

ANALYSIS OF GROUNDWATER QUALITY PARAMETERS IN NIGER STATE USING ORDINARY KRIGING AND COKRIGING APPROACHES

Aminu Kabiru Hassan*, Audu Isah, Abdullahi Usman

Department of Statistics, School of Physical Sciences, Federal University of Technology, Minna, Niger State, Nigeria

*Corresponding Author Email Address: ahassankabiru@gmail.com

ABSTRACT

Groundwater is the primary source of water for households in Niger State. However, increasing urbanisation, poor sanitation, and limited monitoring have raised concerns regarding its quality and safety. This study used data on five Ground Water Quality (GWQ) parameters from twenty locations, which comprised five Boreholes (BH) and fifteen Hand-Dug Wells (HDW). Descriptive statistics were used to measure microbial, physical, and chemical parameters in accordance with the World Health Organization (WHO) standard. Geospatial structures in the data were estimated using ordinary kriging and cokriging. The results showed that, in BH, descriptive statistical analysis of *Escherichia coli* (0 cfu/100mL) concentrations clustered at zero, indicating good water quality. HDW showed a high *E. coli* count (12 cfu/100mL), indicating contaminated water. BH recorded very low microbial contamination with physical and chemical parameters within WHO limits. While HDW exhibited elevated microbial loads (1–56 cfu) and acidic pH (5.80–6.40), these conditions indicate that the water is not safe for direct consumption, despite acceptable Total Dissolved Solids (260–390 mg/L). Boreholes showed moderate spatial dependence ($Z_0 = 0.05$ – 0.15 , $Z = 0.60$ – 1.85 , and $a = 1.20$ – 2.80 km), implying that GWQ changes in a similar pattern. The spherical model best fitted the boreholes, while the spherical and exponential models best fitted the wells. Cross-variogram confirmed pH as the preferred auxiliary variable. The coefficient of determination confirmed that Cokriging outperformed ordinary kriging with ($R^2 \geq 0.96$). These results imply that boreholes represent the safest groundwater source in Minna. The study recommended routine monitoring, sanitation, and use of prediction maps to prioritise targeted interventions.

Keywords: Groundwater quality, Variogram modelling, Ordinary Kriging, Cokriging, and spatial prediction..

INTRODUCTION

Groundwater (GW) is the main source of domestic and agricultural water in Niger State, Nigeria. Still, its quality is increasingly threatened by rapid urbanization, poor sanitation, and intensified agricultural and industrial activities. These pressures contribute to microbial, physical, and chemical contamination, affecting soil productivity, crop yields, and public health through waterborne diseases. The socio-economic burden is particularly severe where alternative water sources and treatment infrastructure are limited. In rural and peri-urban Nigeria, GW is mainly accessed through Hand-Dug Wells (HDW) and boreholes; however, HDW are more vulnerable to surface contamination due to their shallow depth and open structure, while boreholes are generally deeper and relatively better protected. Despite this, both sources in Niger State remain

exposed to contamination from seasonal runoff, waste disposal practices, and geological conditions, and are further compounded by weak monitoring systems and limited access to routine water quality assessments (Matheron, 1963; Goovaerts, 1997; Environmental Protection Agency 2009; Chiles and Delfiner, 2012; Bisiriyu *et al.*, 2020; Odeyemi *et al.*, 2021; Al-Fatlawy *et al.*, 2022; Goovaerts, P. 2023; Kanu *et al.*, 2023; and Egbeyale *et al.*, 2025) Groundwater quality (GWQ) is a natural variable across space and time, requiring systematic assessment of microbial, physical, and chemical parameters to determine suitability for use. Traditional monitoring approaches are often insufficient for capturing this variability, particularly in areas with sparse sampling networks. Geostatistical methods address this limitation by estimating values at unsampled locations based on spatial dependence, assuming that closer points are more similar than distant ones (Goovaerts, 1997; Isaaks *et al.*, 1989; Goovaerts, 2000; Mehrjardi *et al.*, 2008; Isah and Abdullahi, 2015; WHO, 2017).

HDWs are shallow (typically <15 meters), open at the surface, and often susceptible to microbial and chemical contamination due to proximity to surface pollutants and inadequate sanitation infrastructure. In contrast, BH are deeper (20–300 meters), mechanically drilled, and generally better protected from surface infiltration, rendering them less prone to biological contamination (EPA 2009; Tijani *et al.*, 2005). In Niger State, GWQ has emerged as a pressing concern. BH and HDW constitute the principal sources of domestic supply, yet they are highly vulnerable to microbial and physicochemical Contamination. The hydrogeological setting, underlain by sedimentary and basement complex rocks, combined with seasonal variability, agricultural runoff, indiscriminate waste disposal, and rapid urban expansion, heightens the risk of contamination. Limited access to routine monitoring, owing to high costs and logistical barriers, further complicates groundwater resource management (Goovaerts, 2000; Isah and Abdullahi, 2013; Suleiman *et al.*, 2020; Goovaerts, 2023). One of the main global sources of fresh water for drinking and other important domestic purposes is groundwater. Access to potable water in developing nations worldwide is increasingly lacking (Suleiman *et al.*, 2020).

OK is a widely used geostatistical technique that provides best linear unbiased estimates by modelling spatial autocorrelation in groundwater quality data (Cressie, 1993; Goovaerts, 2023). CK extends this approach by incorporating secondary, spatially correlated variables to improve prediction accuracy, especially where such variables are more densely sampled. It reduces estimation error by exploiting cross-correlation structures between parameters (Isaaks and Srivastav, 1989; Singh and Verma,

2019; Webster & Oliver, 2007; Hooshmand, *et al.*, 2011a; Chile and Delfine, 2012; WHO, 2017 and Sarani, *et al.*, 2025) Together, these methods, combined with multivariate statistical tools, enhance the identification of pollution sources, improve spatial prediction, and support more effective groundwater resource management in contaminated or data-scarce environments.

(Isaaks and Srivastav, 1989; Miller, 2007; Isah and Abdullahi, 2015; Al-Fatlawy *et al.*, 2022). revealed deteriorating surface and groundwater quality in Nigeria due to chemical and biological pollution and seasonal changes, among others. As water quality issues become more serious and widespread, the need for water quality monitoring as an important component of health promotion strategy in developing countries cannot be overemphasized. A considerable number of researchers have shown increased interest in using multivariate statistical tools and geostatistical techniques to achieve the sustainable exploitation of water resources (Hooshmand *et al.*, 2011a; Al-Fatlawy *et al.*, 2022). The

combined use of multivariate statistics and geostatistical techniques enables the identification of potential sources affecting water environmental systems. It provides a valuable tool for reliable management of water resources, as well as a rapid solution to pollution issues (Miller, 2007; Hooshmand *et al.*, 2011b; Isah & Abdullahi, 2015)

MATERIALS AND METHODS

Study Area

The study was conducted in Minna and its environs, located in Niger State, the North-Central Geopolitical zone of Nigeria (latitudes 8°20'N–11°30'N, longitudes 3°30'E–7°20'E), with an estimated land area of about 76,363 km².

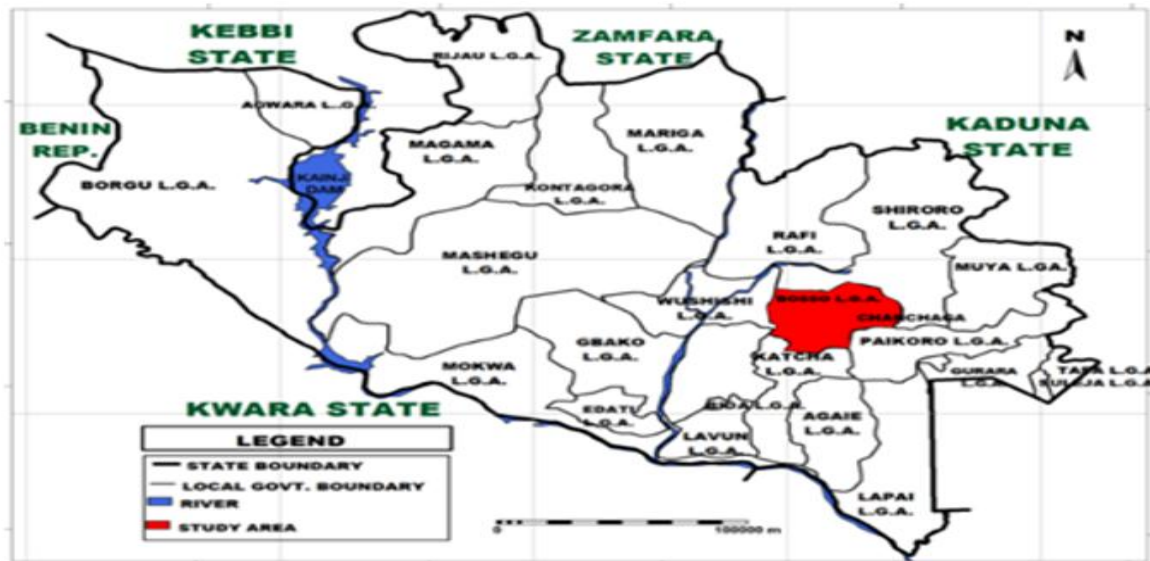


Figure 1: Groundwater quality study area in Niger state.

Sampling Techniques

A Stratified Random Sampling (SRS) and Purposive Sampling (PS) techniques were adopted in this study. The SRS is given by:

$$s_h = \left(\frac{N_h}{N} \right) \times s \quad (1)$$

where: s_h is the sample size for stratum h , N_h is the population size for stratum h , N is the total population size, and s is the total sample size.

The PS selects samples based on specific study objectives. It does not use a formula to select an individual sample. The study area was divided into strata (Residential and Agricultural zones), from which samples were randomly selected, while key sites at risk of contamination were deliberately included. This approach improves good representation and reduces sampling bias (Creswell & Creswell, 2018).

