An Application of Geostatistics to Analysis of Water Quality Parameters in Rivers and Streams in Niger State, Nigeria

Isah Audu[1], Abdullahi Usman[2]

[1]Department of Mathematics & Statistics, School of Physical Sciences, Federal University of Technology, Minna, Nigeria
[2]Academic Planning Unit, Vice Chancellor's Office, Federal University of Technology, Minna, Nigeria

Email address:
aisah@futminna.edu.ng (A. Isah), u.abdullahi@futminna.edu.ng (U. Abdullahi)

Abstract: Assessment of surface water quality using multivariate statistical techniques does not incorporate the spatial locations of data into their defining computations. Information on spatial continuity of surface water concentrations can help in identifying the magnitude of contamination by runoff and anthropogenic pollutions. In the present study, spatial behavior of five (5) surface water quality parameters of some rivers/streams in Niger State of Nigeria was studied using R geostatistical package *gstat*, in conjunction with packages *sp, rgdal, spatstat* and *maptools*. The variograms and ordinary krigged spatial maps were generated for rainy and dry seasons. The characteristics of the best variable models; range; sill and nugget effects of each parameter were obtained. The variogram analysis indicated a high spatial coherence for E.co, Mg and TDS, whereas TCo and TH indicated a low spatial coherence. The nugget to sill ratios of experimental and linear fitted variogram models in all cases were less than 0.25 indicating that the rivers/streams water level has strong spatial coherence in both seasons. This result shows that linear model is the best for both seasons. Krigged spatial variability maps revealed that an average range of 48km variograms for dry season changes more rapidly than it does in rainy season with an average range of 4.3 km and R^2 values of 0.80 to 0.92.

Keywords: Kriging, Predictions, Experimental Variogram, Nugget, Water Parameters

1. Introduction

Niger State is underlain by sedimentary and basement complex rocks which have different capacities of retaining water all year round [23, 24]. Niger State like the rest of Nigeria and other tropical lands has two seasons, the dry and rainy seasons. The rainy season is influenced by the south west wind or the tropical maritime air mass. This wind involves Nigeria between February and June, depending on the location. The dry season is accompanied by a dust laden air mass from the Sahara desert, locally known as the harmattan. During the rainy season, the whole area is often flooded with water while in dry season some of the rivers do dry up. This gives rise to difficulties in accessing adequate safe quality water supply. With the increase in population, the situation of scrambling for domestic water is aggravated. Most of the medium-sized towns have been encountering similar problem of lacking adequate quality water supply since 1980. [20] is of the view that access to portable water in Niger State has been on continuous decrease since 1980s. On the average, less than 20% of the inhabitants of the study area currently have access to portable water.

Water quality is the main factor controlling healthy and diseased states in both humans and animals. Surface-water sources may be extremely difficult to survey adequately, particularly in remote rural areas and where land-use patterns are changing rapidly. Not only may there be daily and seasonal changes in flow to consider but, in addition, variations in physical, chemical, and microbiological characteristics necessitate analysis throughout the year to take account of the effect of changes in rainfall patterns [35]. Surface water quality is an essential component of the natural environment and a matter of serious concern today. The variations of water quality are essentially the combination of both anthropogenic and natural contributions. In general, the anthropogenic discharges constitute a constant source of pollution, whereas surface runoff is a seasonal phenomenon which is affected by climate within the water catchment basin [1]. Among them, because of the intensive human activities, the anthropogenic inputs from a variety of sources are commonly the primary factors affecting the water quality of most rivers, lakes, estuaries, and seas, especially for those close to highly urbanized regions.

Research findings indeed reveal deteriorating surface and ground water quality in Nigeria, Uganda and India due to chemical and biological pollution and seasonal changes among others [11, 17, 25, 28]. As water quality issues become more serious and widespread, the need for water quality monitoring as an important component of health promotion strategy in the developing countries cannot be overemphasized.

Recently, a considerable number of researchers have shown an increased interest in the use of multivariate statistical tools and geostatistical techniques to achieve a sustainable exploitation of water resources [2, 4]. The combined use of multivariate statistics and geostatistical techniques provide the identification of possible sources that affect water environmental systems and offer a valuable tool for reliable management of water resources as well as rapid solution for pollution issues [16, 30]. Moreover, multivariate statistical tools and geostatistical techniques also provide a way of handling large data sets in the environmental studies [5, 29]. [14] applied multivariate methods for assessment of variations in rivers/streams water quality in Niger State of Nigeria. [10] used multivariate statistical analysis for the assessment of water quality changes in a Karstic aquifer in the rainy (winter) and dry (summer) seasons.

Geostatistics have been applied in different fields of study such as water quality [1, 19, 21]. They applied ordinary kriging (OK) to determine the spatial distribution of water quality parameters in urban areas in Konya, Turkey.

Geostatistical analysis provides a series of statistical models and tools for spatial data exploration and surface generation of groundwater quality [27]. Hence, the objective of the present study is to provide an overview for the most significant parameters identified in [14] such as (escherichia coli(E.co), magnesium(Mg), total coliform(TCo), total dissolved solid (TDS), and total hardness(TH)) and integrate the multivariate statistical analysis results to determine the spatial continuity of these river water quality parameters in the study area using geostatistical techniques.

2. Materials and Methods

2.1. Study Area

Niger State of Nigeria lies between Latitudes 8^0 20' N and 11^0 30' N and Longitudes 3^0 30' E and 7^0 20' E with twenty-five local government area councils. The state is endowed with some large rivers, but with no major water bodies. A sizeable amount of rainwater is lost through percolation to the ground while bulk of it flows as runoff into rivers and streams with some of it lost to the atmosphere by evapo-transpiration [22]. There are two major categories of settlements in the state, urban and rural settlements. The samples were taken from rural settlements. The rural dwellers engage basically in agriculture. Two of the hydro-electric power stations in the state are located within the sampled locations.
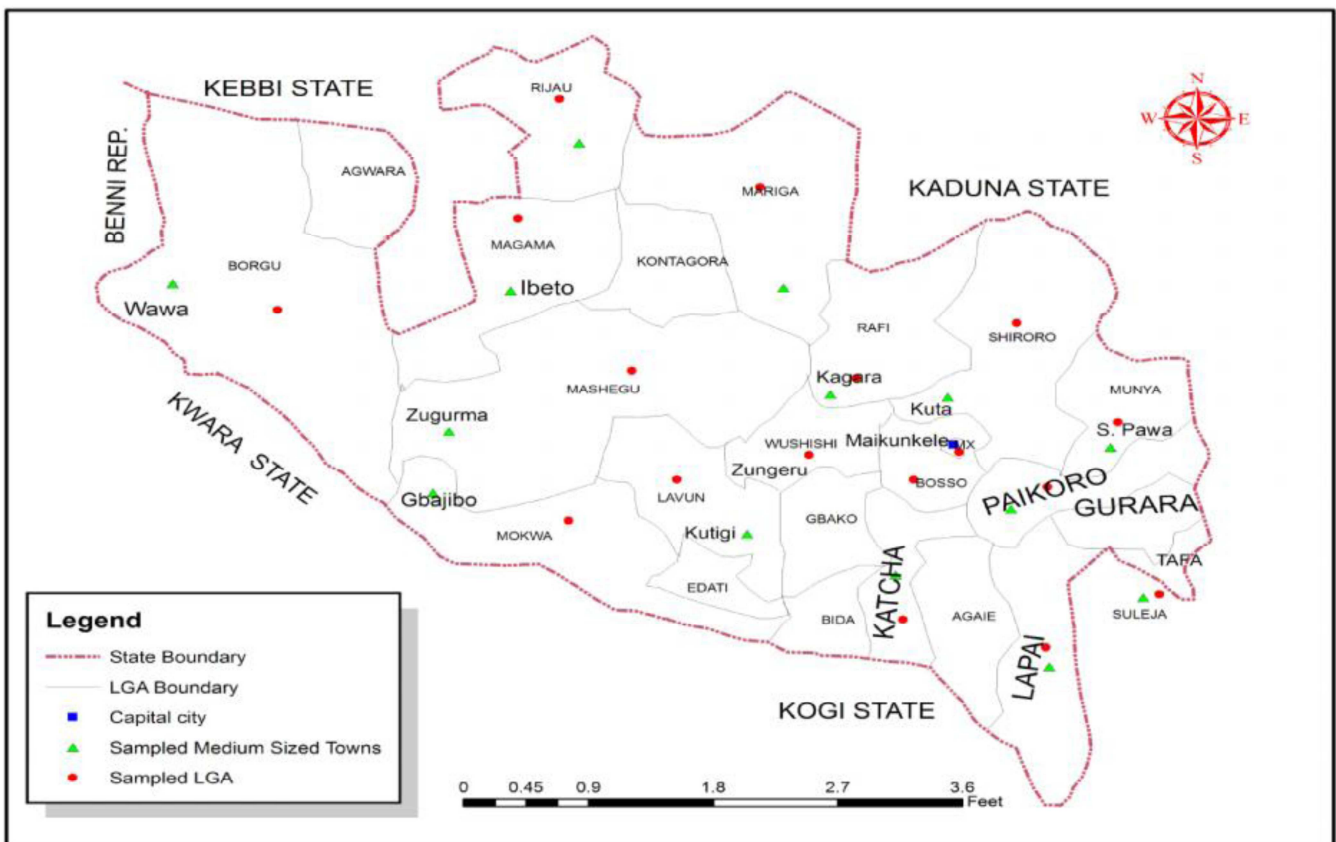

Figure 1. *Sampled Local Government Areas and Medium-sized Towns. (Source: Niger State Ministry of Land and Surveys, Minna)*

2.2. Data

Sixteen towns with the population between 5,000 and 20,000 people across the state were sampled as medium sized towns (see figure 1). Each sampled town was divided into wards and ten percent of the total wards were systematically sampled. Water samples were collected from rivers/streams of each of the sampled medium sized towns in both dry and rainy seasons. Seventeen (17) water quality parameters were monitored between December 2010 and November 2011 to cover the two seasons of the year. This was based on the guidelines of [31] procedure for analyzing the quality of surface water which involves information on chemical, physical and biological parameters. Five out of the seventeen water quality parameters that were most influential in [14] were used in this study.

2.3. Geostatistical Techniques

Geostatistics is a branch of statistics that specializes in the analysis and interpretation of any spatially (temporally) referenced data [13]. It is based on observations that are similar within certain proximity that is, mutually correlated. It can also be said to be a collection of techniques and theories that can be used to generate sampling designs, build statistical models, make spatio-temporal predictions at unsampled locations, extract spatio-temporal patterns in the data and analyze the associated uncertainties [26]. Its basic tool is variogram analysis which involves the study of the variogram function of a specific variable physical value or of water quality parameters under study. The variogram function, with its specific parameters (nugget value, threshold and correlation range), presents the behavior of the variable under study called the "regionalized variable" [9, 12] and thus permits the formulation of conclusions concerning areas that are not represented by any measurement data.

