Geospatial Analysis of Spatial Variability of Groundwater Quality Using Ordinary Kriging: A Case Study of Dungarpur Tehsil, Rajasthan, India

Seema Jalan[1*], Devendra Singh Chouhan[2,*] and Shailesh Chaure[3]
[1]Professor of Geography, Mohanlal Sukhadia University, Udaipur, Rajasthan
[2] Assistant Professor of Geography, Mohanlal Sukhadia University, Udaipur, Rajasthan
[3] Associate Professor of Geology, Government Holkar Science College, Indore, Madhya Pradesh
*Email: seemajalan1@gmail.com

(Received: Dec 21, 2021; in final form: Sep 3, 2022)

Abstract: Groundwater is one the major sources of natural water being exploited excessively for various uses in India. Thus, it is very essential to monitor the spatial and temporal variability of groundwater quality. Geo-Statistical Interpolation using GIS has been considered as the best and most advanced method for the interpolation and prediction studies of groundwater pollution and quality, and is adopted universally. In this paper, ordinary Kriging with logarithmic data transformation has been used to interpolate and predict the spatial variation of groundwater quality parameters - EC, TDS, pH, Na^+, Ca^{2+}, Bi-Carbonate, Fluoride, Chloride, Sulphate and Nitrate using data pertaining to 48 well locations in the Dungarpur tehsil. Data was transformed and normalized using Logarithmic Transformation Method and Semivariograms were drawn and analyzed for selecting the suitable model. The best Semivariogram model was obtained based upon cross validation and on the lesser RMSE criterion and Coefficient of Determination. The results show that the best semivariogram model based on RMSE varied for each water quality parameter. For log transformed data Exponential model was found suitable for EC, TDS, Na^+, TH etc.; Spherical model for Ca^{2+} ; Chloride Gaussian Model for Chloride. For original or raw for non-transformed data Exponential Model was found suitable for Fluoride, Sulphate and Nitrate; and Gaussian Model for pH and Bi-Carbonates.

Keywords: Interpolation, Spatial Distribution, Ground Water Quality, Kriging, Transformation

I. Introduction

Water is crucial for the existence and propagation of life on earth. However, its reserves are depleting day by day due to either over exploitation and contamination through natural and anthropogenic agents or industrial activities. Groundwater quality is severely affected in the regions with low rainfall, high temperatures, and lithology with high concentration of various contaminants and chemical compounds such as Sodium, Fluoride, Bi-Carbonates, various other metals and their compounds. The extraction of ground water is increasing day by day due to limited amount of good quality potable water resources from which water is available to human beings. Further, the water quality of various surfaces water bodies has deteriorated day by day due to industrialization (Nas, 2009). It is essential now that serious efforts must be made for sustainable development, management, monitoring and preparation of government policies for the development of ground water and its quality. In understanding the spatial and temporal variability in distribution of ground water quality parameters based upon the data known or collected, Geostatistics and Interpolation are universally accepted as most suitable methods.

In the recent past, various researches used advanced methods of Statistics and Geo-Statistical Interpolation, IDW, Spline, Kriging, Global Polynomial Interpolation, Radial Basis Function etc. which proved very promising in the accuracy and prediction of unknown parameters from known parameters at variable levels (Shankar et al, 2018; Ali , 2019; Zeybek et al, 2020; Katipoğlu, 2021)
The objectives of the study are to select suitable and appropriate spatial interpolation models and methods for studying the spatial variability of ground water quality parameters. Most commonly used methods of data transformation have been executed and analysed to predict the general trend of the spatial distribution of selected groundwater quality parameters in Dungarpur tehsil of Dungarpur district in southern Rajasthan, India. The performance and suitability of different Geo-Statistical data processing and semivariogram models has been assessed and ordinary Kriging has been employed for various ground water quality parameters to study their spatial variations and patterns.