Laboratory Procedures

Groundwater samples were collected using sterile methods, with boreholes purged prior to sampling. Microbial parameters (*E. coli*,

F. coli, and *T. coli*) were preserved at 4°C and analyzed within 6 hours using membrane filtration, as described by the American Public Health Association (APHA, 2017). pH was measured at the site using a calibrated portable digital pH meter (handheld device) equipped with a glass electrode. The meter is calibrated with standard buffer solutions, and TDS is determined with a calibrated meter. Results were compared with the World Health Organization (WHO) standards (2017).

Data sources

The data utilised in this study comprised five key groundwater quality parameters, namely *E. coli*, *F. coli*, *T. coli*, TDS, and pH, which were compared with WHO drinking water standards. The study area covered Minna and its environs in Niger State, with a population increase from approximately 513,000 in 2024 to about 532,000 in 2025, representing an estimated growth rate of 3–4%. The dataset was compiled from field sampling and laboratory analyses conducted by the Department of Geology, Federal University of Technology, Minna, as well as published geostatistical

studies, applied geostatistical techniques to analyse spatial variability of water quality parameters in rivers and streams across Niger State using variogram modelling and ordinary kriging and theoretical basis for incorporating secondary variables in spatial estimation was informed by research work, which demonstrated that cokriging reduces estimation variance relative to ordinary

kriging by exploiting spatial cross-correlation between primary and secondary variables (Bisiriya, *et al.*, 2020; Isah & Abdullahi, 2015). These parameters were selected for their strong relevance to microbial contamination and the physical and chemical assessment of groundwater quality in urban and peri-urban environments.

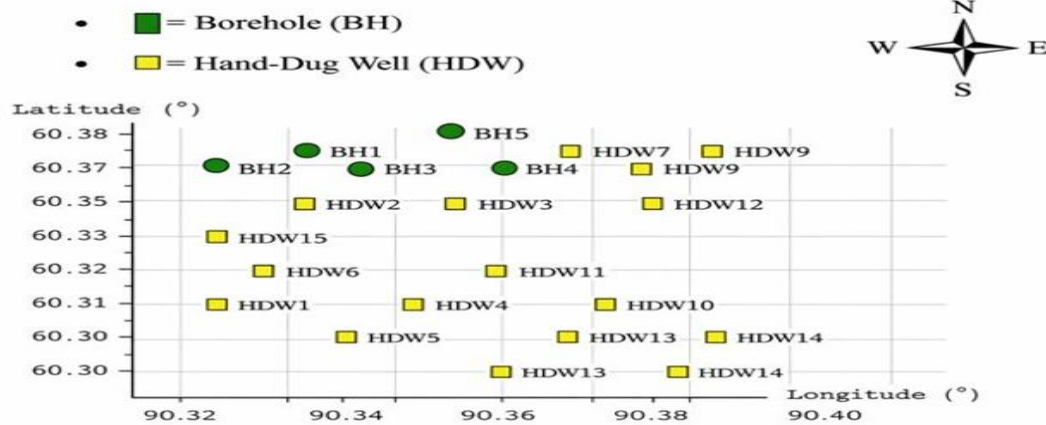


Figure 2: Location of the boreholes and hand-dug wells

Figure 2 displays the location of BH and HDW, where green symbols denote BH mechanically drilled water points and yellow

symbols represent h HDW. The BH are clustered, while the HDW are more dispersed across the study area.

DATA PRESENTATION

Table 1: Original Data of Ground Water Quality Parameters (GWQP)

Locations	Sources	Elev. (m)	Depth (m)	Water Lev(m)	E. Coli (cfu/100ml)	Faecal Col (cfu/100ml)	Total Col (cfu/100ml)	TDS (mg/L)	pH
Maitumbi	BH1	328.5	60.00	51.0	0	0	3	220	7.1
New-Gurusu	BH2	339.5	65.00	27.0	0	1	5	230	6.8
Old-Gurusu	BH3	339.5	27.00	25.0	0	0	2	210	7.0
Ruga Fulani	BH4	301.5	70.00	16.5	2	6	18	280	6.3
Shakauta	BH5	355.5	60.00	25.0	0	0	4	215	7.2
Bosso Area 11	HDW1	251.5	3.80	4.8	12	25	55	340	5.9
Chanchaga	HDW2	229.5	3.80	4.0	8	19	43	370	6.0
Dana-Pharmaceutical	HDW3	293.5	3.50	6.4	4	11	30	295	6.2
Duse-Kura Gwari	HDW4	254.5	2.50	4.2	10	18	40	310	5.8
FGC Tunga	HDW5	222.5	1.00	4.0	3	8	25	300	6.4
Govt. Day Sec. School	HDW6	250.5	3.00	5.0	6	13	33	325	6.3
Gwadeiyi	HDW7	289.5	1.50	3.0	9	22	50	355	6.1
Kinkapa	HDW8	293.5	35.00	7.0	2	5	20	270	6.7
Saukakauta	HDW9	298.5	7.80	12.0	1	3	12	260	6.9
Shanu	HDW10	247.5	4.00	4.0	7	15	38	330	6.0
Shapata	HDW11	312.5	2.40	4.3	4	9	28	310	6.5

Tudunwada -Minna	HDW12	277.5	2.00	2.8	6	14	34	305	6.1
Tunga- Alade	HDW13	219.5	0.6	1.5	11	24	56	390	5.8
Tunga- Goro	HDW14	206.5	2.5	5.3	5	10	29	315	6.2
Tunga Lowcost	HDW15	268.5	1.5	3	8	16	36	340	6.0

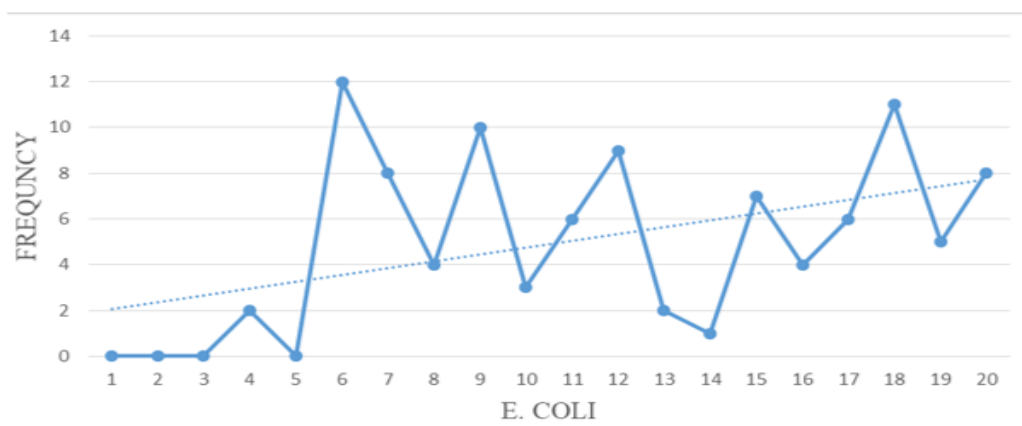


Figure 3: Stationarity assessment of original Data

Figure 3 showed that E. coli concentrations across sampled locations exhibited irregular spatial variability, with no consistent pattern of increase or decrease, indicating non-stationarity.

Detrending is therefore required before stationarity testing and spatial modeling.

Table 2: Detrended Data of Ground Water Quality Parameters (GWQP)

Locations	Sources	Elev. (m)	Depth (m)	Water Lev(m)	E. Coli (cfu/100ml)	Faecal Col (cfu/100ml)	Total Col (cfu/100ml)	TDS (mg/L)	pH
Maitumbi	BH1	328.50	60	51	0.9	0.9	-0.1	4.5	-0.1
New-Gurusu	BH2	339.5	65	27	-0.8	-1.3	-0.1	-8.6	-0.2
Old-Gurusu	BH3	339.5	27	25	-0.5	-0.1	0.5	3.1	0.4
Ruga Fulani	BH4	301.5	70	16.5	-0.4	0.7	-0.5	1.6	0.1
Shakauta	BH5	355.5	60	25	0.8	-0.1	0.1	-0.6	-0.1
Bosso Area 11	HDW1	251.5	3.8	4.8	1.5	3.2	6.4	12.0	-0.3
Chanchaga	HDW2	229.5	3.8	4.0	-0.8	-1.1	-3.5	18.0	-0.2
Dana-Pharmaceutical	HDW3	293.5	3.5	6.4	-1.2	-2.0	-1.5	-25.0	0.1
Duse-Kura Gwari	HDW4	254.5	2.5	4.2	0.6	-0.5	-2.4	-10.0	-0.4
FGC Tunga	HDW5	222.5	1.0	4.0	-2.1	-1.8	-0.6	-15.0	0.2
Govt. Day Sec. School	HDW6	250.5	3.0	5.0	-0.7	0.6	1.8	5.0	0.1
Gwadeiyi	HDW7	289.5	1.5	3.0	2.3	4.8	5.5	20.0	-0.1
Kinkapa	HDW8	293.5	35	7.0	-3.0	-4.5	-6.0	-35.0	0.3
Saukakauta	HDW9	298.5	7.8	12	-2.5	-3.2	-5.1	-28.0	0.4
Shanu	HDW10	247.5	4	4	-0.2	-0.8	-2.2	-5.0	0.0
Shapata	HDW11	312.5	2.4	4.3	-1.8	-2.1	-0.5	-12.0	0.2
Tudunwada-Minna	HDW12	277.5	2.0	2.8	0.9	1.4	2.6	8.0	-0.2
Tunga-Alade	HDW13	219.5	0.6	1.5	3.1	5.2	7.0	30.0	-0.5
Tunga-Goro	HDW14	206.5	2.5	5.3	-1.4	-2.0	-1.2	-10.0	0.1
Tunga Lowcost	HDW15	268.5	1.5	3.0	0.6	0.9	1.4	15.0	0.0