2.3.1. Kriging

Kriging is a means of spatial, temporal or spatio-temporal prediction and of estimating unknown local values of variables that are distributed in a space of one, two or three dimensions from more or less sparse data. It is an exact linear interpolating procedure. It is based on a spatial linear model for the data which specifies a parametric spatial mean function and spatial dependence structure [3]. The basic assumption in Kriging is that the data comes from a stationary stochastic process and some methods require that the data be normally distributed. Kriging differs from other methods (such as IDW), in which the weight function (w_i) is no longer arbitrary, being calculated from the parameters of the fitted semivariogram model under the conditions of unbiasedness and minimized estimation variance for the interpolation. The method of Kriging attempts to model the variability in the data as a function through the variogram [13]. A data point estimated by Kriging will have exactly the same magnitude as the original observation. This is because in the estimation procedure Kriging weights each observation according to the distance and direction between that point and the point to be estimated or kriged. If the weights are equal, we have the classical estimate of the mean. The weights are distributed using any of the following methods; inverse of the square of the distances, the inverse of the distance, and the inverse of the number of values. It also uses the information from the semivariogram to find an optimal set of weights. They are chosen to minimize the Kriging variance or the square root of the Kriging error. In this sense, the estimates are optimal [33]. Thus, Kriging is regarded as a best linear unbiased estimation (BLUE).

Kriging is divided into two distinct tasks: viz. quantifying the spatial structure of the data and producing a predicted surface. In order to predict an unknown value for a specific location, Kriging will use the fitted model from variogram, the spatial data configuration, and the values of the measured sample points around the prediction location [6]. Because Kriging uses statistical models, it allows a variety of map outputs, including predictions, prediction standard errors, probability, and quantile maps. Today, a number of variants of Kriging are in general use, these are: Simple Kriging (SK), Ordinary Kriging (OK), Universal Kriging (UK), Block Kriging (BK), Co-Kriging (CK) and Disjunctive Kriging (DK). Among the various forms of Kriging, Ordinary Kriging (OK) has been used widely as a reliable estimation method [23].

2.3.2. Interpolation by Ordinary Kriging (OK)

OK is used to model the spatial variability of each of the five influential parameters and to perform their estimation in sampled locations. It is based on the concept of a variable $Z(p)$ that is both random and spatially autocorrelated [12]. The predictions are based on the model:

$$Z(p_i) = \mu + \varepsilon'(p_i) \qquad (2.1)$$

where μ is the constant stationary function (global mean) and $\varepsilon'(p_i)$ is the spatially correlated stochastic part of variation. The predictions are obtained using:

$$\hat{Z}_{OK}(p_0) = \sum_{i=1}^{n} w_i(p_0).Z(p_i) = \lambda_0^T.a \qquad (2.2)$$

where λ_0 is the vector of kriging weights (w_i), a is the vector of n observations at primary locations.

The semivariogram is a convenient tool in geostatistics for the analysis of spatial dependence structure [8]. It is based on simple measure of dissimilarity and is defined by:

$$\gamma(h) = \frac{1}{2} Var(Z(p_i) - Z(p_i + h)) \qquad (2.3)$$

where $Z(p_i)$ is the value of random variable at some sampled location and $Z(p_i + h)$ is the value of the location at distance $p_i + h$.

In order to determine the spatial coherence of each of the parameters and to identify the best model variable mode, the variogram for each parameter was drawn through linear, spherical, exponential and Gaussian models using a

$C_0/(C_0 + C)$ relationship [5]. The nugget to sill ratio as described by [31] was used to analyze the spatial structure. A variable is said to have strong spatial dependence if the ratio is less than 0.25, and has a moderate spatial dependence if the ratio is in between 0.25 and 0.75; otherwise the variable has weak spatial dependence.

2.3.3. Variogram Models

Because the kriging algorithm requires a positive definite model of spatial variability, the experimental variogram cannot be used directly. Instead, a model must be fitted to the data to approximately describe the spatial continuity of the data [18]. Experimental variogram for Escherichia coli, Magnesium, Total Coliform, Total Dissolved Solid (TDS) and Total Hardness were calculate at a lag distance of 500m. Thereafter, the models of spatial variability were fitted to the experimental variogram by minimizing the sum of squares between the experimental values and those of the model. Some important models [7, 8] are linear, spherical, exponential, and Gaussian models.

Linear model

$$\gamma(h) = \begin{cases} \tau^2 + \sigma^2 h, & \text{if } h > 0 \\ 0, & \text{otherwise} \end{cases}$$

Spherical model

$$\gamma(h) = \begin{cases} \tau^2 + \sigma^2, & \text{if } h \geq range \\ \tau^2 + \sigma^2 \left[\dfrac{1.5 * h}{range} - \dfrac{0.5 * h^3}{range^3} \right], & \text{if } 0 < h \Leftarrow range \\ 0, & \text{otherwise} \end{cases}$$

Exponential model

$$\gamma(h) = \begin{cases} \tau^2 + \sigma^2 (1 - \exp(-\phi h)), & \text{if } h > 0 \\ 0, & \text{otherwise} \end{cases}$$

Gaussian model

$$\gamma(h) = \begin{cases} \tau^2 + \sigma^2 (1 - \exp(-\phi^2 h^2)), & \text{if } h > 0 \\ 0, & \text{otherwise} \end{cases}$$

Where $\tau^2 + \sigma^2$ is the value of the semivariogram at the sill, h is the separation distance and $\emptyset$ equals $r\sqrt{3}$.

In this study, OK is applied to each parameter data set using linear, spherical, exponential, and Gaussian models. This is used for spatial prediction of data values of the five water parameters.

2.4. Cross Validation

The semivariogram models were tested for each parameter data set. The quality of prediction performances were assessed by cross validation. Cross validation was conducted to assess the accuracy of the OK through some statistical measurements of the prediction error: the mean error (ME), the root-mean-square error (RMSE) and the root-mean-square standardized error (RMSSE) defined as follows:

$$ME = \frac{1}{L}\sum_{i=1}^{L} [\,\hat{z}(p_i) - z^*(p_i)] \tag{2.4}$$

$$RMSE = \sqrt{(\frac{1}{L}\sum_{i=1}^{L} [\,\hat{z}(p_i) - z^*(p_i)]^2)} \tag{2.5}$$

$$RMSSE = \sqrt{(\frac{1}{L}\sum_{i=1}^{L} \left[\frac{\hat{z}(p_i) - z^*(p_i)}{\hat{\sigma}_i}\right]^2)} \tag{2.6}$$

where $\hat{z}(p_i)$ are estimated values, $z^*(p_i)$ are actual observations, L is the number of validation points and $\hat{\sigma}_i$ is the prediction standard error in location p_i.

For a model that provides accurate predictions, the ME should be close to zero, the RMSE should be as small as possible (this is useful when comparing models), and the RMSSE should be close to one for good prediction [15].

2.5. R Geostatistics Packages

This study introduces the functionality of five (5) R geostatistics packages that were used to run the processing and display the results: *gstat, sp, rgdal, spatstat and maptools*. All these are available as open source or as freeware and no licenses are needed to use them. By combining the capabilities of the five packages, the study harnessed the best out of each package and optimized preparation, processing and the visualization of the spatial maps. In this case, *gstat* calculates sample (experimental) variograms; plots an experimental variogram with automatic detection of lag spacing and maximum distance; iteratively fits an experimental variogram; a generic function to make predictions by inverse distance interpolation, ordinary kriging and runs krige with cross-validation; package *sp* provides general purpose classes and methods for visualizing spatial data; *rgdal* produces map projections; s*patstat* used for various types of statistical and geostatistical analysis; and *maptools* used for getting shape files into R and converts some *sp* objects for use in *spatstat*.

3. Results and Discussion

According to [34], assessing water quality values using geostatistical techniques require a normal distribution of the parameter values under investigation. In this study, histogram and normal QQPlot analysis were applied to each water quality parameter and it was found that E. coli, Magnesium and TH parameters shows normal distribution. It was also found that Total Coliform and TDS parameters (see figures 7 and 8 under the appendix) exhibited non-normal distributions and therefore do not satisfy the basic assumption of normality which is a condition for geostatistical analysis. Logarithmic transformation was performed on Total Coliform and Total Hardness parameters to make them closer to normal distribution (see figures 9 and 10). The deviations from the straight line are minimal. After the transformation, Kolmogorov-Smirnov test was performed and the result shows that the histograms do not differ much and are normally distributed.
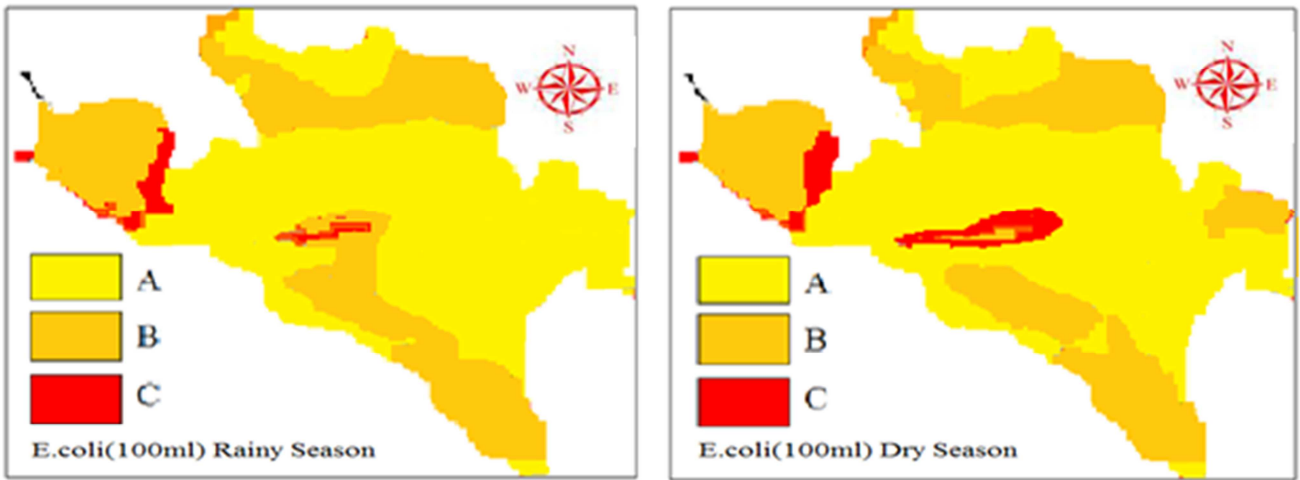
Figure 2. Spatial Distribution of E.coli Surface Water Parameter for Rainy and Dry Seasons.
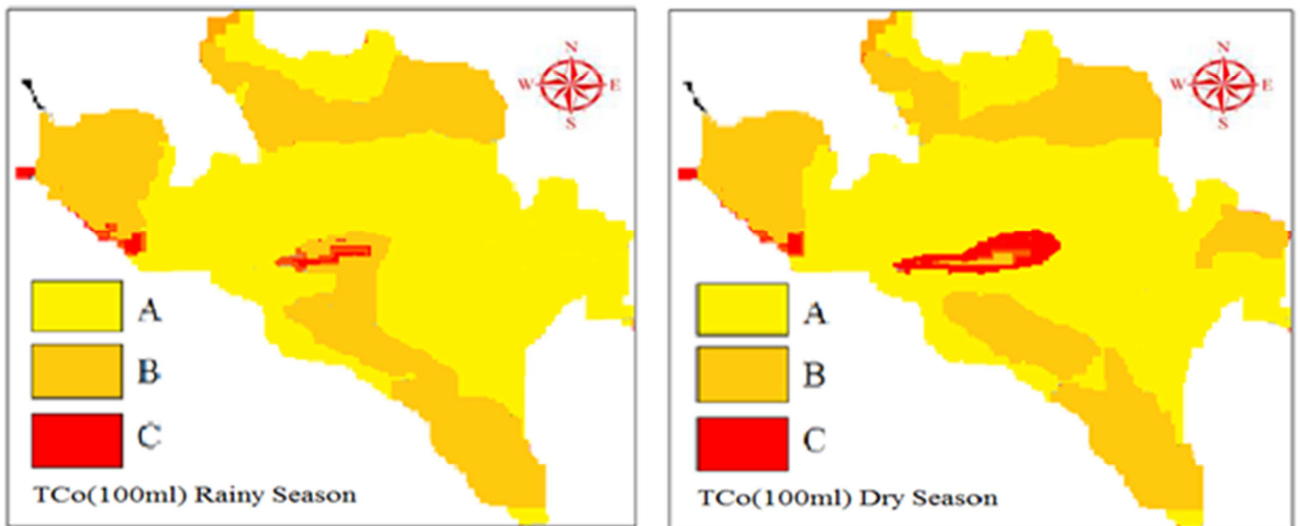