2. Study area, data used, materials and methods

2.1. Study Area

Dungarpur tehsil is situated in the north-western part of Dungarpur district, one of the southernmost districts of the state of Rajasthan in India (Figure 1) between 23° 20' 1.56" N latitude and 73° 21' 6.70" E longitude to 24° 0' 41.4144" N latitude and 74° 22' 50.484" E longitude. It shares political boundaries with the state of Gujarat in the south as well as in west direction. In 2022, the estimated total population of the district is 15,59,120 persons (as per aadhar uidai.gov.in December 2020 data).
Dungarpur tehsil is a tribal dominated area being economically, socially and educationally backward. The estimated population of the tehsil in 2022 is 6,34,141 persons. As per Census 2011, the total population of the tehsil was 4,95,423 persons residing in 98,876 households. It is a predominantly rural area with 90.4 percent of population living in 305 villages. In 2011, ST population comprised 77.71 percent of the total population. Main workers comprised only one-third of the population of the tehsil having gainful employment for more than 06 months in a year. The population is mostly dependent on groundwater for drinking requirements.

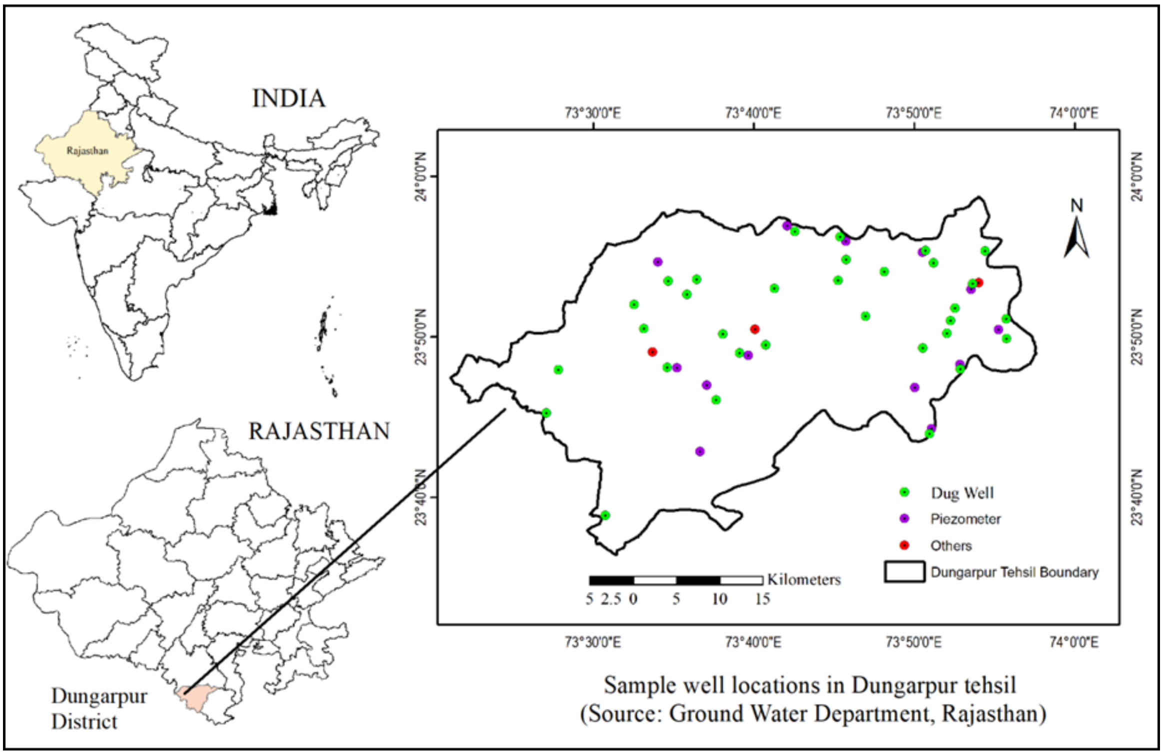

Figure 1. Location of study area and sample well locations in Dungarpur tehsil.

Occurrence of high fluoride in ground water is a matter of great concern. Particularly the areas adjoining blocks of Aspur, Dungarpur and Sagwara tehsils are severely affected by Flouride contamination in groundwater. The high amount of fluoride in the groundwater is reflected in the bone related diseases widely prevalent in the resident population.