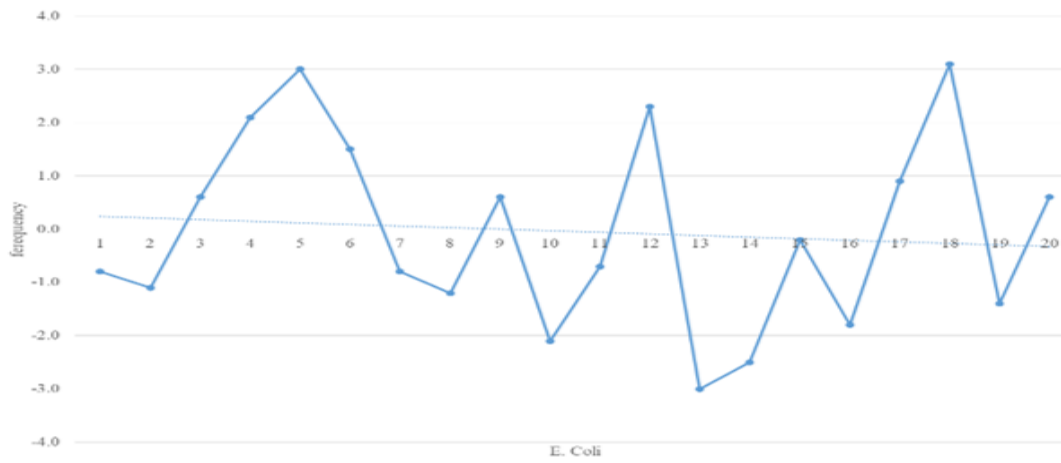


Figure 4: Stationarity assessment of detrended data

Figure 4 presents the stationarity assessment of detrended data. The series oscillates about a constant mean with no discernible trend. Variability remains approximately constant throughout the locations. These features satisfy the conditions of stationarity. The series is suitable for variogram modeling and kriging-based spatial analysis.

METHODS

Variogram Models

In the variogram, 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, Isah and Abdullahi (2013). Experimental variograms for E. coli, T. coli, TDS, and pH were calculated at different lag distances. The models of spatial variability were fitted to the experimental variogram by optimizing the model parameters to achieve the best fit, as indicated by maximizing the coefficient of determination (R^2) between the observed and modeled values. (Goovaerts, 1997; Goovaerts, 2000, and Hooshmand *et al.*, 2011b). Some important models are spherical, exponential, and Gaussian models.

The spherical model:

$$\gamma(h) = Z_0 + Z \left(1.5 \left(\frac{a}{b} \right) - \frac{1}{2} \left(\frac{a}{b} \right)^3 \right) \quad \text{for } 0 < a \leq b \quad (2)$$

where: $\gamma(h)$ Semivariance at distance h , quantifying the difference in groundwater quality parameters (E. coli, F. col, T. Col, TDS, and pH) between two points, Z_0 is the Nugget effect, representing small-scale variability or measurement errors in water quality monitoring, Z (sill) structural variance (partial sill), representing the variance explained by spatial dependence, a (range) practical range, the distance at which approximately 95% of the sill is reached, indicating the extent of spatial correlation of groundwater quality parameters, and h (the lag distance): The separation distance between two sampling locations in the study area.

The Exponential model:

$$\gamma(a) = Z_0 + Z \left(1 - e^{-\left(\frac{a}{b}\right)} \right) \quad (3)$$

The parameters are as in equation (2) above.

Gaussian model:

$$\gamma(a) = Z_0 + Z \left[1 - e^{-\left(\frac{a}{b}\right)^2} \right] \quad (4)$$

The parameters are also as in equation (2) above.

Semivariograms

The semivariogram is central to geostatistical theory, enabling the quantification of spatial dependence and guiding the choice of interpolation models. This technique is applied when there is irregular variation in attributes and the sample density is such that simple interpolation methods may produce reliable predictions. This method provides a probabilistic estimate of interpolation quality not only for the sample numbers and sampling patterns but also for the assumptions made in the geostatistical methods, which are included in the variable's built-in structure. This spatial structure is characterized by spatial autocorrelation and the semivariograms. Using the semivariograms, ordinary kriging and Cokriging are used to interpolate at unknown points (Goovaerts, 1997; Goovaerts, 2000, and Hooshmand *et al.*, 2011b).

Correlation analysis

Correlation analysis measures the strength and direction of the relationship between WQP, helping to identify shared contamination sources and key interactions. It supports the selection of suitable auxiliary variables, improves the accuracy of geostatistical modeling, and informs effective groundwater monitoring and management (Wali *et al.*, 2024). The formula is expressed in equation (5)

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ki} - \bar{x}_i)^2 \sum_{k=1}^n (x_{kj} - \bar{x}_j)^2}} \quad (5)$$

Where: r_{ij} Denotes the correlation coefficient between parameter i and parameter j ; n is the number of groundwater samples; x_{ki} and x_{kj} represent observed values of the i^{th} and j^{th} parameters at sample k ; while \bar{x}_i and \bar{x}_j are their respective mean values. The correlation coefficient ranges from ± 1 (Zero, perfect negative, and positive correlations).

General Kriging theory

Kriging is a method for spatial, temporal, or spatio-temporal prediction and for estimating unknown local values of variables distributed in one, two, or three dimensions from relatively sparse

data. It is an exact linear interpolating procedure based on a spatial linear model that specifies a parametric spatial mean function and a spatial dependence structure [15 and 27]. The fundamental assumption of Kriging is that the data are derived from a stationary stochastic process, and some variants require normally distributed data. Unlike other methods, such as inverse distance weighting (IDW), the Kriging weight function is not arbitrary. However, it is calculated from the parameters of the fitted semivariogram model to ensure unbiasedness and minimize estimation variance. The Kriging approach models the variability in the data as a function of the variogram (Goovaerts, 1997). The general Ordinary Kriging model is expressed as:

$$Z(s_o) = \sum_{i=1}^n \lambda_i Z(s_i) + \varepsilon(s) \quad (6)$$

Where: $Z(s_o)$ is the groundwater quality at an unsampled location, $Z(s_i)$ is observed values at BH and HDW, λ_i is the Kriging weights that satisfy the unbiasedness condition and minimise estimation variance and ensure reliable spatial predictions.

Ordinary Kriging (OK)

Ordinary Kriging is used to model the spatial variability of each of the five influential parameters and to estimate them at sampled locations. It is based on the concept of a variable $Z(x_i)$ that is both random and spatially autocorrelated [10, 18 and 30]. The predictions are based on equation (5) as:

$$Z(s_i) = \mu + \varepsilon(s_i) \quad (7)$$

Where: $Z(s_i)$ is the observed groundwater quality parameters at location s , μ is the constant but unknown global mean (stationary function, and $\varepsilon(s_i)$ represents the spatially correlated stochastic component.

Cokriging Model

Cokriging is also an interpolation technique that relies on the same premise as kriging; the technique uses autocorrelation and cross-correlation to create an interpolated surface that predicts values at unmeasured locations. Additionally, cokriging includes a secondary variable in the interpolation model (Queiroz et al., 2008), and assumes that some autocorrelation exists between the primary and secondary variables; the stronger the autocorrelation among the multivariate results, the greater the accuracy of predicting the primary variable in the cokriging model (Sarani et al., 2025). Using two or more variables, cokriging produces the best model by eliminating bias between estimated and true values and minimizing the variance of estimates. Using CK, the accuracy of the interpolated surface can be increased when compared to univariate kriging. In this study, pH was employed as a secondary variable in the cokriging model due to its significant spatial and hydrogeological influence on E. coli contamination in groundwater. Distance captures the horizontal attenuation of microbial concentrations with increasing separation from contamination sources. The inclusion of auxiliary variables strengthens the CK estimator by integrating both spatial structure and subsurface characteristics, thereby improving the accuracy and reliability of the estimated values (Hooshmand et al., 2011b; Goovaerts, 2000; Isaaks & Srivastava, 1989; Queiroz et al., 2008; Yalcin, 2005). The CK model is expressed in equation (7) as;

$$Z(s_o) = \sum_{i=1}^n a_i \cdot Z_i(s_i) + \sum_{j=1}^m b_j \cdot Z_j(s_j) + \varepsilon(s) \quad (8)$$

Where: $Z(s_o)$ is the parameter concentration (cfu/100 ml) at an unsampled groundwater location u_i , $Z_i(s_i)$ is the observed concentration at the i^{th} sampled locations s_i (primary variable),

$Z_j(s_j)$ is the Observed secondary information at the j^{th} neighbouring location s_j , a_i and b_j are CK weights assigned to the primary and secondary variable, n and m are the number of primary and secondary observations, and $\varepsilon(s)$ the error term component. The development of the cokriging system is identical to the development of the ordinary kriging system. The variance of the estimation error is expressed in terms of cokriging weights and covariance between the variables [30], and is expressed in equation (8) as:

$$\begin{aligned} \text{Var}(R_{gw}) = & \sum_{i=1}^n \sum_{j=1}^m a_i a_j \text{Cov}(S_i S_j) + \\ & \sum_{i=1}^n \sum_{j=1}^m b_i b_j \text{Cov}(U_i U_j) + \sum_{i=1}^n \sum_{j=1}^m a_i b_j \text{Cov}(S_i U_j) - \\ & 2 \sum_{i=1}^n a_i \text{Cov}(S_i U_j) - 2 \sum_{i=1}^n b_j \text{Cov}(S_j S_o) + \\ & \text{Cov}(U_j S_o) + \text{Cov}(S_o S_o) \end{aligned} \quad (9)$$

Where $\text{Var}(R)$ is the estimation variance of the CK predictor for the groundwater quality parameter at the target (unsampled) location, n and m are number of sampled observations of the primary groundwater quality parameter S and secondary variable U . $\text{Cov}(S_i S_j)$ Spatial covariance between primary GWQP values at locations i^{th} and j^{th} , $\text{Cov}(U_i U_j)$ Spatial covariance between secondary variable values at locations i^{th} and j^{th} , $\text{Cov}(S_i U_j)$ Cross-covariance between the primary GWQP and the secondary variable at locations i^{th} and j^{th} , $\text{Cov}(S_i S_o)$ Covariance between the primary GWQP at sampled location i^{th} and prediction location, $\text{Cov}(U_j S_o)$ Cross-covariance between the secondary variable at sampled location, $\text{Cov}(S_o S_o)$ Variance of the primary GWQP at the prediction location, equal to the sill of the variogram of S , and a_i and b_i are the kriging weights for primary and secondary GWQP samples.