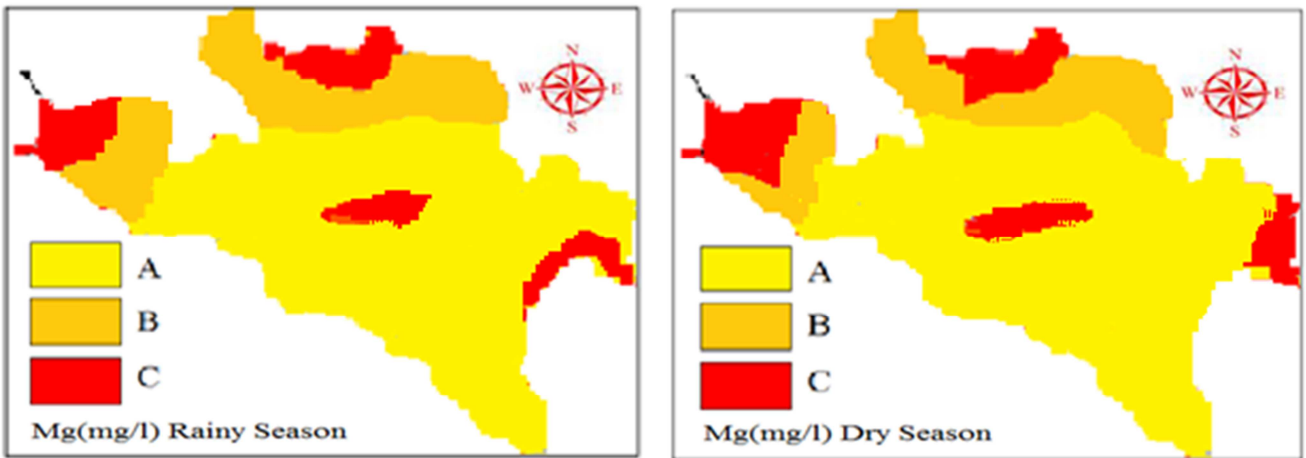*Figure 3. Spatial Distribution of Total Coliform Surface Water Parameter for Rainy and Dry Seasons.*

Figure 4. Spatial Distribution of Magnesium Surface Water Parameter for Rainy and Dry Seasons.
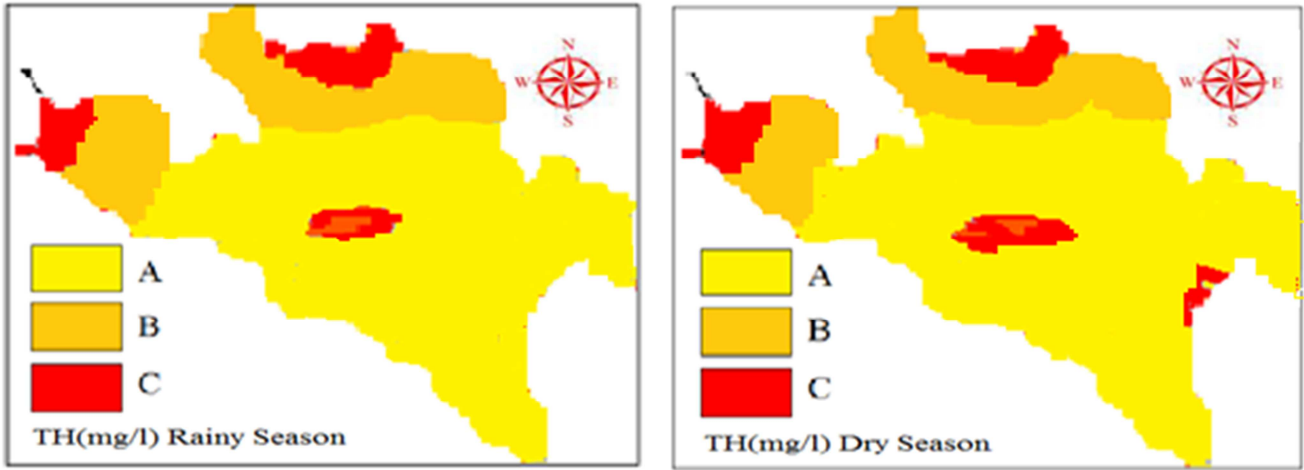
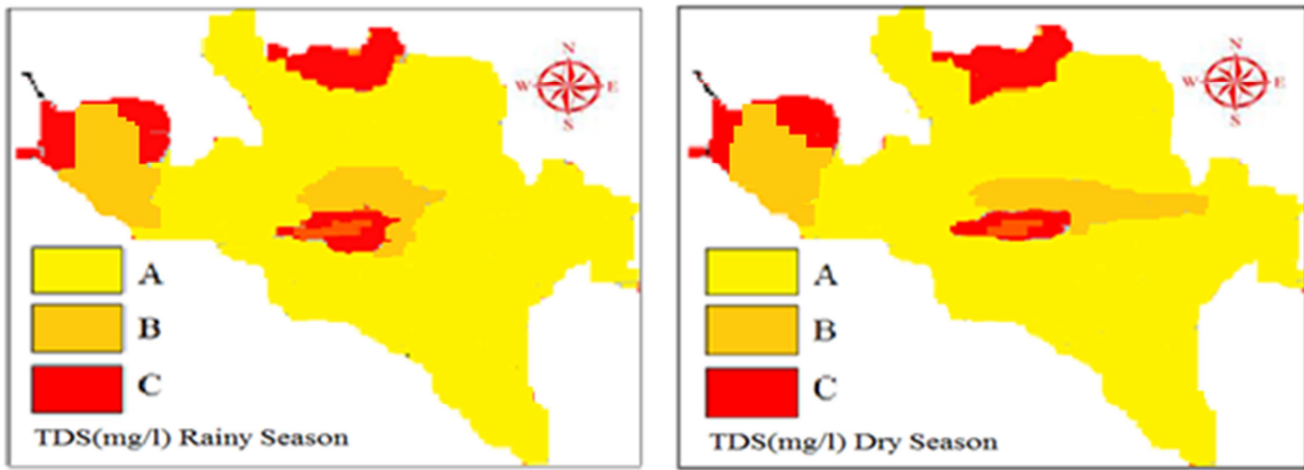***Figure 5.*** *Spatial Distribution of Total Hardness Surface Water Parameter for Rainy and Dry Seasons.*
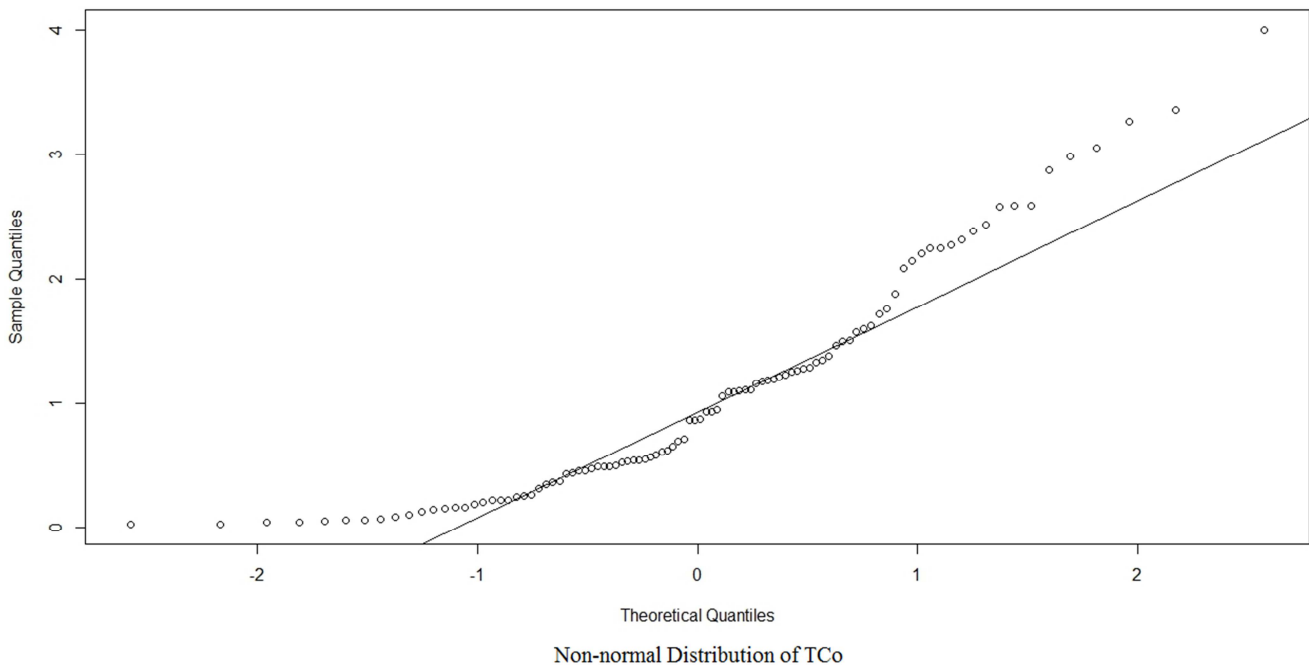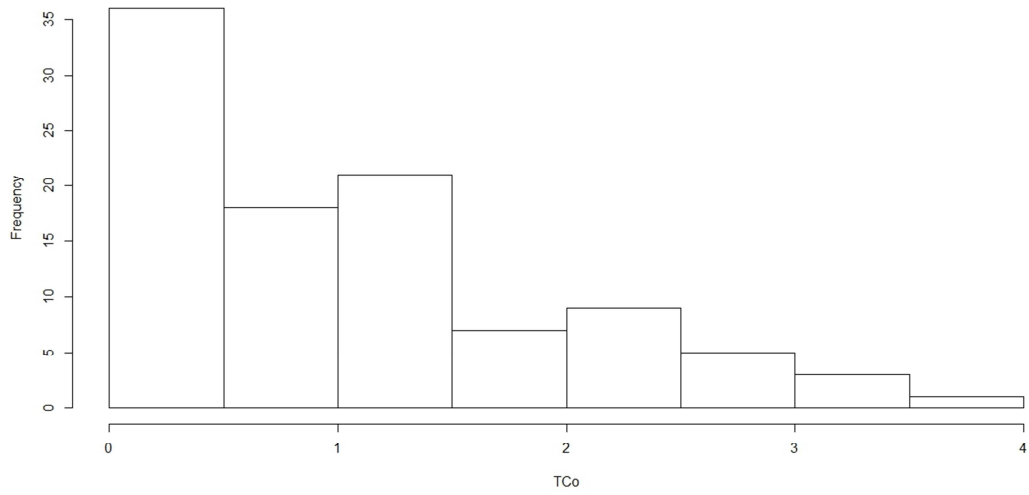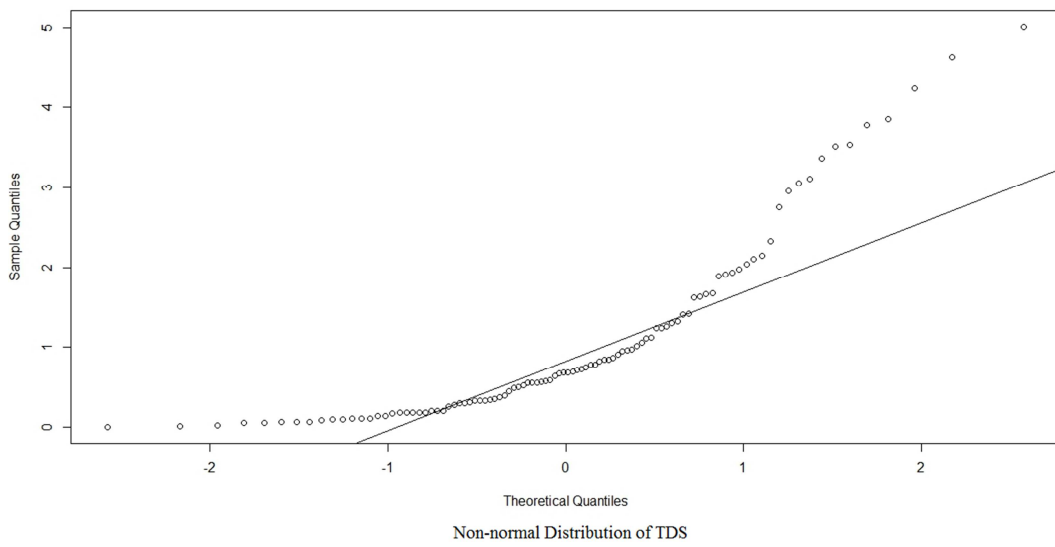