The overall geology of the area is highly complex. Dungarpur district is underlain mainly by hard rock formations such as Gneisss, Schists, Phyllites and Slates. The Fluorides in groundwater apparently originate from the rock formations.

2.2. Data used

A total number of 48 sites were selected using the water quality observation well locations used by the Ground Water Department (GWD), Rajasthan and mentioned in the 'Water Quality Report of Pre-Monsoon 2014', issued by the GWD (Figure 1). The well location sites were identified with help of their names and searching their geographic latitude and longitude values by querying and identify tool on the GWD Assets layer in the 'Rajdharaa-Ground Water GIS' Portal of the Government of Rajasthan. The pre-monsoon water quality data of 11 parameters - Electric Conductivity (EC), Total Dissolved Solids (TDS), pH, Sodium (Na^+), Calcium (Ca^{2+}), Bi-Carbonate (HCO_3^-), Fluoride (F^-), Total Hardness (TH), Chloride (Cl^-), Sulphate (SO_4^2) and Nitrate (NO_3^-) - were taken from the above mentioned Water Quality Report.

Since the primary objective of the study is to assess the performance of spatial interpolation models, only representative pre-monsoon data of the region has been used. The spatial variability of some of the physicochemical parameters may be altered due to the precipitation in the post monsoon season.

3. Method of spatial correlation analysis

3.1. Semi-variable function

In statistics, semi variable function (Equation 1) is used for the calculation and understanding the pattern of spatial correlation, if the sampling data values are normally distributed and the skewness of the sample data value distribution lie within a range of - 1 to +1. (Hu et. al, 2009)

Semivariogram analysis is the main tool to graphically represent the spatial correlation among the neighboring sample data values and observations (Bàrdossy, 1997). It shows the relationship between the lag distance on the horizontal axis and the semivariogram value on the vertical axis (Figure 2). Lag distance is the distance between the measurements of a particular property. The semivariogram value increases from low to high values indicating higher spatial autocorrelation at the small lag distance (Nayanaka et al., 2010).

$$\gamma(h) = \frac{1}{2N(h)}\sum_{i=0}^{N(h)}[\,Z\,(x_i)\ - Z(x_i\ + h\,)]^2$$

.......Equation (1)

Where,

$\gamma(h)$ = Semi-Variable Function,

$Z\,(x_i)$ = Sample or Data Value,

$Z(x_i\ + h\,)$ = Sample or data value with 'h' distance from x_i ,

$N(h)$ = No. of attribute pairs that are separated by distance 'h'.

Degree of spatial correlation can be evaluated by calculation of nugget effect which is the ratio of nugget to sill

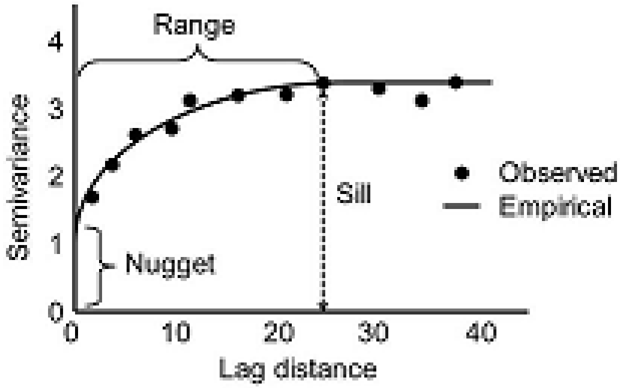

Figure 2. Semi-Variogram

3.2. Common Semi-Variogram Model Fitting Equations

The equations of some of the commonly used models for the fitting of semi-variograms are:

a) Spherical model

$\gamma(h) = C(\frac{3h}{2a} - \frac{1}{2}\frac{h^3}{a^3})$, $if\, h \le a$

$\gamma(h) = C$, $if\ h > a$

.......Equation (2)

Where,
a = Range or distance between correlated and uncorrelated data,
h = Lag distance and
C ≠ 0

If any two points are separated by a distance greater than range '**a**', then the corresponding points or data locations are independent of each other. The value of the Sill '**C**' is the value of the variogram for the distances greater than the range. It is equal to $C(0)$, the variance of the random variable (Bàrdossy, 1997)

b) Exponential model

$$\gamma(h) = C\ (1 - e^{-\frac{h}{a}})$$

.... Equation (3)