The set of CK weights must satisfy two conditions: First, the weights must be such that the estimate given in Equation is unbiased. Second, the weights must be such that the error variances given in Equation are the smallest possible (Yalcin, 2005; Goovaerts, 2007; and Hooshmand et al., 2011b). One way of guaranteeing unbiasedness is to ensure that the weights in the first term sum to 1 while those in the second sum to 0, as shown in equation (10):

$$\sum_{i=1}^n a_i = 1 \text{ and } \sum_{j=1}^m b_j = 0 \quad (10)$$

where: $a_1 + \dots + a_n = 1$ and $b_1 + \dots + b_n = 0$.

Model Cross-Validation Metrics

The predictive performance of geostatistical interpolation models (Ordinary Kriging and Cokriging model), were evaluated by various researchers using cross-validation metrics: the coefficient of determination (R^2) to quantify variance explained by model, Mean Error (ME) to assess bias, Mean Absolute Error (MAE) to measure prediction accuracy, and the Root Mean Squared Standardised Error (RMSSE) to verify that the kriging variance derived from the fitted variogram accurately reflected observed errors, thereby ensuring the spatial structure and correlation were properly captured, as demonstrated by [8 and 29]. The Mean Absolute Error is the identification of models producing largely unbiased estimates,

The Mean Absolute Error [10];

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |Z(x_i) - \hat{Z}(x_i)| \quad (11)$$

Where: MAE is the average magnitude of prediction errors in groundwater quality estimates, $Z(x_i)$ is the observed concentration of a water quality parameter (E. coli, F. col, T. col,

TDS, or pH) at sampling point x_i , $\hat{Z}(x_i)$ is the Predicted concentration of unsampled location and n is the number of sampling locations.

The Root Mean Squared Error (Goovaerts, 2007; Goovaerts, 2023).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n [Z(s_i) - \hat{Z}(s_i)]^2} \quad (12)$$

RMSE Reflects the average size of the errors in predicted microbial water quality values (E. coli, faecal coliform counts), $Z(s_i)$ is the true measured microbial concentration (OCFU/100ml) at a sampling location x_i , $\hat{Z}(s_i)$ is the predicted microbial concentration at the same sampling location, calculated using spatial interpolation methods (kriging or cokriging), n is the total number of groundwater sampling points used for model validation, and $[Z(x_i) - \hat{Z}(x_i)]^2$ is the square of the difference between observed and predicted concentrations, ensuring larger differences contribute more to the overall error.

The Coefficient of determination is given by equation (13):

$$R^2 = 1 - \frac{\sum(Z_i - \hat{Z}_i)^2}{\sum(Z_i - \bar{Z})^2} \quad (13)$$

R^2 Describes how well the predicted microbial concentrations match the observed measurements across the sampling locations (A high R^2 suggests the model effectively predicts contamination risks at unsampled location), Z_i is the true measured concentration of a WQP at the water sampled point, \hat{Z}_i is the estimated value by the spatial model at the same location, \bar{Z} is the average measured microbial concentration across all sampled locations, and $(Z_i - \hat{Z}_i)^2$ is total squared difference between observed and predicted

microbial concentrations, represents the unexplained variance in the model predictions

R Geostatistics Packages

This study introduces the functionality of five (5) R geostatistics packages used for processing and displaying results: gstat, sp, rgdal, spatstat, and maptools. All of these are available as open-source software 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) variogram; 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; spatstat 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.

RESULTS AND DISCUSSION

Results

Descriptive statistics

The descriptive statistics of borehole water samples and hand-dug well water samples obtained from various locations within the Minna and its environs were presented in Tables 3 and 4, respectively.

Descriptive Statistics results for Five Boreholes

Table 3: Descriptive Statistics results for Five Boreholes

Statistic	E.coli (cfu/100ml)	F.col (cfu/100ml)	T.col (cfu/100ml)	TDS (mg/L)	pH
Total	15.00	15.00	15.00	15.00	15.00
Minimum	1.00	3.00	12.00	260.00	5.80
Maximum	12.00	25.00	56.00	390.00	6.90
Mean	6.40	14.13	35.27	321.00	6.19
Variance	10.83	44.27	152.35	1222.14	0.10
SD	3.29	6.65	12.34	34.96	0.32
Skewness	0.07	0.08	0.09	0.18	0.80
Kurtosis	-0.99	-0.94	-0.52	-0.37	-0.17
25th Percentile	4.00	9.50	28.5	302.50	6.00
Median	6.00	14.00	34.00	315.00	6.10
75th Percentile	8.50	18.50	41.50	340.00	6.35

Table 3 shows the descriptive statistics results and indicates that BH generally has low to moderate microbial contamination and stable physical and chemical conditions. E. coli ranged from 0.00 – 2.00cfu/100 ml (mean = 0.40, median = 0.00 and SD = 0.89), F.col from 0.00–6.00 cfu/100 ml (mean = 1.40 and SD = 2.61), indicated mostly non-detects but occasional faecal contamination and T.col levels were higher 2–18 cfu/100 ml (mean = 6.40, median = 4.00 and SD = 6.58), suggested higher microbial presence. TDS concentrations were 210.00–280.00 mg/L (mean = 231 mg/L; SD

= 28.37), and pH values were near-neutral, 6.30–7.20 (mean = 6.88; SD = 0.36); both were within WHO drinking water guidelines, indicating acceptable physical and chemical quality. The detection of microbial parameters above the WHO guideline of 0.00 cfu/100 ml in drinking water indicated that the water is not safe for direct human consumption without treatment. While physical and chemical parameters support suitability for general domestic use, microbiological results necessitate appropriate treatment before human consumption.

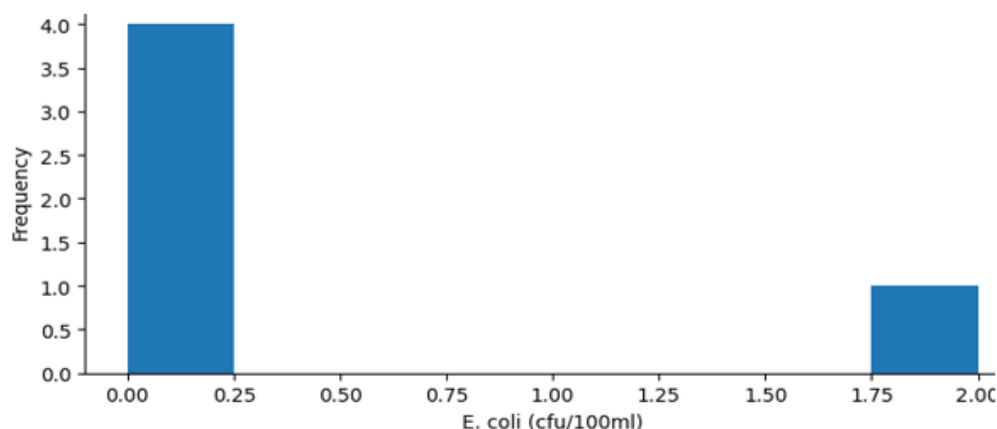


Figure 5: Histogram of E. coli (cfu/100 ml) in Boreholes

Figure 5 presents the E. coli contamination concentrations in BH, with a strong cluster at zero, and only one sample recorded a value above 1 cfu/100 ml. This distribution is yielding a low mean (0.40 cfu/100 ml), variance ($\sigma^2 = 0.80$), highly right-skewed, as confirmed by the positive skewness value (SK = 1.25) and kurtosis (K = 0.25), indicating a peaked distribution with heavy tails, reflecting the dominance of very low values and the presence of rare higher observations. The mean concentration of 0.40 cfu/100

ml and median of 0.00 cfu/100 ml further confirm that most BH are microbiologically safe, with minimal faecal contamination. The low variance (0.80) and standard deviation (0.89) suggest limited variability, implying relatively uniform water quality across borehole locations. The histogram indicates that borehole water in the study area is largely uncontaminated by E. coli, consistent with deeper groundwater sources being better protected from surface pollution.