Figure 6. *Spatial Distribution of Total Dissolved Solid Surface Water Parameter for Rainy and Dry Seasons.*
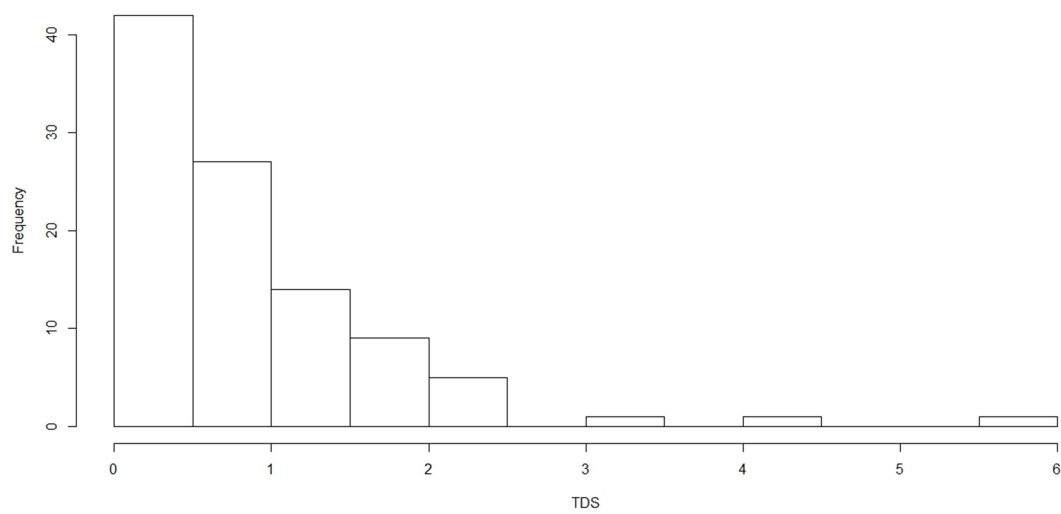
Histogram of TCo for Non-normal Distribution

Figure 7. *Theoretical Quantile and Histogram of TCo for Non-Normal Distribution.*

Non-normal Distribution of TDS
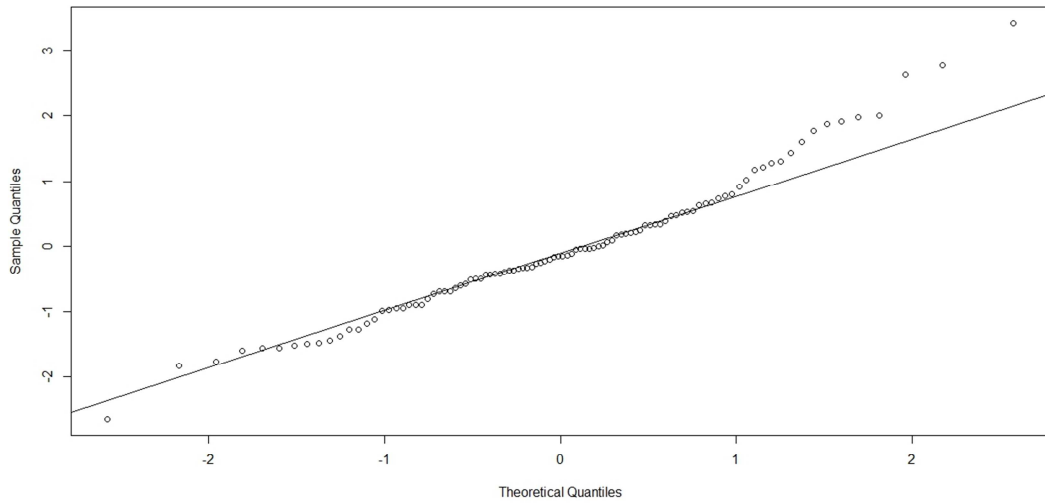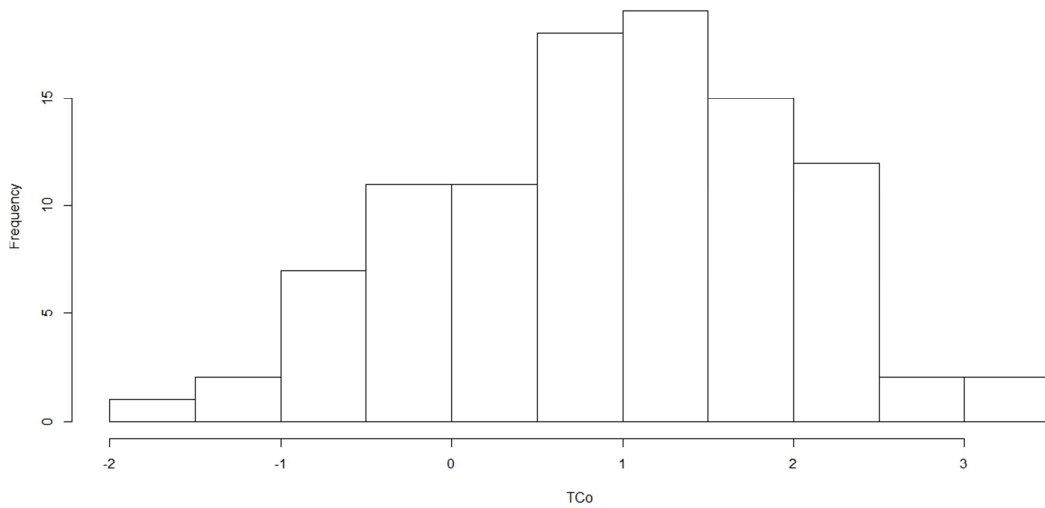

Histogram of TDS for Non-normal Distribution

Figure 8. *Theoretical Quantile and Histogram of TDS for Non-Normal Distribution.*

