equation (22) is the general model relating the electrical conductivity (EC), alkalinity (Alk), total hardness (TH) and chloride ion (Cl⁻) in the water samples used and it coincides favorably with the experimental values. This was validated by plotting actual values obtained in the field against the predicted values obtained from the regression equation using the same scale via JMP version 14.0 as shown in figure 1. Since most of the plotted points fell on the diagonal of figure 1, it suggests that the regression equation is highly reliable.

A programming langauge was written for equation (22) using Visual Basic. Net (VB. Net) version 2018 and the flowchart of the model is shown in figure 2.

3.2. Trend in Water Quality During Retention Period

The turbidity of both water sources (tap water and borehole water) in all the storage containers during the retention period were within the WHO permissible limit (5.00 NTU) except for water stored in uncoated steel metal tanks (USM_t and USM_b). Likewise, the pH of all the water samples were within the WHO limit (6.5 – 8.5) in the first four weeks of storage thereafter, little deviations from the permissible limit occurred especially for the water samples stored in uncoated steel metal tanks. The electrical conductivity, chloride ion and alkalinity of the water samples throughout the retention period were within the WHO standards (750 µS/cm, 200 mg/l and 250 mg/l respectively) irrespective of the type of water and storage container. Dissolved oxygen (DO) is an important water quality parameter as it is a respiratory gas. The DO content in all the water samples improved (increased) during the first week of storage and afterward, it fluctuated between values higher and lower than the permissible limit (1.5 mg/l). Notwithstanding, the minimum concentration of DO (1.0 mg/l) was recorded on the second week of retention in tap water stored in uncoated steel metal tank (USM_t) while the maximum concentration (2.7 mg/l) was noted on the first week in tap water stored in clay pot (CLP_t). The variations in hardness (total) level of the water samples during the retention period were not much. Nevertheless, water samples drawn from storage containers filled with borehole water recorded higher concentrations (308.77 mg/l to 661 mg/l) than samples drawn from storage containers filled with tap water (77.44 mg/l to 127.60 mg/l). Water hardness above 200 mg/l may cause scale deposition in treatment works and pipe distribution system as well as excessive soap consumption (Ogbozige et al., 2018b).

4. CONCLUSION

Based on the results obtained in this research, it could be concluded that the relationship between electrical conductivity (EC), alkalinity (Alk), total hardness (TH) and chloride ion (Cl$^-$) is given as; EC = -224.8066493 + 6.244028022(Alk) + 0.28204735(TH) + 0.000518108(Cl$^-$). Hence, prospective researchers on water quality of the selected sampling points could make use of the model while researchers on water samples drawn from other sources should calibrate the model before making use of it.

5. ACKNOLEDGEMENTS

The role played by Prof. O.J. Mudiare in writing this paper is well appreciated.

6. CONFLICT OF INTERESTS

There is not conflict of Interests.

7. REFERENCE

Ahamad, K.U., Nazary, B.K., Mudoi, P., Rani, R., Bharati, V & Singh, H. 2015. Application of Regression Analysis for Surface Water Quality Modeling. *IOSR Journal of Mechanical and Civil Engineering,* **1(1)**: 65-70.

Akulari, D.D., Steven, T & Gbenga, A.O. 2014. Effects of Short Term Storage on Water Quality. *Nigerian Journal of Science,* **30(2)**: 147-152.

Andrew, F. 2014. Water Quality Deterioration: A Study of Household Drinking Water Quality in Rural Honduras. *International Journal of Environmental Health Research,* **14(4)**: 273-284.

APHA. (2012). Standard Methods for Examination of Water and Wastewater (25th Edition). American Public Health Association, Washington DC, USA.

Chen, W & Liu, W. 2015. Water Quality Modelling in Reservoirs Using Multivariate Linear Regression and Two Neural Network Models. *Advances in Artificial Neural Systems,* **1(1)**: 31-40.

Helder J. F., Marconio S. S., Jório, B. C & Maria H. C. 2016. Modeling of reference evapotranspiration by multiple linear regression. *Journal of Hyperspectral Remote Sensing,* **6(2)**: 44-58.