Descriptive Statistics for Fifteen Han-dug Wells

Table 4: Descriptive Statistics for Fifteen Han-dug Wells

Statistic	E.coli (cfu/100ml)	F.col (cfu/100ml)	T.col (cfu/100ml)	TDS (mg/L)	pH
Total	5.00	5.00	5.00	5.00	5.00
Minimum	0.00	0.00	2.00	210.00	6.30
Maximum	2.00	6.00	18.00	280.00	7.20
Mean	0.40	1.40	6.40	231.00	6.88
Variance	0.80	6.80	43.30	805.00	0.13
SD	0.89	2.61	6.60	28.37	0.36
Skewness	1.50	1.40	1.40	1.26.00	-0.93
Kurtosis	0.25	0.10	0.11	-0.09	-0.55
25th Percentile	0.00	0.00	3.00	215.00	6.80
Median	0.00	0.00	4.00	220.00	7.00
75th Percentile	0.00	1.00	5.00	230.00	7.10

Table 4 shows the descriptive statistics for HDW samples, indicating elevated microbial contamination and moderate physicochemical variability. E. coli concentrations ranged from 1.00 to 12.00 cfu/100 ml, (mean = 6.40, median = 6.00 and SD = 3.29), F.col levels were similarly high, varying from 3 to 25 cfu/100 ml (mean = 14.13 and SD = 6.65) and also T.col concentrations were widespread (12.00–56.00 cfu/100 ml (mean = 35.27 cfu/100 ml; SD = 12.34), indicated persistent faecal contamination. The near-zero skewness values (0.07–0.09) implied approximately symmetric distributions, reflecting consistently elevated contamination across

the study area. TDS values ranged from 260 to 390 mg/L (mean of 321.00 and SD = 34.96), within the recommended guideline of ≤ 500 mg/L and pH values were slightly acidic to near neutral, 5.80–6.90 (mean = 6.19 and SD = 0.32), with some samples falling below the recommended limit, indicating acceptable mineralisation. The consistently high microbial concentrations exceed the drinking water guideline of 0.00 cfu/100 ml, rendering the water unsuitable for direct human consumption without appropriate treatment.

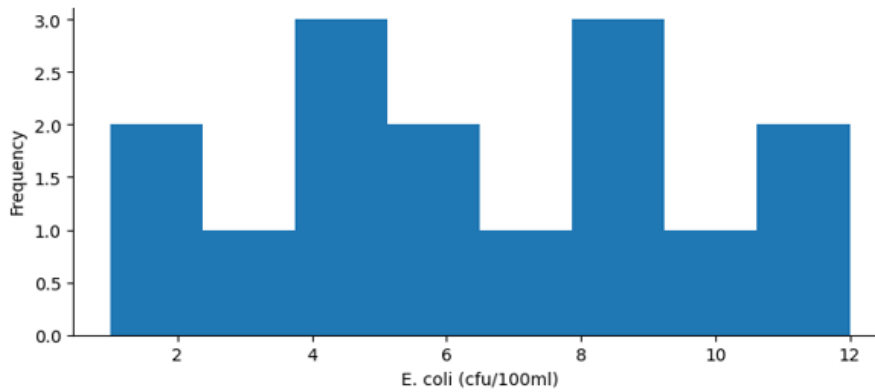


Figure 5: Histogram of E. coli (cfu/100 ml) in HDW

Figure 6 reveals that E. coli concentrations in HDW are consistently elevated, ranging from 1.00 to 12.00 cfu/100 ml (mean = 6.40, sigma squared = 10.83, and SD = 3.29). The distribution appears approximately symmetrical, as confirmed by the low skewness (SK = 0.07), and the negative kurtosis (K = -0.99) suggests a flatter

distribution. This implies that hand-dug wells are more vulnerable to surface-derived contamination, posing a greater public-health risk. This may be attributed to differences in well depth, construction quality, and proximity to pollution sources.

Boreholes and hand-dug wells parameters compared with the WHO (2017)

Table 5: Boreholes and hand-dug wells parameters compared with WHO (2017)

Location / Source	E.coli (cfu/100m)	F.col (cfu/100m)	T.col (cfu/100mL)	TDS (mg/L)	pH (mg)
WHO Standard	0 cfu/100 ml	0 cfu/100 ml	0 cfu/100 ml	≤500 mg/L	6.5 – 8.5
BH1	0.00	0.00	3.00	220.00	7.10
BH2	0.00	1.00	5.00	230.00	6.80
BH3	0.00	0.00	2.00	210.00	7.00
BH4	2.00	6.00	18.00	280.00	6.30
BH5	0.00	0.00	4.00	215.00	7.20
HDW6	12.00	25.00	55.00	340.00	5.90
HDW7	8.00	19.00	43.00	370.00	6.00
HDW8	4.00	11.00	30.00	295.00	6.20
HDW9	10.00	18.00	40.00	310.00	5.80
HDW10	3.00	8.00	25.00	300.00	6.40
HDW11	6.00	13.00	33.00	325.00	6.30
HDW12	9.00	22.00	50.00	355.00	6.10
HDW13	2.00	5.00	20.00	270.00	6.70
HDW14	1.00	3.00	12.00	260.00	6.90
HDW15	7.00	15.00	38.00	330.00	6.00
HDW16	4.00	9.00	28.00	310.00	6.50
HDW17	6.00	14.00	34.00	305.00	6.10
HDW18	11.00	24.00	56.00	390.00	5.80
HDW19	5.00	10.00	29.00	315.00	6.20
HDW20	8.00	16.00	36.00	340.00	6.00

Table 5 demonstrates significant differences in groundwater quality between BH and HDW across sample locations, as assessed against the World Health Organization (WHO) drinking-water guidelines. BH generally exhibits superior microbiological quality, with most recording zero E. coli and low faecal and total coliform counts, although occasional detections still contravene WHO zero-tolerance guidelines. HDW consistently shows high levels of E. coli, F. col, and T. col, indicates a higher vulnerability to faecal and environmental contamination, likely due to their shallow depths,

inadequate sanitary protection, and exposure to surface runoff and anthropogenic activities while Physical parameters (TDS) in both water sources remain within WHO limits. However, high concentrations were observed in HDW, indicating increased influence from surface inputs and local geological conditions. The chemical parameter (pH) generally falls within the WHO-recommended limits for BH, whereas HDW shows slightly acidic conditions below the WHO guidelines.

Variogram and cross-variogram parameters

Table 6: Variogram and cross-variogram parameters

Parameter	Source	Nugget (C_0)	Sill (C)	Range (a, km)	Model Type
E. coli	BH	0.15	1.85	1.20	Spherical
F. col		0.20	2.10	1.50	Spherical
T. col		0.25	2.60	1.80	Exponential
TDS		0.10	1.40	2.30	Spherical
pH		0.05	0.60	2.80	Gaussian
E. coli + TDS		0.12	1.50	1.80	Spherical
E. coli + pH	0.08	1.25	2.00	Spherical	
E. coli	HDW	0.20	2.10	1.30	Spherical
F. col		0.25	2.50	1.40	Exponential
T. col		0.30	2.80	1.60	Exponential
TDS		0.15	1.60	2.10	Spherical
pH		0.07	0.70	2.50	Gaussian
E. coli + TDS		0.14	1.60	1.90	Spherical
E. coli + pH	0.10	1.35	2.10	Spherical	

Table 6 reveals that the experimental variogram and cross-variogram parameters describe the geospatial structure of groundwater quality for BH and HDW, as well as the cross-relationships between E. coli and pH. Boreholes exhibited generally low nugget values ($Z_0 = 0.05 - 0.15$), moderate sill values ($Z = 0.60 - 1.85$) and relatively short to moderate ranges ($1.20 - 2.80m$), indicated weak short-range variability, limited microscale effects and with the spherical model providing the best overall description of spatial dependence (while a Gaussian model best fitted pH, HDWs showed higher nugget values ($0.07 - 0.30$), larger sills ($0.70 - 2.80$) and ranges ($1.30 - 2.50 km$), reflecting stronger micro-scale heterogeneity and greater influence of localised surface contamination with spatial structures best represented by spherical models (E. coli, TDS, pH) and exponential models for coliform indicators. demonstrate the value of auxiliary variables, with E. coli + TDS showing nugget, sill and range values of 0.12, 1.50 and 1.80 km in BHs and 0.14, 1.60 and 1.90 km in HDWs, and E. coli + pH exhibiting lower nuggets and longer ranges (BH: $Z_0 = 0.08, Z = 1.25, a = 2.00 km$; HDW: $Z_0 = 0.10, Z = 1.35, a = 2.10 km$), both best described by the spherical cross-variogram model.

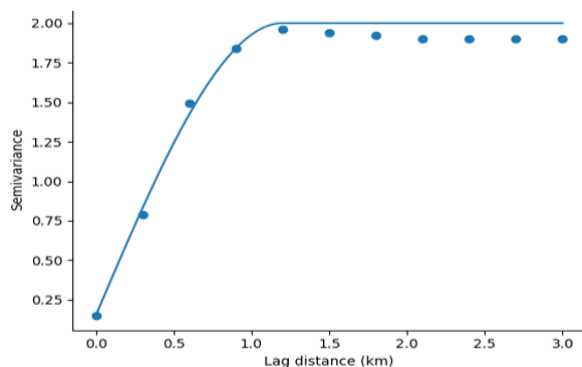


Figure 7: experimental variogram of E. coli in BH

Figure 7 reveals that the borehole E. coli experimental variogram exhibits a clear distance-dependent increase in semivariance, indicating a well-defined spatial structure. At short lag distances ($\leq 0.5 km$), low semivariance values reflect strong spatial continuity among neighbouring boreholes. A stable sill is reached at approximately 1.2 km, beyond which spatial dependence declines. The spherical model provided an adequate fit to the experimental points and is therefore considered the most suitable model for representing E. coli spatial variability in boreholes.

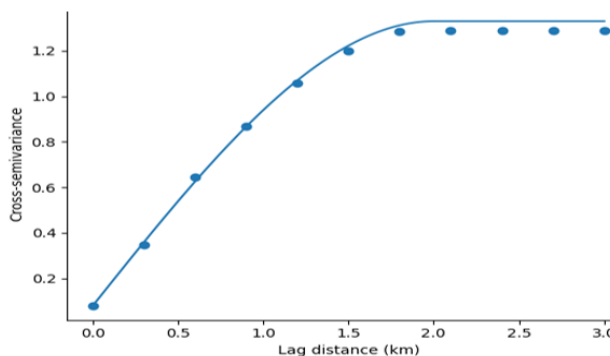


Figure 8: Cross-variogram of E. coli and pH in BH

Figure 8 shows that the E. coli + pH cross-variogram for boreholes displayed a consistent positive spatial dependence. Low cross-semivariance at short lag distances ($\leq 0.7 km$) indicates strong local coupling between microbial contamination and pH conditions. Cross-semivariance increases progressively with distance and stabilises at approximately 2.0 km. The spherical model best captures cross-variogram spatial interaction and supports using pH as a reliable auxiliary variable for cokriging E. coli in boreholes.