Where,
a = Range or distance between correlated and uncorrelated data,
h = Lag distance and
C ≠ 0

In case of exponential model all the random variables are supposed to be dependent. But there is an effective range '3a' such that random variables related to points more distant than '3a' can be treated as independent. The Sill 'C' is equal to 'C(0)', the variance of the random variable (Bàrdossy, 1997).

c) Gaussian model

$$\gamma(h) = C(1 - e^{-\frac{h^2}{a^2}}\ )$$

.......Equation (4)

Where,
a = Range or distance between correlated and uncorrelated data,
h = Lag distance and
C ≠ 0

Again, in Gaussian model the Sill 'C' is equal to 'C(0)', the variance of the random variable. Parameter 'a' is again related to effective range of the variogram. There is a limit to the effective range in Gaussian model which is equal to '√3a' (Bàrdossy, 1997).

4. Interpolation method

There are two most common categories in which statistical interpolation methods viz. Deterministic and Geostatistical. (Chen et al., 2013) Geo-statistics can be regarded as a collection of numerical techniques that deal with the characterization of spatial attributes employing primarily random models in a manner similar to the way in which time series analysis characterizes temporal data (Olea et al., 2012). It deals with spatially autocorrelated data that have a basic structure or spatial patterns which can be manifested in semivariogram analysis.

Kriging is a Geostatistical interpolation method used to study and predict the values for the locations where no data is available, on the basis of data collected and measured known locations. Kriging also facilitates in the calculation of the uncertainty of error and estimation of accuracy in the surface of values generated. Moreover, various researchers have found that Kriging provides better results as compared to other deterministic and stochastic methods. Therefore, Kriging has been selected for the present study.

4.1. Kriging

Kriging method works on the principle of existence of spatial autocorrelation for the application of geostatistics. To study the spatial pattern of the data values and parameters, semivariograms are used as a descriptive tool (Goovaerts, 1999). The spatial dependence between measured points is calculated on the basis of the distance between these points with the help of semivariance (Tatalovich, 2006). Thereafter, the fitting of semivariograms is done using different models such as Gaussian, Spherical and Exponential to check the best suitable model to be adopted for the study and to generate optimum interpolation weights (Burroughs and

McDonnell, 1998). Kriging is a very flexible interpolator that can be exact or smooth. It allows a variety of output surfaces including predictions, prediction standard errors, and probability (Johnston et al., 2004). Kriging technique eases out the optimal, unbiased estimation of the regionalized variables of unsampled sites with the properties of the semivariogram, using initially taken data values. (Mehrjardi et al., 2008).

4.2. Logarithmic transformation

To perform Kriging operation the data value distribution of random variable which is taken must be normally distributed and follow stationarity principle. But it may not be practically necessary. To account for the non-normality, the mathematical process of data transformations are used such as 'Logarithmic' transformation. Logarithmic transformation is basically a class of mathematical operation in which the log values of all the random variable values are taken to make the whole distribution follow normality principle or comes closer to it. In this study natural logarithmic transformation has been used for all different water quality parameters whenever the data distribution for any particular parameter was not found normally distributed by taking the natural logs of all the individual parameter values (after Osbourne, 2002).

4.3 Cross-Validation

Generally, the method used for the validation of interpolation processes is called Cross-Validation (Voltz and Webster, 1990). The key procedure in the process of cross-validation is the temporary removal of one data point and the removed point is estimated with the help of remaining data points, thereafter the difference between the actual and estimated values is calculated. This procedure is repeated for all the remaining set of data points.

In this study, the cross-validation of interpolation has been performed for all the water quality parameters on the basis of least Root Mean Square Error (RMSE).

5. Methodology

Suitable hydro geochemical parameters for modeling the ground water quality have been selected based on literature review (after McNeely et al., 1979). Geo-statistical analysis (Ordinary Kriging) has been used for spatial interpolation to model the spatial distribution of groundwater chemistry.