Log Transformation to Normal Distribution of TCo

Histogram of Log Transformation to Normal Distribution of TCo

Figure 9. *Theoretical Quantile and Histogram of Log Transformed TCo.*
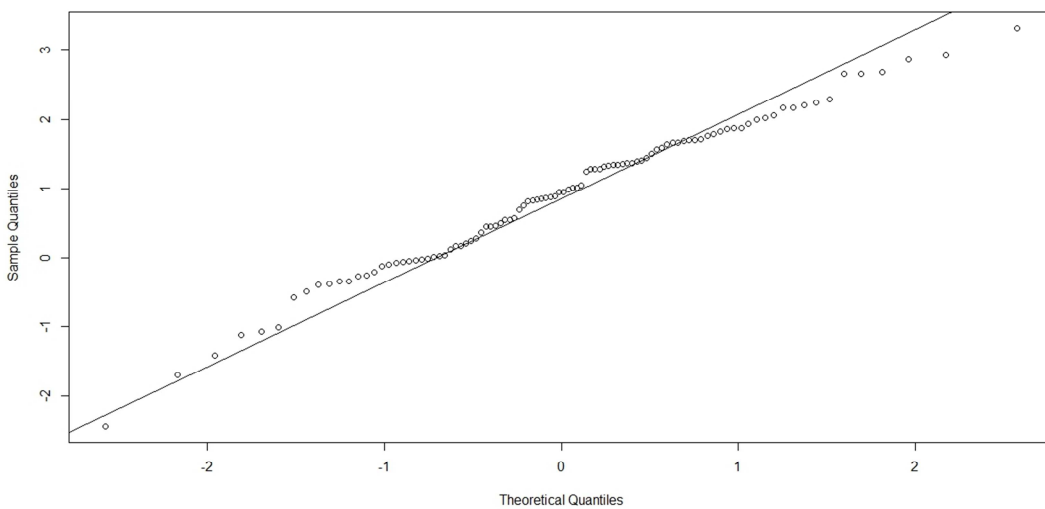
Log Transformation to Normal Distribution of TDS
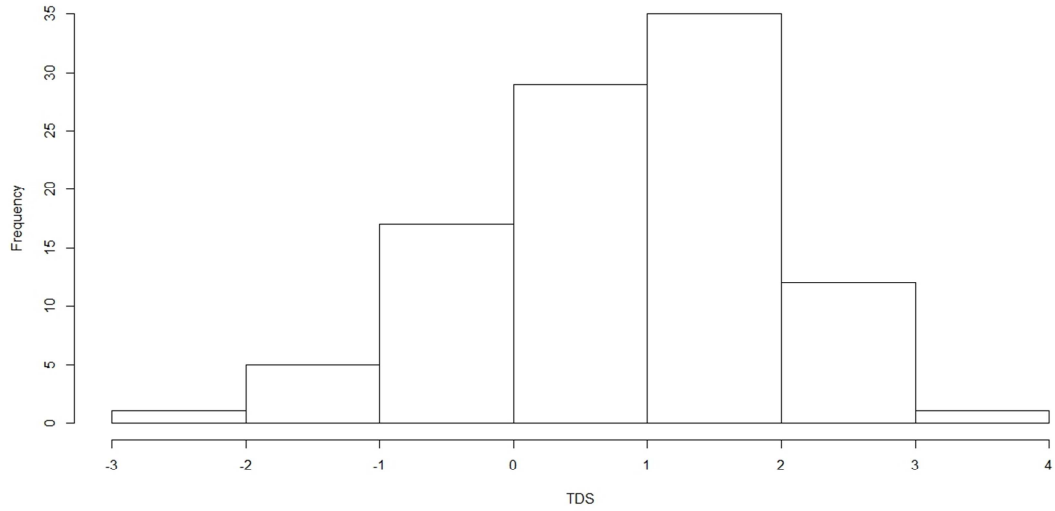
Histogram of Log Transformation to Normal Distribution of TDS

Figure 10. *Theoretical Quantile and Histogram of Log Transformed TDS.*
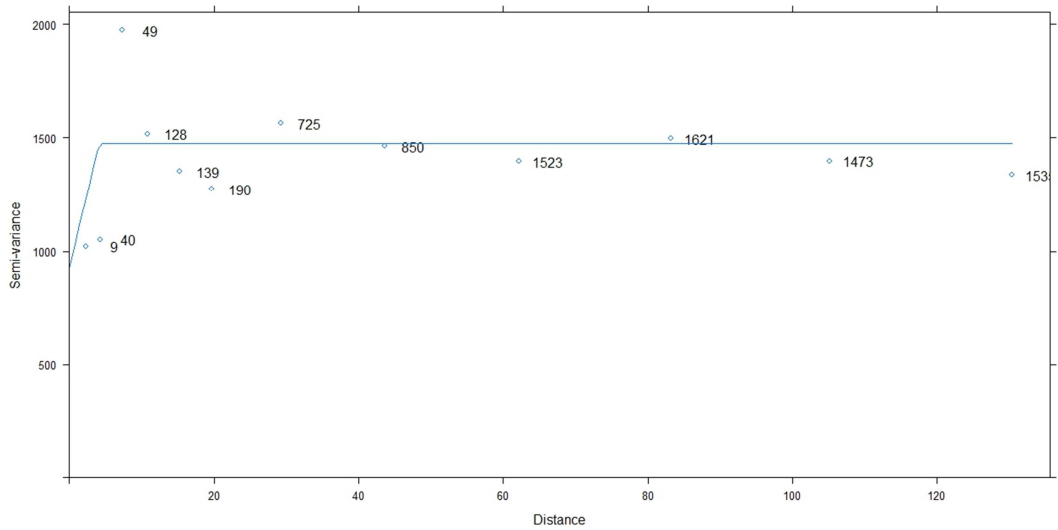
Experimental and Fitted Linear Variogram Model of E.coli During Rainy Season
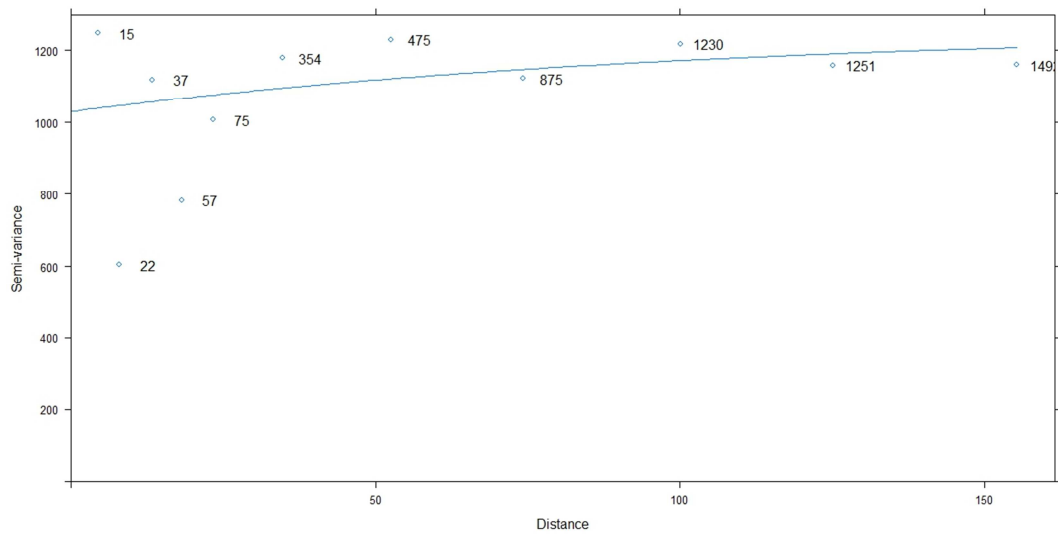
Experimental and Fitted Exponential Variogram Model of E.coli During Dry Season

Figure 11. *Experimental and fitted variogram models of E. coli for rainy and dry seasons.*
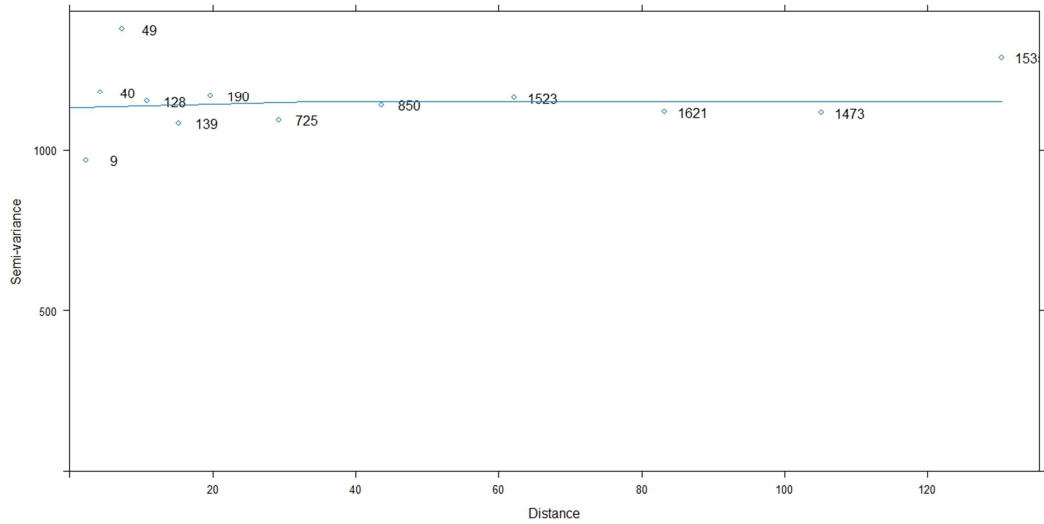
Experimental and Fitted Linear Variogram Model of TCo During Rainy Season
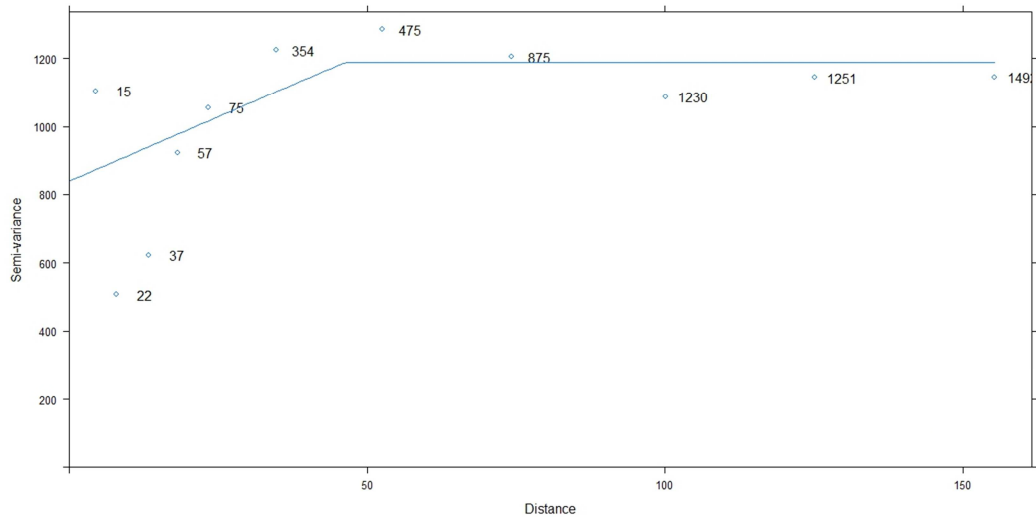
Experimental and Fitted Linear Variogram Model of TCo During Dry Season