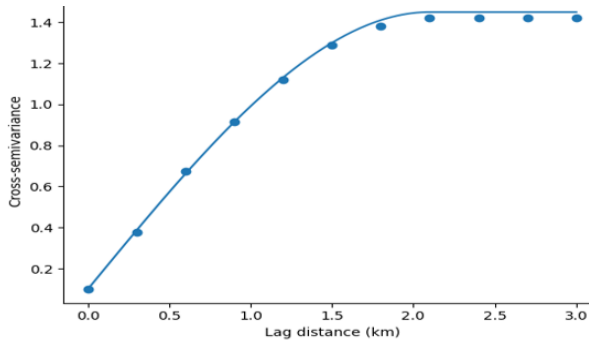


Figure 9: variogram of E. coli of HDW

Figure 9 shows that the HDW E. coli variogram exhibited distance-dependent spatial structure (Nugget = 0.10 and sill = 1.35), with semivariance increasing at short lag distances (≤ 0.60 km). This behaviour indicates strong local similarity followed by rapid spatial divergence. The sill is approximately 1.30 km, beyond which spatial dependence diminishes. Compared to BH, higher nugget and sill values indicate greater microscale variability linked to surface influence. The spherical model provides the best representation of E. coli spatial continuity in hand-dug wells.

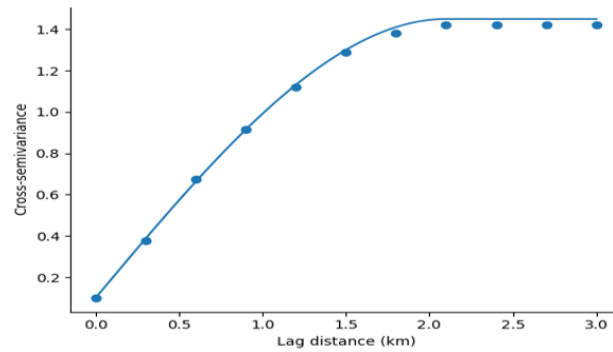


Figure 10: HDW cross-variogram of E. coli and pH

Figure 10 reveals that the cross-variogram (E. coli + pH) for HDW shows stable, positive spatial dependence (Nugget = 0.1, sill = 1.35). Low cross-semivariance at short lag distances (≤ 0.8 km) reflects strong local interaction, while increasing cross-semivariance with distance indicates weakening similarity. A stable sill is reached at approximately 2.1 km. The spherical cross-variogram model best fits the experimental data, confirming pH as an effective secondary variable for cokriging in hand-dug wells.

Correlation analysis results for Boreholes (BH) and Hand-dug wells

Table 7: Correlation analysis results for Boreholes (BH) and Hand-dug wells

Parameter	E. coli	F. Coli	T. Coli	TDS	pH
Boreholes					
E. coli	1.00(0.000)				
F. Coli	0.98 (0.005)	1.00 (0.000)			
T. Coli	0.96 (0.008)	0.99 (0.001)	1.00 (0.000)		
TDS	0.65 (0.237)	0.71 (0.176)	0.70 (0.185)	1.00 (0.000)	
pH	-0.69 (0.195)	-0.74 (0.155)	-0.72 (0.168)	-0.41 (0.492)	1.00 (0.000)
Hand-dug wells					
E. coli	1.00 (0.000)				
F. Coli	0.98 (<0.001)	1.00 (0.000)			
T. Coli	0.93 (<0.001)	0.95 (<0.001)	1.00 (0.000)		
TDS	0.86 (<0.001)	0.83 (<0.001)	0.80 (<0.001)	1.00 (0.000)	
pH	-0.91 (<0.001)	-0.88 (<0.001)	0.84 (<0.001)	-0.91 (<0.001)	1.00 (0.000)

Table 7 presents the correlation results with strong justification for selecting pH as an auxiliary variable in the analysis. In BH, there are strong positive and statistically significant relationships between microbial parameters ($r = 0.96$ to 0.99 , $p < 0.01$), indicating a similar water source. At the same time, pH exhibits moderate negative correlations and non-significant relationship with microbial parameters (r , equals, minus 0.69 , t_o , minus 0.74 , p , greater than, 0.05), suggest that lower pH is associated with high microbial presence. At the same time, DW displays a strong negative correlation and statistical significance between pH and microbial parameters ($r = -0.88$ to -0.91 , $p < 0.001$), implying that more acidic conditions strongly correspond to high levels of microbial contamination. The results confirm that pH is a strong auxiliary variable, given its statistically significant inverse

relationship with microbial contamination, particularly in HDW ($p < 0.001$). This indicates that high acidity is associated with higher pollution levels. Therefore, incorporating pH into geostatistical modeling strengthens the model, improves hotspot detection, and supports groundwater quality management decision-making to inform appropriate interventions.

Ordinary kriging and cokriging estimation results for Boreholes

To compare the estimation results of ordinary and cokriging, model validation techniques (Variance, Error, Absolute error, and R^2) were used to compare estimated and true values using only information available in the sample data set.

Ordinary Kriging and Cokriging estimation results for Boreholes

Table 8: Ordinary Kriging and Cokriging estimation results for Boreholes

Method	N	Mean	Standard Deviation	Minimum	Maximum
True E. coli	5.00	0.00	0.79	-0.78	0.94
Ordinary Kriging					
Estimated E. coli	5.00	0.01	0.77	-0.70	0.90
Kriging variance	5.00	0.02	0.01	0.02	0.03
Error	5.00	0.01	0.05	-0.05	0.08
Absolute error	5.00	0.05	0.02	0.02	0.08
Relationship between estimates and true E. coli; Slope = 0.95, Intercept = 0.05, R ² = 0.92.					
Cokriging (E. coli + pH)					
Estimated E. coli	5.00	0.00	0.78	-0.75	0.92
Kriging variance	5.00	0.02	0.01	0.01	0.03
Error	5.00	0.00	0.02	-0.02	0.03
Absolute error	5.00	0.02	0.01	0.01	0.03
Correlation (E. coli vs pH)		-0.69			
Relationship between estimates and true E. coli; Slope = 0.98, Intercept = 0.02, R ² = 0.96.					

Table 8 shows moderate observed E. coli concentrations in BH (mean = 0.00, and SD = 0.75), indicating heterogeneous sound water contamination within the study locations. Ordinary Kriging (OK) overestimates the observed mean (0.01) and displays a low standard deviation (SD = 0.77), consistent with the smoothing effect inherent in kriging interpolation. The low mean error (0.01), moderate absolute error (0.05) with a strong relationship between estimated and observed values (slope = 0.98, and intercept = 0.02), indicates good performance, that is, the model explained about 96% of the observed spatial variability R-squared of E.coli suggesting a R-squared of $R^2 = 0.96$ suggest underestimation of higher values. CK improves performance, yield estimates closer to observed values (mean = 0.0 and SD 0.78), negligible mean errors (0.00) and absolute error (0.02) with a strong association between estimated and measured concentration (slope = 0.98 and intercept = 0.02),

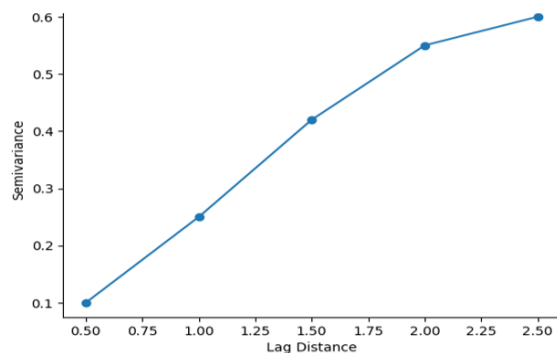


Figure 11: Experimental Variogram of Ordinary Kriging (OK) model in BH

Figure 11 shows that the experimental variogram of E. coli exhibits increasing semivariance from 0.10 to 0.60 as lag distance increases, indicating strong spatial dependence at short ranges. Low semivariance at small lags reflects similarity among nearby observations, while the gradual approach to a sill suggests a finite range of spatial correlation. Ordinary Kriging predictions have low kriging variance (mean = 0.02) and absolute error (0.05), with regression slope = 0.95, intercept = 0.05, and $R^2 = 0.92$. These results confirmed that OK captured the spatial structure of E. coli in the study area.

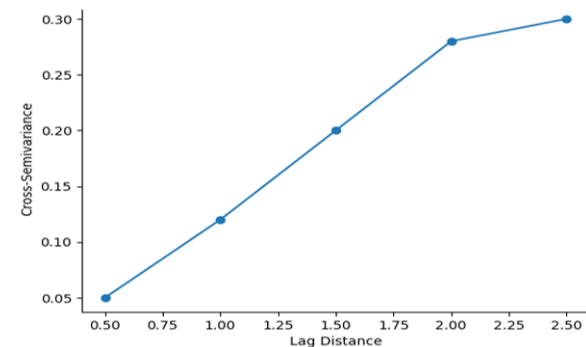


Figure 12: Experimental Cross-variogram of E. coli & pH for Cokriging model for BH

Figure 12 shows that the cross-variogram between E. coli and pH increases from 0.05 to 0.30, reflecting partial spatial correlation despite a moderate negative correlation (-0.69). Implies that CK reduces prediction uncertainty (mean = 0.02, absolute error = 0.02), with (slope equals 0.98, intercept = 0.02, and R-squared = 0.96). This demonstrates that incorporating pH improves the estimation accuracy of CK over OK.