Geostatistical analysis has been implemented in the following three steps:

- Semivariogram creation and analysis.
- Validation of Semivariogram models.
- Generation of the groundwater quality prediction surfaces and maps.

Geostatistical Wizard of ArcGIS 10.5 was used for the exploratory data analysis, data transformation, and semivariogram analysis.

RMSE criterion has been used for comparison between the different semivariogram models and data transformation. The smallest value of RMSE indicates the most suitable model for the data. Figure 3 illustrates the methodological workflow of the study.

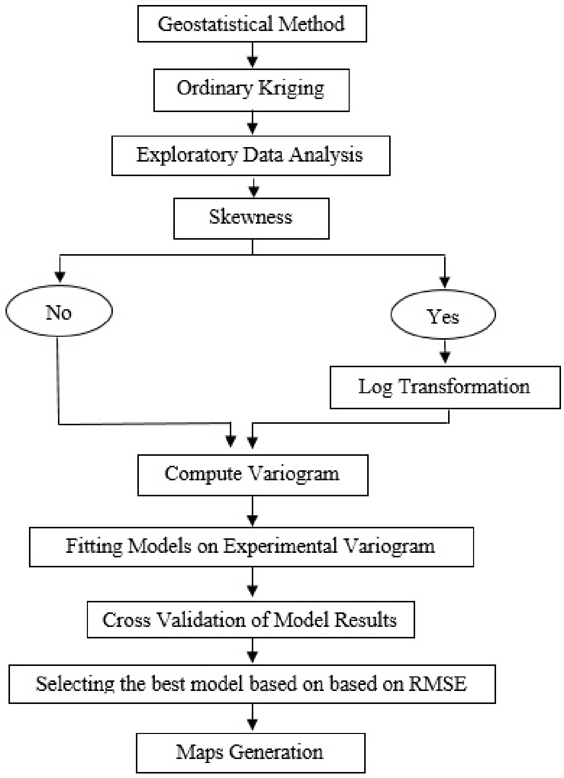

Figure 3. Methodological flowchart

6. Result and discussion

Distribution Characteristics

The following tables summarize the distribution characteristics of the selected hydro-geochemical parameters based on raw and log transformed data. Descriptive statistics describing distribution based on natural log transformed and raw untransformed values has been given in Tables 1.

Results reveal that the raw data distributions of EC, TDS, TH, Na^+, Cl^- and NO_3^- were highly positively skewed, whereas those of pH, Ca^{2+}, F^-, HCO_3^-, CO_3^{2-} and SO_4^{2-} were more or less normally distributed. The natural log transformed distributions of all the parameters obeyed normal distribution. The acceptable range of skewness considered was -2 to + 2 and the acceptable range of kurtosis considered was -3 to +3 (after Garson, 2012).

Table 1. Descriptive statistics of hydro-geochemical parameters for raw and log transformed distribution.

Indices	Minimum		Maximum		Average		Standard Deviation		Skewness		Kurtosis	
	lgN	Raw	lgN	Raw	lgN	Raw	lgN	Raw	lgN	Raw	lgN	Raw
EC	2.66	460	3.38	2390	2.9	835.73	0.14	335.59	0.98	2.42	1.35	8.75
TDS	2.40	252	3.12	1315	2.65	472.10	0.15	193.74	0.96	2.21	1.07	6.76
pH	0.89	7.8	0.99	9.7	0.93	8.62	0.02	0.40	0.12	0.35	1.46	1.72
TH	2.0	100	2.83	680	2.36	245.93	0.16	99.44	0.39	1.93	0.59	6.50
Na^+	1.46	29	2.40	251	1.87	81.19	0.19	37.86	0.06	1.92	0.03	7.31
Ca^{2+}	1.30	20	1.95	90	1.61	45.29	0.20	19.83	-0.14	0.44	-1.16	-0.91
F^-	-0.92	0.12	0.25	1.8	-0.10	0.93	0.26	0.49	-0.72	0.49	0.59	-0.8
Cl^-	1.63	43	2.72	525	2.04	126.67	0.21	79.85	0.78	2.93	0.85	12.41
HCO_3^-	1.38	24	2.58	378	2.20	178.29	0.22	80.90	-0.98	0.75	3.0	0.15
SO_4^{2-}	1.04	11	2.15	141	1.57	46.42	0.29	33.05	0.13	1.39	-0.54	1.25
NO_3^-	0	1	2.06	116	1.24	27.35	0.44	29.06	-0.47	2.08	1.29	3.84