Figure 12. *Experimental and fitted variogram models of TCo for rainy and dry seasons.*
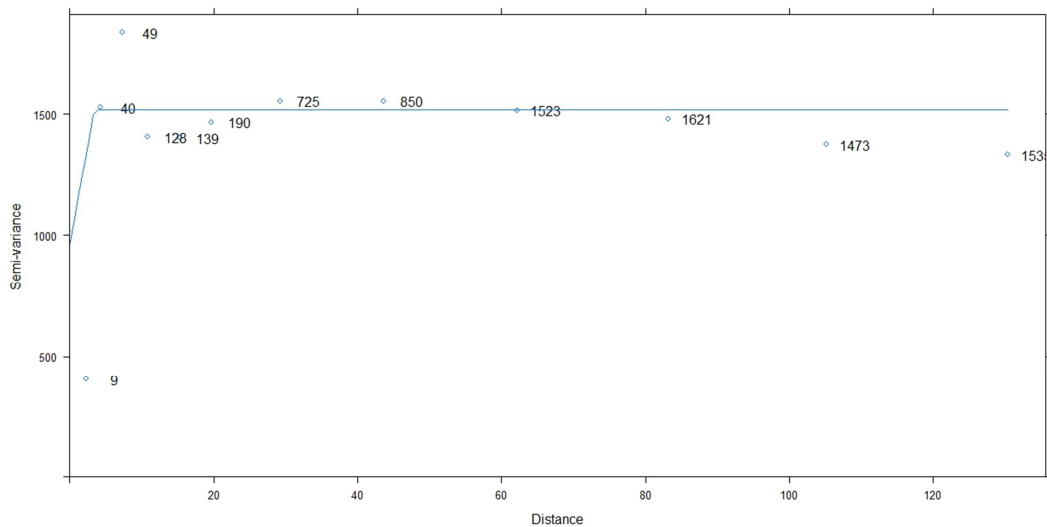
Experimental and Fitted Linear Variogram Model of Magnesium During Rainy Season

Figure 13. *Experimental and fitted variogram models of Magnesium for rainy and dry seasons.*

Figure 14. *Experimental and fitted variogram models of TH for rainy and dry seasons.*
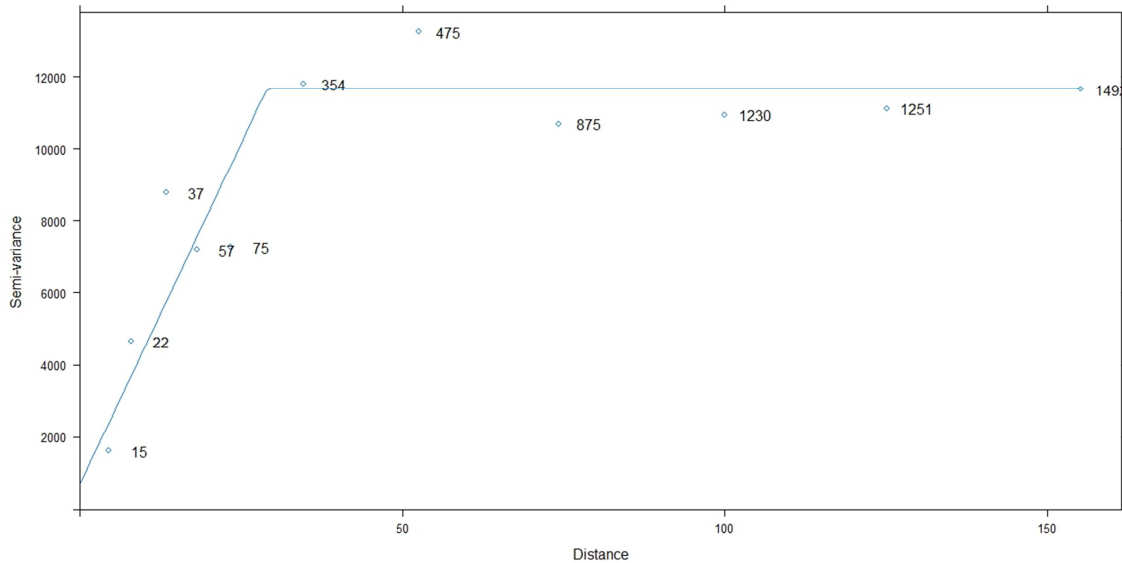
Experimental and Fitted Linear Variogram Model of TDS During Rainy Season
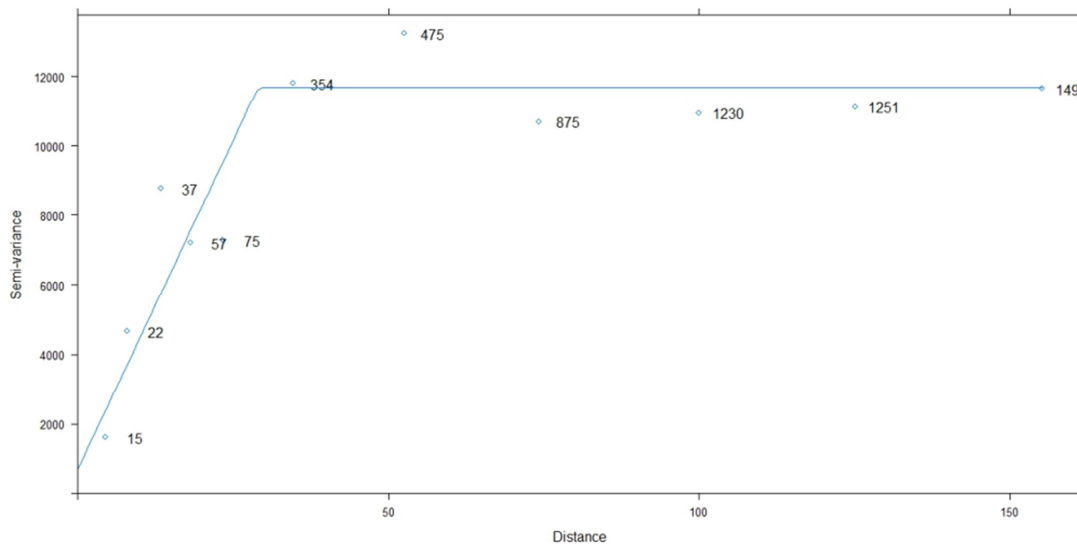
Experimental and Fitted Linear Variogram Model of TDS During Dry Season

Figure 15. *Experimental and fitted variogram models of TDS for rainy and dry seasons.*

A total of 125 surface water samples were collected from 16 sampled medium sized towns during rainy and dry seasons. The descriptive statistics for both seasons can be seen in Table 3.1. Fro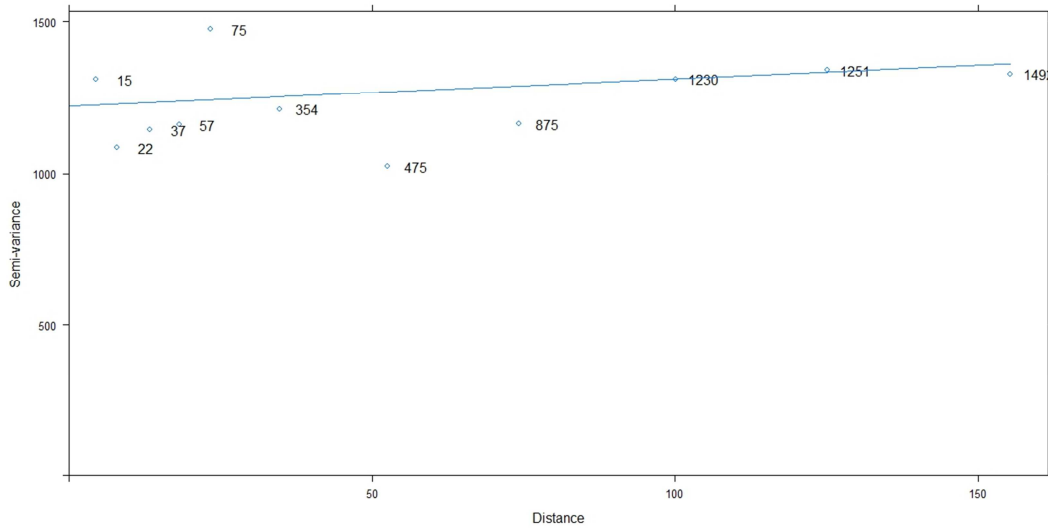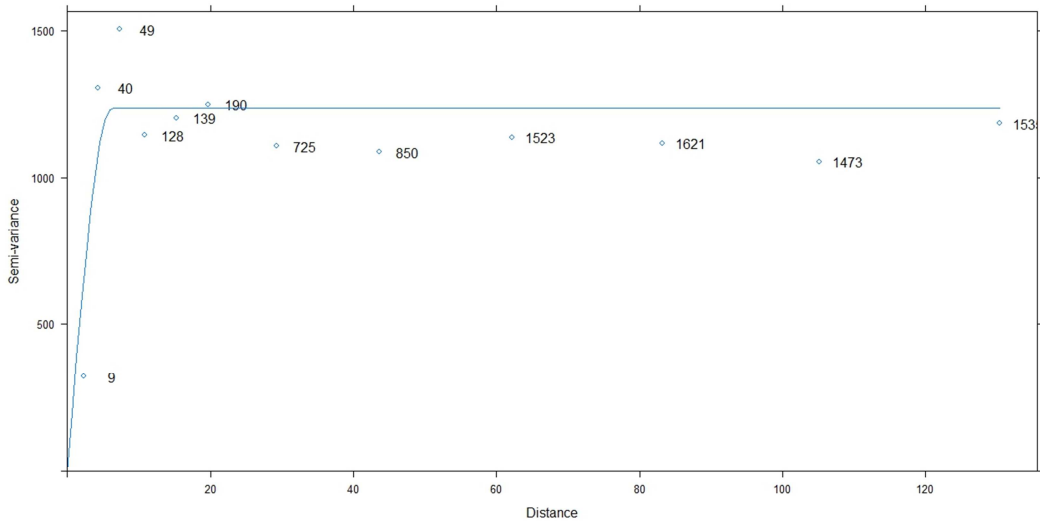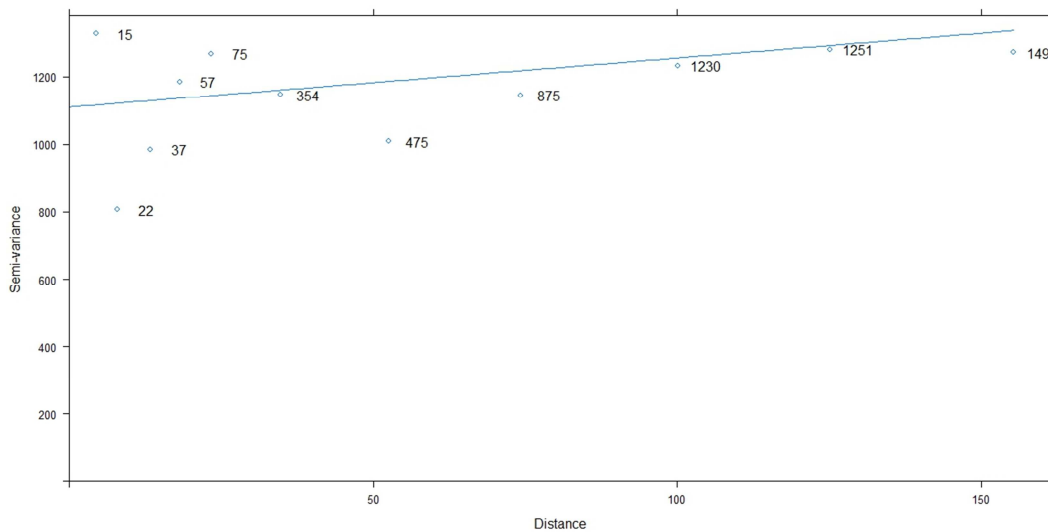m the results, the two seasons are almost identical. However, these two seasons are significantly different in ways that do not incorporate the spatial locations of data into their defining computations by the common descriptive statistics. The spatial distribution of E.coli, total coliform, magnesium, total hardness, and TDS concentrations developed from the cross validation process are given in figures 2 to 6, respectively.