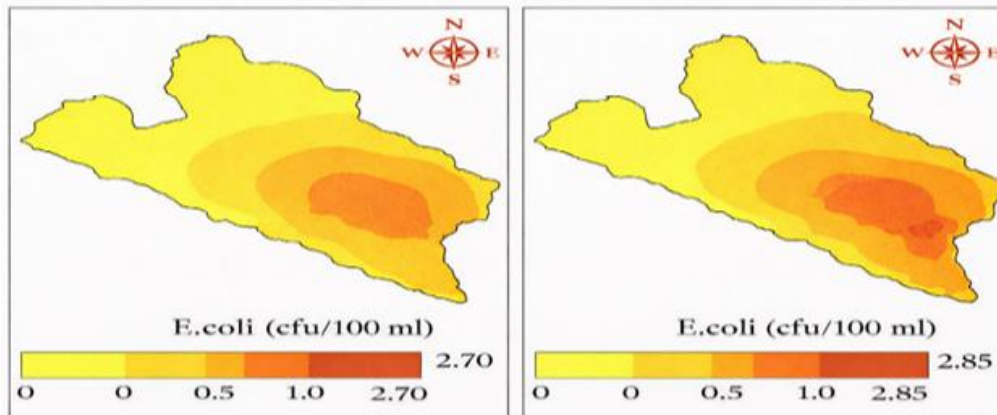


Figure 13: OK and CK Estimation Maps of E. coli in Borehole

Figure 13 displays the OK and CK estimation maps using colour gradient from light yellow, through light-orange, to deep red. The color scheme represents E. coli concentrations: light yellow indicates low contamination, light orange indicates moderate

contamination, and dark red indicates high contamination. Thereby supporting more effective public health planning and water resource management decisions for appropriate intervention.

Ordinary Kriging and Cokriging estimation results for HDW

Table 9: Ordinary Kriging and Cokriging estimation results for HDW

Method	N	Mean	Standard Deviation	Minimum	Maximum
True E. coli	15.00	-0.31	1.79	-3.00	3.10
Ordinary Kriging (E. coli)					
Estimated E. coli	15.00	-0.28	1.40	-2.20	2.40
Kriging variance	15.00	1.96	0.42	1.20	2.70
Error	15.00	0.03	0.74	-1.10	1.30
Absolute error	15.00	0.62	0.38	0.10	1.30
Relationship with true E. coli; Slope = 0.59, Intercept = 0.12, $R^2 = 0.44$					
Cokriging (E. coli + pH)					
Estimated E. coli	15.00	-0.30	1.15	-2.00	2.10
Kriging variance	15.00	1.32	0.31	0.90	1.90
Error	15.00	0.01	0.53	-0.80	0.90
Absolute error	15.00	0.44	0.29	0.05	0.90
Correlation (E. coli vs pH)	15.00	-0.91			
Relationship with true E. coli; Slope = 0.78, Intercept = 0.07, $R^2 = 0.66$					

Table 9 shows moderate observed E. coli concentrations in HDW (mean = $-0.3SD$, equals, 1.79), indicating homogeneous groundwater contamination within the study area. OK slightly underestimates the mean (-0.28) and displays a low standard deviation ($SD = 1.40$), consistent with the smoothing effect inherent in kriging interpolation. The low mean error (0.03), moderate absolute error (0.62), coupled with a weak correlation between estimated and observed values (slope = 0.59 and intercept = 0.12), suggest minimal bias (moderate predictive accuracy), with model explaining 44% of the observed spatial variability in E.coli concentrations ($R^2 = 0.44$) and a tendency to underestimate higher contamination levels. CK incorporate an auxiliary variable, significantly improved predictive

performance by producing estimates closer to the observed values (mean = -0.30 and $SD = 1.15$), lower errors (0.01) and lower absolute error (0.44), and The relationship between estimated and observed concentrations was more stronger when using the auxiliary variable (slope = 0.78, intercept = 0.07 and $r = -0.91$), indicates that CK model explains approximately 57% of the spatial variability of E.coli concentration in HDW ($R^2 = 0.66$), these results confirms that CK outperforms OK and is therefore more reliable approach for groundwater quality analysis across sample locations.

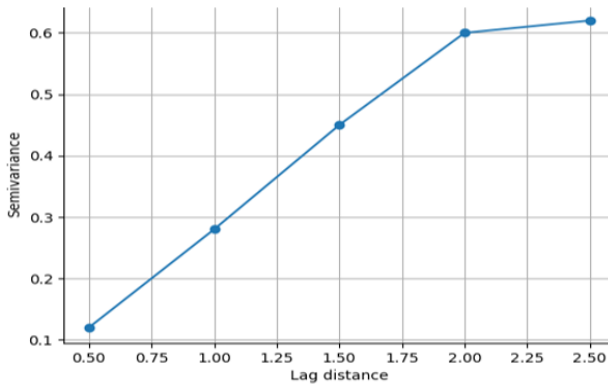


Figure 14: Experimental Variogram of the OK model for HDW

Figure 14 shows that the E. coli variogram in HDWs increases from ~0.7 at short lags to ~2.2 at longer distances, indicating moderate spatial dependence among wells. The high sill corresponds to a large mean kriging variance (1.96), indicating high dispersion and spatial heterogeneity of E. coli concentrations across the study area. Consequently, Ordinary Kriging explains only 44% of the observed variability ($R^2 = 0.44$), confirming its limited precision for estimating E. coli concentration in hand-dug wells.

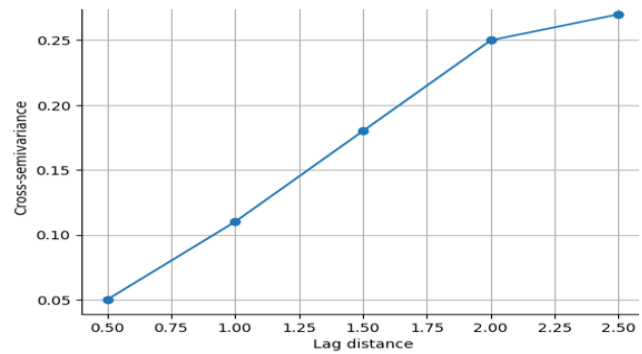


Figure 15: experimental Cross-variogram of E. coli & pH for CK (HDW)

Figure 15 shows that the cross-variogram (E. coli + pH) for HDWs increases from ~0.35 at short lags to ~1.35 at longer distances, indicating strong spatial association. The low cross-semivariance and mean kriging variance (1.32) reflected lower prediction uncertainty. At the same time, the strong negative correlation ($r = -0.86$) between CK performance and the auxiliary variable explained 66% of the variability in E. coli concentration ($R^2 = 0.66$). These results imply that incorporating pH as a secondary variable significantly enhances E. coli estimation in HGW.

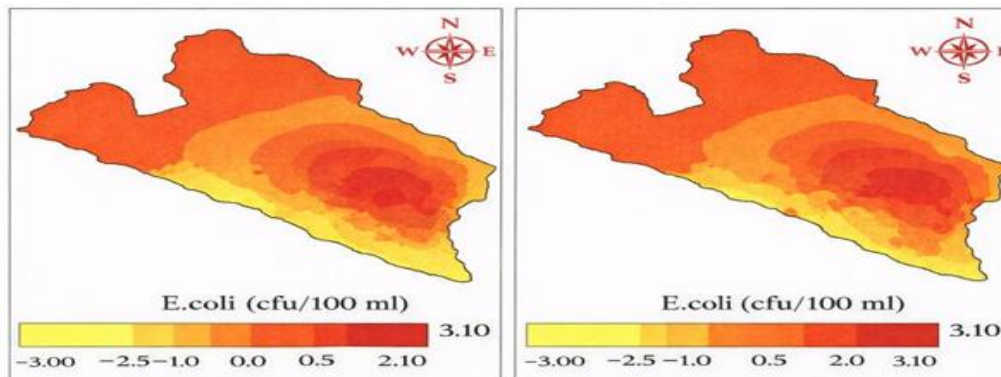


Figure 16: Ordinary Kriging and Cokriging Estimation Maps of E. coli in hand-dug wells

Figure 16 demonstrates that OK and CK maps use a color scale ranging from light yellow through orange to deep red. The color scheme represents E. coli concentrations: light yellow denotes low contamination levels, and darker orange to red indicate moderate to high microbial concentrations, respectively. At the same time, both methods reveal similar patterns. These color patterns indicate widespread microbial contamination in HGW, with CK providing a clearer depiction of contamination intensity, supported by a correlation coefficient ($r = -0.91$) and explanatory power of approximately 66% of the spatial variability in E. coli concentration ($R^2 = 0.66$).

DISCUSSION OF RESULT

The results presented in Tables 3–9 are briefly discussed in relation to the existing literature, with an emphasis on groundwater quality characteristics, compliance with WHO standards, geospatial dependence of parameters, and the comparative performance of geostatistical methods (ordinary kriging and cokriging).

The descriptive statistics in Table 3 indicate that the boreholes had low microbial contamination, with a mean E. coli concentration of 0.40 cfu/100 ml, while physicochemical parameters remained within WHO limits. This indicated that the groundwater conditions in the boreholes were protected and stable. This finding is consistent with the study of Bisiriyu *et al.* (2020) in Niger State. While the descriptive statistics for HDW in Table 4 showed microbial concentrations exceeding the WHO guideline (0.00 cfu/100 ml) for drinking-water standards, this indicated greater vulnerability of HDW to surface contamination, rendering the water unsafe for direct human consumption without appropriate treatment. This finding aligned with the study of Egbeyale *et al.* (2025) in Kwara State.