The log normal values of EC range between 2.66 to 3.38 μScm^{-1}; for TDS range between 2.40 to 3.12 mg/l; for pH range from 0.89 to 0.99 ; for TH range from 2 to 2.83 mg/l ; for Na^+ range from 1.46 to 2.40 mg/l ; for Cl^- range from 1.63 to 2.72 mg/l ; and for NO_3^- range from 0 to 2.06 mg/l.

The log transformed values of Ca^{2+} range from 1.30 to 1.95 mg/l ; for F^- range from -0.92 to 0.25 mg/l ; for HCO_3^- range from 1.38 to 2.58 mg/l ; and for SO_4^{2-} range from 1.04 to 2.15 mg/l. The raw values of EC range from 460 to 2390 μScm^{-1} ; that of TDS range from 252 to 1315 mg/l ; and that of TH range from 100 to 680 mg/l. The raw values of Na^+ range from 29 to 251 mg/l ; that of Ca^{2+} range from 20 to 90 mg/l ; that of F^- range from 0.12 to 1.8 mg/l and of Cl^- range from 43 to 525 mg/l. The raw values of HCO_3^- range from 24to 378 mg/l, that of SO_4^{2-} range from 11 to 141 mg/l and of NO_3^- range from 1 to 116 mg/l.

Higher spatial variability, with pockets of extremely high concentration, as reflected in the range of raw values, high standard deviation, high positive skewness and kurtosis, is observed in EC, TDS, TH, Na^+, Cl^- and NO_3^-However positive skewness and leptokurtic distribution indicates higher concentration of values towards lower end of the scale. The distribution of Ca^{2+}, F^-, HCO_3^- and SO_4^{2-} is normally distributed. However the range and standard deviation values indicate high spatial variability in concentration of these parameters tooacross the tehsil.

Semivariogram creation and validation of models

Attempt has been made to identify the best optimum model by comparing the Spherical, Gaussian and Exponential model for both raw data distributions as well as natural log normalized data distributions for all the twelve indices. The minimum Root Mean Square Error (RMSE) values across different models (Gaussian, Spherical and Exponential variogram models using raw and log transformed data) have been compared for each parameter to identify the most suitable model. The RMSE values obtained for all water quality parameters for different Semivariogram models have been given in Table 2.

Table 2. Best fit Semivariogram model for the selected water quality parameters (bold and asterisk indicates lowest RMSE).

Parameters	Angle Tolerance (in°)	Ordinary Kriging					
		Model on Raw Data			Model on Log Transformed Data		
		Spherical	Exponential	Gaussian	Spherical	Exponential	Gaussian
EC	45	329.53	312.91	412.35	321.73	**306.19***	378.78
TDS	45	182.14	171.14	222.6	178.02	**167.60***	209.53
pH	45	0.3463	0.3426	**0.3351***	0.3486	0.3379	0.3381
TH	45	107.2	101.7	134.76	106.3	**100.76***	102.46
Na^+	45	32.89	31.42	36.49	32.66	**31.41***	32.16
Ca^{2+}	45	18.96	19.26	19.21	**18.95***	19.28	19.25
F^-	45	0.303	**0.288***	0.349	0.319	0.309	0.311
Cl^-	45	89.93	83.66	95.34	87.84	83.46	**80.52***
HCO_3^-	45	73.75	71.54	**71.29***	74.43	72.44	80.71
SO_4^{2-}	45	26.97	**26.57***	27.77	28.36	27.23	29.95
NO_3^-	45	19.7	**18.07***	24.9	22.44	20.05	22.33

The best semivariogram model varies for each water quality parameter. Exponential model of logarithmically transformed data was most suitable method for EC, TDS, Na^{+} and TH. For Ca^{2+} Spherical model for logarithmically transformed data was found suitable. For Cl^{-} Gaussian model of logarithmically transformed data was found suitable. For F^{-}, $SO4^{-}$, and $NO3^{-}$ the Exponential model with original (raw) data has been found most suitable. The Gaussian model with original data has been found most suitable for pH and $HCO3^{-}$. The best fit Semivariance models have been illustrated in Figure 4 (a - k).