The cross validation reports, that examined the validity of the fitting models and parameters of semivariograms for river water parameters are given in Table 3.2. For example, during rainy season and using E.coli and TCo parameters as an example, the best fit model for E.coli and TCo is the linear model with a 0.148 and 0.308 ME, respectively. Also,

the experimental and fitted linear variogram models plot never level out, therefore, the linear model is considered the best. While in dry season, exponential model is the best fit for E.coli with an ME value of 0.100 and RMSS value of 0.620 whereas linear model fitted well for TCo with ME value of 0.303 and RMSS value of 0.683. This result shows that linear model is the best for both seasons.

After performing kriging cross-validation for different models for each water quality parameter, the prediction errors were calculated and models giving best results were determined. Table 3.3 shows the most suitable models and their prediction error values for each parameter.

The variograms for the OK are presented in Table 3.4. The parameters were obtained by using measurement error to estimate the nugget, global variance to estimate the sill and the mean distance to nearest neighbor to estimate the range. The fitted models have the following structure:

Table 3.1. *Descriptive Statistics.*

Parameter	N	Min	Max	Mean	Median	S.D	Skew	Kurt	WHO
Rainy Season									
E. coli(100ml)									
<50	54								
50-100	31	0.00	124.00	57.06	24.00	43.03	1.03	2.95	0
>100	40								
T. Coliform(100ml)									
<50	71								
50-100	41	0.00	124.00	98.32	52.66	64.75	0.24	2.13	10
>100	13								
Magnesium(mg/l)									
<50	76								
50-100	22	4.00	144.13	3.16	33.02	2.04	2.01	7.24	50
>100	27								
T. Hardness(mg/l)									
<50	85								
50-100	31	17.01	157.14	40.11	44.03	24.63	1.92	6.67	500
>100	9								
TDS(g/l)									
<100	76								
100-200	36	18.27	498.48	86.57	50.25	48.39	3.57	16.19	500
>200	13								
Dry Season									
E. coli(100ml)									
<50	52								
50-100	35	1.00	23.16	58.91	21.81	42.38	1.03	2.95	0
>100	38								
T. Coliform(100ml)									
<50	70								
50-100	45	0.00	122.20	112.22	53.00	72.25	0.24	2.13	10
>100	10								
Magnesium(mg/l)									
<50	76								
50-100	25	4.00	138.53	3.33	33.02	2.27	2.01	7.24	50
>100	24								
T. Hardness(mg/l)									
<50	86								
50-100	33	17.01	156.64	39.03	44.03	23.83	1.92	6.67	500
>100	6								
TDS(mg/l)									
<100	74								
100-200	40	16.08	496.49	85.60	48.33	47.86	3.57	16.19	500
>200	11								

Table 3.2. *Cross-validation Report of E.coli and Total Coliform Parameters.*

Parameter	Model	Prediction Error					
		Rainy Season			Dry Season		
		ME	RMSE	RMSSE	ME	RMSE	RMSSE
E. coli	Linear	0.148	9.282	0.689	0.118	7.803	0.326
	Spherical	0.207	9.377	0.536	0.187	7.664	0.419
	Exponential	0.261	9.314	0.531	0.100	7.130	0.620
	Gaussian	0.180	9.356	0.570	0.135	7.361	0.533
TCo	Linear	0.308	7.749	0.722	0.303	5.968	0.683
	Spherical	0.374	7.603	0.683	0.370	5.725	0.617
	Exponential	0.466	7.620	0.590	0.326	5.294	0.653
	Gaussian	0.394	7.652	0.601	0.312	5.416	0.644

Table 3.3. Most Suitable Models and their Prediction Error Values by Parameter.

Parameter	Prediction Error								
	Model	Rainy Season			Model	Dry Season			
		ME	RMSE	RMSSE		ME	RMSE	RMSSE	
E.coli	Linear	0.148	9.282	0.689	Exponential	0.100	7.130	0.620	
TCo	Linear	0.308	7.749	0.722	Linear	0.303	5.968	0.683	
Mg	Linear	0.408	14.037	0.664	Linear	0.166	10.268	0.597	
TH	Spherical	0.206	10.033	0.714	Exponential	0.201	9.003	0.590	
TDS	Linear	0.216	5.661	0.768	Linear	0.154	4.037	0.911	

Table 3.4. Experimental Variogram and Fitted Variogram Models of E.coli and Total Coliform Parameters.

Model	Rainy Season					Dry Season				
	E.coli					E.coli				
	Nugget	sill	range	$C_o/(C_o+C)$	R^2	Nugget	sill	range	$C_o/(C_o+C)$	R^2
Linear	321	1476	4.1	0.22	0.84	103	1257	42	0.08	0.86
Exponential	-	-	-	-	-	101	1162	56	0.09	0.76
Spherical	477	1436	7.4	0.33	0.72	611	1173	97	0.52	086
Gaussian	-	-	-	-	-	391	1029	49	0.38	0.78
	TCo					TCo				
Linear	113	1150	3.6	0.09	0.88	248	1189	46	0.20	0.85
Gaussian	-	-	-	-	-	646	850	49	0.76	0.71

Experimental variogram and fitted variogram models evaluation in Table 3.4 for rainy and dry season's results indicate a high spatial coherence for magnesium and total hardness parameters, while E.coli and total coliform parameters indicate a medium coherence and TDS parameter indicate a low spatial coherence.

Results of semivariogram analysis are provided in Table 3.5. Linear model fitted best in rainy and dry seasons in all the parameters, except for magnesium. The nugget to sill ratios of linear model in all cases were less than 0.25 indicating that the river water level have strong spatial coherence in both seasons.

The range is the distance within which the parameters are spatially correlated. The R^2 values of 0.80 to 0.92 indicate that the variograms were chosen correctly and the predictions were accurate.

As with sanitary inspection, data on E.coli, total coliform, magnesium, total hardness and TDS water quality may usefully be divided into a number of categories; the levels of contamination associated with each category should be selected in the light of local circumstances. A typical classification scheme is presented in Table 3.6, based on increasing orders of magnitude of contamination [35].

Table 3.5. Best-Fitted Experimental Variogram and Fitted Variogram Models.

Parameter	Rainy Season						Dry Season					
	Best-Fit Model	Nugget	sill	range	$C_o/(C_o+C)$	R^2	Best-Fit Model	Nugget	sill	range	$C_o/(C_o+C)$	R^2
E.coli	Linear	321	1476	4.1	0.22	0.84	Linear	103	1257	42	0.08	0.86
TCo	Linear	113	1150	3.6	0.09	0.88	Linear	248	1189	46	0.20	0.85
Mg	Exponential	132	1239	4.2	0.10	0.83	Exponential	110	358010	58	0.003	0.82
TH	Linear	303	1235	6.2	0.25	0.80	Linear	1222	175573	65	0.006	0.89
TDS	Linear	289	1518	3.4	0.19	0.92	Linear	498 m	11682	29	0.04	0.81

Table 3.6. Classification and Color-code Scheme for the Five Parameters.

Count Per 100ml for E.coli & T.Coliform	Count Per mg/L for Mg, TH & TDS	Category & Color-code	Remark
0	0-1000	A (Yellow)	In conformity with WHO guidelines
1-10	1000-3000	B (Orange)	Low risk
10-1000	3000-10000	C (Red)	High risk

Source: WHO Geneva 2011- Guidelines for Drinking-Water Quality 3rd ed.

3.1. Escherichia Coli

Table 3.1 indicates that the mean value of E.coli is 57.06 cfu/100ml in rainy season; it increases slightly to 58.91cfu/100ml in dry season. The spatial distribution of E.coli shows that some rivers did not meet the standard of zero tolerance indicated by [35]. The continuous high E.coli

concentration occurs within Northwest and city center.

3.2. Total Coliform

The presence of total coliform in surface water may indicate that the surface water has been affected by surface runoff and anthropogenic pollution. Based on [35], the total coliform must be 10 (100mL) to protect human from diseases, such as

diarrhoea, nausea, vomiting, cramps or other gastrointestinal distress. Table 3.1 shows that the mean values of total coliform ranges from 112.22 and 114.36 100ml in the dry and rainy seasons, respectively. However, the total coliform range from 0 to 124 (100mL) in the study areas. The spatial distribution of total coliform shows high concentrations in both seasons and occurs within Northwest and the city center (see figure 3).