The study, as shown in Table 5, found that boreholes provide safe groundwater. In contrast, hand-dug wells pose significant public health risks, indicating the need for improved well protection, sanitation, and appropriate water treatment interventions. This finding is consistent with Egbeyale *et al.* (2025): boreholes

demonstrated lower levels of microbial contamination and greater stability, confirming their suitability as a safe drinking water source compared with hand-dug wells, which remain more exposed to surface-derived pollution and substandard physical and chemical conditions, including acidic pH. This finding is consistent with the research of Ojo *et al.* (2021) in Osogbo.

The results in Table 6 reveal that the spherical model best fits the microbial parameters and cross-variograms for boreholes. In contrast, both the spherical and exponential models fit the data well. Cross-variogram confirmed pH (BH 0.08, 1.25, 2.00 km; HDW 0.10, 1.35, 2.10 km) as the preferred auxiliary variable. This finding aligned with the studies by Isah and Abdullahi (2013) in Niger State and by Hooshmand *et al.* (2011a) in Iran.

The Correlation analysis results in Table 7 indicate strong positive and significant relationships among microbial parameters ($r = 0.93-0.99$, $p < 0.05$), suggesting common faecal contamination sources. The table also revealed a strong, significant negative relationship between pH and microbial contamination, particularly in HDW ($r = -0.91$, $p < 0.001$). This indicates that higher acidity is associated with higher pollution levels and that incorporating pH information strengthens geostatistical modeling. This finding is also consistent with the research of Suleiman *et al.* (2020; Sarani *et al.*, 2025)

Ordinary Kriging and Cokriging estimation results for BH in Table 8 found that the cokriging model explained about 96% of the geospatial variability of boreholes in *E. coli* concentration ($R^2 = 0.96$), supported by lower absolute error (0.02) and moderate negative correlation between *E. coli* and pH ($r = -0.69$). This demonstrated the superior predictive performance of cokriging over ordinary kriging ($R^2 = 0.92$). Therefore, incorporating pH improved the accuracy of estimation in groundwater quality analysis. This finding is in agreement with the study of (Sarani *et al.*, 2018; Hooshmand *et al.*, 2011a; Yalcin, 2005).

Ordinary Kriging and Cokriging estimation results for HDW in Table 9 found that ordinary kriging produced low predictive performance in HDW with ($R^2 = 0.44$), while cokriging improved prediction accuracy ($R^2 = 0.66$) with minimum absolute error (0.44) and strong negative correlation between *E. coli* and pH ($r = -0.86$). This indicated that incorporating pH as an auxiliary variable significantly improved contamination prediction and reduced estimation uncertainty in HDW. Validation results further showed that cokriging consistently outperformed ordinary kriging by reducing prediction errors and improving explanatory power. (BH; R^2 improved from 0.92 to 0.96, while HDW: R^2 increased from 0.90 to 0.95). This demonstrated that incorporating pH as a secondary variable significantly enhanced the accuracy of groundwater quality prediction. This finding is consistent with studies by Hooshmand *et al.* (2011a), Kumari *et al.* (2018), Singh and Verma (2019), Khan *et al.* (2023), and Sarani *et al.* (2025).

Conclusion and Recommendation

Conclusion

The study reveals clear differences in groundwater quality between BH and HDW. BH exhibits low microbial contamination, and physical and chemical parameters are within the WHO drinking-water standards. At the same time, HDW recorded high microbial concentrations above the WHO limit of (0.00 cfu/100 ml), indicating greater exposure to surface contamination and associated public health risks. Correlation analysis further shows strong positive

relationships among microbial indicators and strong negative relationships between pH and microbial indicators, confirming pH as a suitable auxiliary variable for spatial prediction.

Geostatistical analysis demonstrates clear geospatial dependence among groundwater quality parameters, with spherical models providing the best fit for microbial parameters and cross-variogram, while a Gaussian model best fitted pH. The study further established that cokriging consistently outperformed ordinary kriging, yielding lower prediction errors and greater explanatory power. Prediction accuracy improved from $R^2 = 0.92$ to 0.96 in boreholes and from $R^2 = 0.44$ to 0.66 in HDW following the incorporation of pH as a secondary variable. In conclusion, these findings confirmed that boreholes are safer and more reliable for domestic water supply, demonstrated the suitability of cokriging for groundwater quality prediction, and highlighted the need for routine well monitoring, sanitary well protection, and appropriate groundwater management interventions to minimize public health risk.

Recommendation

Based on the research findings, it is recommended that: Communities should adopt household water treatment methods (chlorination, boiling, and filtration) to reduce microbial contamination and improve water quality by constructing deeper wells with fitted covers.

The government should organize and train Community-based monitoring groups to conduct routine microbial tests, properly manage waste (Record proximity to latrines, waste dumps, livestock pens, and drainage patterns near wells and boreholes), and prioritize the use of boreholes where available.

The government should prioritize constructing protected boreholes and deeper wells to provide safer, more reliable water for human consumption.

Future studies should apply Cokriging with multiple explanatory variables (Record proximity to waste dumps and land-use) to improve predictive performance.

REFERENCES

- Al-Fatlawy, F., & Al-Ansari, N. (2022). Groundwater quality assessment using WQI. *Groundwater for Sustainable Development*, 16, 100700.
- American Public Health Association, American Water Works Association, & Water Environment Federation. (2017). *Standard methods for the examination of water and wastewater* (23rd ed.)
- Bisiriyu, M. T., Ganiyu, M. A., Audu, L. I., Abdullahi, A. S., & Animashaun, I. M. (2020). An assessment of groundwater quality in Tudun Fulani, Niger State, Nigeria. *Sule Lamido University Journal of Science and Technology (SLUJST)*, 1(2), 7-16.
- Chiles, J.-P., & Delfiner, P. (2012). *Geostatistics: Modelling spatial uncertainty* (2nd Ed.) Wiley (online)
- Cressie, N. (1993). *Statistics for spatial data*. Wiley (online).
- Egbeyale, G. B., Yussuff, Q. B., Ogunseye, T. T., & Bello, A. K. (2025). Comparison of water quality from boreholes and hand-dug wells around and within Malete, Kwara State, Nigeria. *IOSR Journal of Environmental Science, Toxicology*

- and Food Technology (IOSR-JESTFT), 19(5) Ser. 1, 01–07
- Environmental Protection Agency (2009). National Water Quality Inventory Report to Congress (2004 Reporting Cycle). U.S. Environmental Protection Agency, Office of Water, EPA 841-R-08-001 (online).
- Goovaerts, P. (1997). Geostatistics for natural resources evaluation. Oxford University Press.
- Goovaerts, P. (2000). Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of Hydrology*, 228, 113–129.
- Goovaerts, P. (2023). Advances in cokriging for environmental applications. *Spatial Statistics*, 55, 100767.
- Hooshmand, A., Delbari, M., & Nourani, V. (2011a). Comparative study of kriging and cokriging for groundwater quality mapping. *Arabian Journal of Geosciences*, 4, 173–181.
- Hooshmand, A., Mohammadi, K., & Nakhaei, M. (2011b). Application of cokriging in groundwater quality assessment. *Environmental Monitoring and Assessment*, 175, 651–663.
- Isaaks, E. H., & Srivastava, R. M. (1989). An introduction to applied geostatistics. Oxford University Press.
- Isah, A., & Abdullahi, U. (2015). An application of geostatistics to the analysis of water quality parameters in rivers and streams in Niger State, Nigeria. *American Journal of Theoretical and Applied Statistics*, 4(5), 373–388.
- 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*.
- Ishaku, J. M., Ahmed, A. S., & Abubakar, M. (2011). Assessment of groundwater quality in parts of Yola, north-eastern Nigeria. *Journal of Water Resource and Protection*, 3(12), 925–933.
- Kanu, I., Opara, C., & Okorie, N. (2023). Evaluation of spatial interpolation methods for groundwater quality assessment. *Environmental Science and Pollution Research*, 30, 54451–54463.
- Matheron, G. (1963). Principles of geostatistics. *Economic Geology*, 58, 1246–1266.
- Mehrijardi, R. T., Jahromi, M. Z., Mahmodi, S., & Heidari, A. (2008). Spatial analysis of groundwater quality using geostatistics. *Environmental Monitoring and Assessment*, 137(1–3), 237–245.
- Miller, H. J. (2007). Tobler's first law and spatial analysis. *Annals of the Association of American Geographers*, 97(2), 284–289.
- Queiroz, M. M. A., Alcântara, E., & Andrade, E. (2008). Mapping groundwater salinity using cokriging. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 12, 501–508.
- Sarani, A., Khatibi, R., & Amini, M. (2025). Improving groundwater quality prediction in arid regions using Cokriging with hydrochemical auxiliary variables. *Environmental Earth Sciences*, 84(2), 1–16.
- Suleiman, A. A., Ibrahim, A., & Abdullahi, U. A. (2020). Statistical explanatory assessment of groundwater quality in Gwale LGA, Kano State, Northwest Nigeria. *Hydrospatial Analysis*, 4(1), 1–13.
- Singh, P., & Verma, P. (2019). A comparative study of spatial interpolation techniques (IDW and kriging) for determining groundwater quality. In *GIS and geostatistical techniques for groundwater science* (pp. 43–56). Elsevier.
- Tijani, M. N., Onugba, A., & Akinyemi, O. D. (2005). Groundwater contamination assessment around wells in Nigeria. *Water Resources Journal*, 29, 35–48.
- Webster, R., & Oliver, M. A. (2007). *Geostatistics for environmental scientists* (2nd Ed.). Wiley.
- World Health Organization (2017) *Guidelines for drinking-water quality* (4th Ed.) WHO Press (Online)
- Yalçın, E. (2005). Cokriging and its effect on the estimation precision. *Journal of the South African Institute of Mining and Metallurgy*, 105, 223–232.