The spatial structure indices of the best fit Semivariogram model selected for each parameter have been summarized in Tables 3. The EC shows high spatial correlation, and there is no nugget effect, which has a range of 0.947 degrees. TDS also shows high spatial correlation with zero nugget effect and effective range of 0.965 degrees. The pH shows high spatial correlation, 0.09% nugget effect and in effective range of 3.662 degrees. TH also shows high spatial correlation with zero nugget effect, where range is 9.051 degrees.

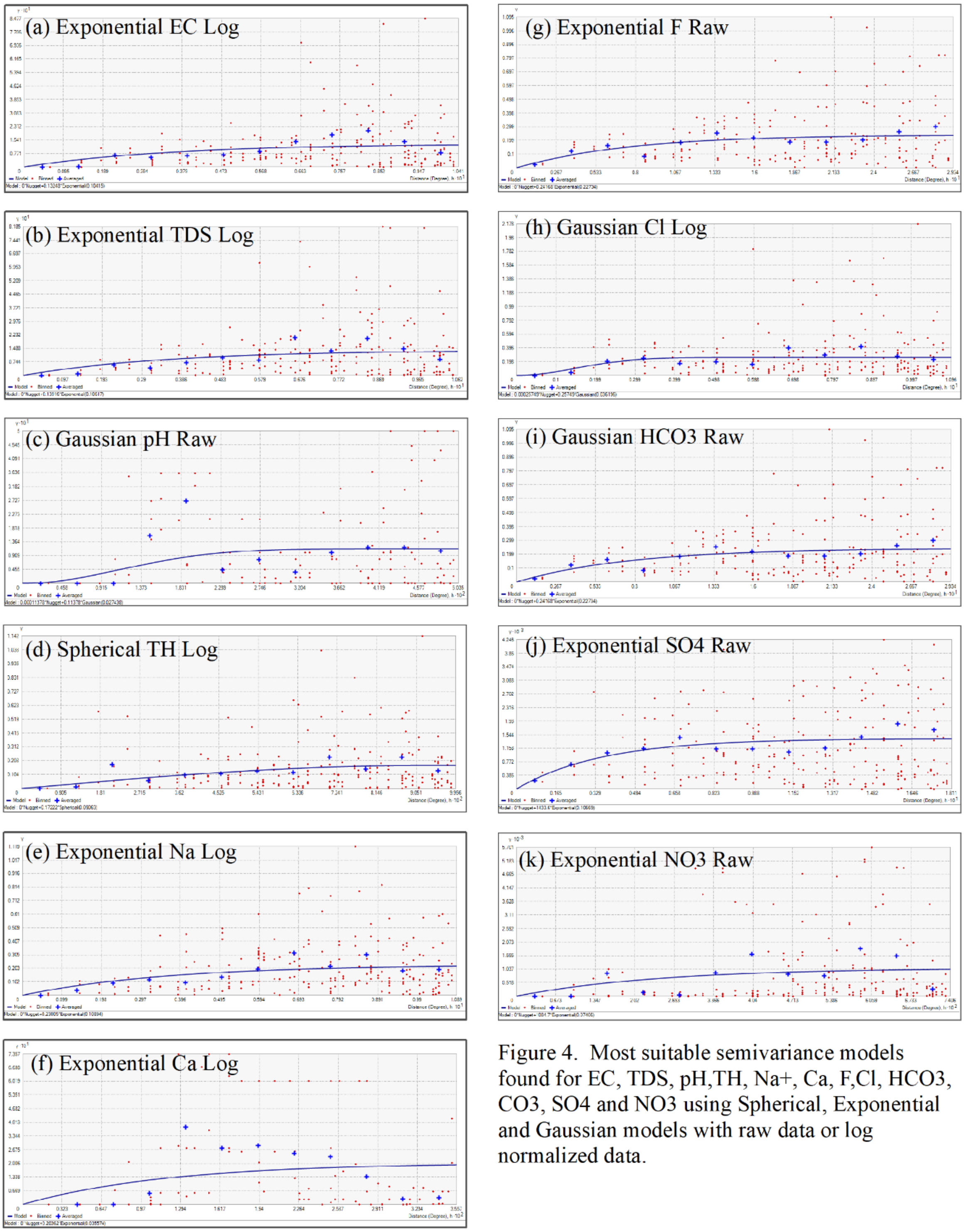

Figure 4. Most suitable semivariance models found for EC, TDS, pH,TH, Na+, Ca, F,Cl, HCO3, CO3, SO4 and NO3 using Spherical, Exponential and Gaussian models with raw data or log normalized data.

Spatial structure analysis

Na^+ shows high spatial correlation, and there is no nugget effect, which has a range of 1.089 degrees. Values of Ca^{2+} also show high spatial correlation with zero nugget effect and effective range of 3.557 degrees. F^- shows high spatial correlation, zero nugget effect in effective range of 2.667 degrees. The values of Cl^- also show high spatial correlation with 0.1% nugget effect, where range is 0.498 degrees.

HCO_3^- data shows high spatial correlation, and there is no nugget effect, which has a range of 5.225 degrees. SO_4^{2-} also shows high spatial correlation, zero nugget effect in effective range of 1.482 degrees. NO_3^- also shows high spatial correlation with zero nugget effect, where range is 6.733 degrees.

The basis for the describing the spatial correlations in the values of different indices as high spatial correlation is that if the nugget effect is less than 25 percent it suggests that a large portion of the variance in data is introduced spatially and thus emphasizes a strong spatial dependence of the data or variable. A high nugget effect (greater than 75 percent) indicates weak spatial dependency (Mutuma, 2017). Thus all the geochemical parameters are highly spatially correlated.