3.3. Magnesium

Higher concentration of magnesium makes the water unpalatable and act as laxative to human beings. Table 3.1 shows the mean concentrations of magnesium range between 3.33 mg/l and 3.46 mg/l in dry and rainy seasons respectively. In dry season, the maximum magnesium value reaches nearly 138.53 mg/l, which is considerably higher than the permissible limit of 50mg/l in [35]. In rainy season, the maximum concentration of magnesium reaches 144.13mg/l. However, in both seasons, the mean concentrations are higher than the permissible limit of 50(mg/L). The content of magnesium increases from the Northwest to North and Northcentral to Northeast. Figure 4 shows the presence of high magnesium concentration in the three geopolitical zones of the state and in the two seasons.

3.4. Total Hardness

The presence of high calcium and magnesium level shows consistence of water hardness in such sources of water. From Table 3.1, the mean hardness for both seasons is lower than the [35] drinking water standard of 500 mg/l. The total hardness value of the river water ranges from 17.01mg/l to 167.14mg/l in rainy season and from 17.01mg/l to 156.64mg/l in dry season. Figure 5, shows that the value of water hardness concentration is the same as magnesium concentration.

3.5. Total Dissolved Solids (TDS)

[20] reported that high TDS values have the tendency to absorb heat from the sun thereby raising the temperature and increasing the turbidity of water. Table 3.1, the mean values of TDS are less than the [35] standard (500 mg/l) for both seasons. The TDS values range from 18.27 to 498.48 mg/l and from 16.08 to 496.49 mg/l for rainy and dry seasons, respectively. Since both seasons fall within 500 (mg/L) and 1,000 (mgl) they can be tolerated with little health effects. As indicated in figure 6, high concentrations occurred around the rivers in Northwest.

4. Conclusion

The water quality standard of the World Health Organization [35] was used as the basis for the surface water quality evaluation (Table 3.1). Ordinary Kriging (OK) was used to determine the spatial continuity of the river water quality parameters. Different semivariogram models namely; linear, spherical, exponential and gaussian were tested. The semivariogram parameters; nugget, sill, and range, with $C_0/(C_0 + C)$ and R^2 were determined and the performance

of each model was evaluated using cross-validation, which examines the accuracy of the generated surfaces. Thereafter, the models with smallest ME were selected. The spatial prediction maps of river water were calculated using ordinary kriging for both seasons using R software.

The descriptive statistics of the parameters shows that the mean concentrations of E.coli, and Total Coliform in both seasons are greater than permissible limit of 0 ml to 10 ml and is not in conformity with WHO [35] standards of drinking water quality. This means that there is presence of faeces contamination by animals, including birds. While Magnesium, Total Hardness and TDS mean values in both seasons meet the recommended limit of [35]. The nugget to sill ratios of experimental and linear fitted variogram models in all cases were less than 0.25 indicating that the river water level has strong spatial coherence in both seasons and therefore, linear model fitted best. Spatial variability maps of surface water level indicated that the two seasons are almost identical. The maps show that water quality in dry season changes more rapidly than it does in rainy season.

Recommendations

The study only looked at five surface water parameters out of the several parameters. It is recommended that other parameters not covered in this study be further investigated. It is also recommended that robust variogram model be used to improve the predictions at unsampled locations.

References

[1] Agoubi, B., Kharroubi, A., & Abida, H. (2013). Hydrochemistry of groundwater and its assessment for irrigation purpose in coastal Jeffara Aquifer, southeastern Tunisia. *Arabian Journal of Geosciences, 6*(4), 1163–1172. doi:10.1007/s12517-011-0409-1.

[2] Babiker, I. S., Mohamed, M. A. A., & Hiyama, T. (2007). Assessing groundwater quality using GIS. *Water Resources Manage, 21*, 699–715.

[3] Berke O.I. (1999). Estimation and Prediction in the Spatial Linear Model, Water, Air, and Soil Pollution, 110, 215-237.

[4] Bordalo, A., Nilsumranchit, W., & Chalermwat, K. (2001). Water Quality and uses of the Bangpakong River (Eastern Thailand). *Water Research, 35*(15), 3635–3642.

[5] Cambardella, C. A. Moorman, T. B. Novak, J. M. Parkin, T. B. Karlen, D. L. Turco, R. F. and Konopka, A. E. (1994). Field-Scale Variability of Soil Properties in Central Iowa Soils, Soil Science Society of America Journal, Vol. 58, pp. 1501-1511.

[6] Cheesbrough, M. (2003). Water quality analysis. District Laboratory practice in Tropical countries (2) Cambridge University Press, United Kingdom. 146-157.

[7] Chiles, J. P., Delfiner, P. (1999). Geostatistics: modeling spatial uncertainty. John Wiley & Sons, New York.

[8] Cressie, N. A. C. (1993). Statistics for Spatial Data, revised edition. John Wiley & Sons, New York, p. 416.

[9]   De Gruijter, J.J., Marsman, B.A. (1985). Transect sampling for reliable information on mapping units. In: Nielsen, D.R., Bouma, J. (Eds.), Soil Spatial Variability. Pudoc, Wageningen, pp. 150– 163.

[10] Elci, A., & Polat, R. (2010). Assessment of the statistical significance of seasonal groundwater quality change in a karstic aquifer systemnear Izmir-Turkey. *Environmental Monitoring and Assessment, 172*(1), 445–462. doi:10.1007/s10661-010-1346-2.

[11] Galadima, A., & Garba, Z.N. (2012). Heavy metals pollution in Nigeria: Causes and consequences. *Elixir Pollution 45*, 7917-7922

[12] Heuvelink, G.B.M., Musters, P., Pebesma, E.J. (1997). Spatio-temporal kriging of soil water content. In: Baafi, E.Y., Schofield, N.A. Eds., Geostatistics Wollongong '96. Kluwer Academic Publishers, Dordrecht, pp. 1020–1030.

[13] Isah A. (2009). Spatio-Temporal Modeling of Nonstationary Processes, Journal of Science, Education and Technology, Vol.2 No.1 pp372-377.

[14] Isah A., Usman, A., & Mohammed, M., N. (2013). Application of Multivariate Methods for Assessment of Variations in Rivers/Streams Water Quality in Niger State, Nigeria. *American Journal of Theoretical and Applied Statistics.* Vol. 2, No. 6, 2013, pp. 176-183. doi: 10.11648/j.ajtas.20130206.14

[15] Johnston K., Hoef J.M.V., Krivoruchko K., Lucas N. (2001).Using ArcGIS Geostatistical Analyst. ESRI.380 New York Street. Redlands, CA 92373-8100, USA.

[16] Kazi T.G., Arain M.B., Jamali M.K., Jalbani N., Afridi H.I., Sarfraz R.A., Baig J.A., Shah A.Q. (2009). Assessment of water quality of polluted lake using multivariate statistical techniques: A case study. Ecotox. Environ. Safe. 72:301–309.

[17] Kumar, J., & Pal, A. (2012). Water quality monitoring of Ken River of Banda District, Uttar Pradesh, India. *Elixir Pollution* 42, 6360-6364.

[18] Lark, R. M. (2000). Estimating variograms of soil properties by the method-of-moments and maximum likelihood. *European Journal of Soil Science, 51*, 717–728.

[19] Liu, W.C., Yu, H.L., & Chung, C.E. (2011). Assessment of water quality in a subtropical Alpine Lake using multivariate statistical techniques and geostatistical mapping: a case study. *International Journal of Environmental Research and Public Health, 8*(4), 1126–1140.

[20] Morenikeji W., Sanusi Y. A, and Jinadu A. M. (2000). The Role of Private Voluntary Organizations' in Community and Settlement Development in Niger State, A Research Report Submitted to the Centre for Research and Documentation, Kano.

[21] Nas, B., & Berktay, A. (2010). Groundwater quality mapping in urban groundwater using GIS. *Environmental Monitoring and Assessment, 160*(1–4), 215–227. doi:10.1007/s10661-008-0689-4.

[22] Niger State Bureau of Statistics (2012). Facts and Figures about Niger state of Nigeria,

[23] Obaje N. G. (2009). Geology and Mineral Resources of Nigeria. Lecture Notes in Earth Sciences. Published by Springer Dordrecht Heidelberg, New York.

[24] Olasehinde P. I. (2010). The Groundwater of Nigeria: A Solution to Sustainable National Water Needs. Inaugural Lecture Series 17, Federal University of Technology, Minna, Nigeria.

[25] Oluseyi, T., Olayinka, K., and Adeleke, I. (2011). Assessment of ground water pollution in the residential areas of Ewekoro and Shagamu due to cement production. *African Journal of Environmental Science and Technology* 5(10), 786-794.

[26] Robertson G.P. (1987). Geostatistics in ecology: interpolating with known variance. Ecology, 68, 744-748.

[27] Sarukkalige, R. (2012). Geostatistical Analysis of Groundwater Quality in Western Australia, IRACSTEngineering Science and Technology: An International Journal (ESTIJ), Vol. 2, No. 4, pp. 790-794.

[28] Sha'Ato, R., Akaahan, T.J., & Oluma, H.O.A. (2010). Physico-chemical and bacteriological quality of water from shallow wells in two rural communities in Benue State, Nigeria. *Pakistan Journal of Analytical and Environmental Chemistry* 11(1), 73-78.

[29] Shyu, G. S., Cheng, B. Y., Chiang, C. T., Yao, P. H., & Chang, T. K. (2011). Applying factor analysis combined with kriging and information entropy theory for mapping and evaluating the stability of groundwater quality variation in Taiwan. *International Journal of Environmental Research andPublic Health, 8*(4), 1084–1109. doi:10.3390/ijerph8041084.

[30] Singh K.P., Malik A., Mohan D., and Sinha S. (2004). Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti river (India): A case study. Water Res. 38:3980–3992.

[31] Taany R. A., A B Tahboub and G. A. Saffarini (2009). Geostatistical analysis of spatiotemporal variability of groundwater level fluctuations in Amman–Zarqa basin, Jordan: a case study. *Environ. Geol,* 57, pp. 525–535.

[32] UNEP/ WHO, (1996). Water quality Monitoring, *A practical guide to the design and implementation of freshwater quality Studies and Monitoring Programmes* edited by Jamie Bartram and Richard Balance. Printed in Great Britain by TJ Press (Padstow) Ltd, Padsow, Cornwell.

[33] Vieira R.J., Hartfield L.J., Nielsen D. R. and Biggar W.J. (1982). Geostatistical Theory and Application to Variability of some Agronomical Properties, Hilgardia 51,3.

[34] Webster, R. and Oliver, M. (2001). Geostatistics for Environmental Scientists, John Wiley & sons, Ltd.

[35] WHO. (2011). *Guidelines for drinking-water quality. vol 4, 3rd edn.* Geneva: World Health Organization.