Henry, C.I., Chioma, D.E & Clifford, M.K. 2014. Water Quality Behavior Over Time. *Journal of Tropical Scientific Research,* **5(10)**: 38-50.

Magda, M & Marco, C. 2018. A Time Series Model Comparison for Monitoring and Forecasting Water Quality Variables. *Hydrology,* **5(37)**: 1-20.

Michael, F.P. 2015. Effects of Containers and Storage Conditions on Bacteriological Quality of Borehole Water. *Nigeria Journal of Pure and Applied Science,* **17(2)**: 1223-1226.

Mustafa, A & Fuaad, A. 2019. Modeling Water Quality Parameters Using Data-Driven Models, a Case Study Abu-Ziriq Marsh in South of Iraq. *Hydrology,* **6(24)**: 1 – 17.

Nala, N & Jagals, P. 2013. The Effect of a Water-Hygiene Educational Programme on the Microbiological Quality of Container-Stored Water in Households," *Water SA,* **29(2)**: 171-176 (DOI: 10.4314/wsa.v29i2.4852).

Narayanan, C. V., Dileep Kumar, P.G & Ammad, K. K. 2015. Statistical Analysis of Quality of Water in Various Water Shed for Kozhikode City, Kerala, India. *Aquatic Procedia,* **4**: 1078-1085 (doi: 10.1016/j.aqpro.2015.02.136).

Ochekpe, S.R. 2011. Potability of Nigerian Sachet Water. Daily Trust News Paper, Kaduna, Nigeria. December 6th, pp 8-9.

Ogbozige, F.J., Ibrahim, F.B & Adie, D.B. 2018a. Effect of Storage Time and Container Material on Potable Water Quality. *Ife Journal of Science and Technology*, **1(2)**: 59-71.

Ogbozige, F.J., Ibrahim F.B & Adie, D.B. 2018b. Drinkable Water Stored in Hot Climates: Interactions among Water Quality Parameters. *Arid Zone Journal of Engineering, Technology and Environment,* **14(3)**: 381-390.

Patle, G.T., Sikar, T.T., Rawat, K.S & Singh, S.K. 2018. Estimation of Infiltration Rate from Soil Properties Using Regression Model for Cultivated Land. *Geology, Ecology and Landscapes,* **1(1)**:1-14.

Rajib, M. 2018. *Statistical Methods in Hydrology and Hydroclimatology*. Springer Transactions in Civil and Environmental Engineering, 444p (DOI: 10.1007/978-981-10-8779-0).

Rashidi, M., Ahmadbeyki, A & Hajiaghaei, A. 2014. Prediction of Soil Infiltration Rate Based on Some Physical Properties of Soil. *American-Eurasian J. Agriculture and Environmental Science,* **14 (12)**: 1359-1367.

Sahoo, S & Jha, M.K. 2013. Groundwater-Level Prediction Using Multiple Linear Regression and Artificial Neural Network Techniques: A Comparative Assessment. *Hydrogeology Journal* **21(8)**: 1865–1887 (DOI 10.1007/s10040-013-1029-5).

Saito, T., Yasuda, H., Suganuma, H., Inosako, K., Abe, Y & Kojima, T. 2016. Predicting Soil Infiltration and Horizon Thickness for a Large-Scale Water Balance Model in an Arid Environment. *Water,* **8(3)** 96: 1-15 (https://doi.org/10.3390/w8030096).

Salam, H.E., Abed, S.A & Khadum, S.A. 2018. Predicting the Tigris Water Quality within Baghdad, Iraq by Using Water Quality Index and Regression Analysis. *Environmental Technology and Innovation,* **11(1)**: 390-398.

Shirgure, P.S. 2012. Evaporation Modeling with Multiple Linear Regression Techniques – A Review. *Scientific Journal of Review,* **1(5)**: 170-182.

Yan, S., Yu, S., Wu, Y., Pan, D & Dong, J. 2018. Understanding Groundwater Table Using Statistical Model. *Water Science and Engineering,* **11(1)**: 1-7.

Zomlot, Z., Verbeiren, B., Huysmans, M & Batelaan, O. 2015. Spatial Distribution of Groundwater Recharge and Base Flow: Assessment of Controlling Factors. *Journal of Hydrology: Regional Studies,* **4(4)**: 349-368.