Generation of ground water quality prediction surfaces and analysis

The interpolation surfaces generated by Kriging with optimal parameters models are shown in Figure 5 (a) & (b). The interpolation surfaces are smooth and show regional variation patterns corresponding to the hydrogeological conditions in the region.

Table 4 shows the correlation (r) among various water quality parameters. Table 5 shows the Coefficient of Determination (R^2) between the raw and predicted or interpolated values of the different water quality parameters. The measured values explain 20% to ~80 percent variability in predicted values.

Table 3. Spatial structure indices of the selected parameters for the best fit Semivariogram

Parameters	Model	Transformation	Indices				
			Nugget	Sill	Nugget/Sill	Nugget Effect (%)	Range (Decimal Degrees)
EC	Exponential	lgN	0	0.132	0	0	0.947
TDS	Exponential	lgN	0	0.139	0	0	0.965
pH	Gaussian	Raw	0.0001	0.113	0.0009	0.09	3.662
TH	Exponential	lgN	0	0.169	0	0	9.051
Na^+	Exponential	lgN	0	0.23	0	0	1.089
Ca^{2+}	Spherical	lgN	0	0.184	0	0	3.557
F^-	Exponential	Raw	0	0.241	0	0	2.667
Cl^-	Gaussian	lgN	0.0003	0.257	0.001	0.1	0.498
HCO_3^-	Gaussian	Raw	7.8063	7814.12	0.001	0.1	5.225
SO_4^{2-}	Exponential	Raw	0	1433.6	0	0	1.482
NO_3^-	Exponential	Raw	0	1084.7	0	0	6.733

Table 4. Correlation matrix of the selected water quality parameters.

Correlation	EC	TDS	pH	Cl^-	SO_4^{2-}	HCO_3^-	NO_3^-	F^-	TH	Na^{+1}	Ca^{+2}
EC	1										
TDS	**0.99**	1									
pH	-0.48	-0.43	1								
Cl^-	**0.9**	**0.87**	-0.421	1							
SO_4^{2-}	**0.7**	**0.77**	-0.188	0.535	1						
HCO_3^-	**0.63**	**0.58**	-0.53	0.384	0.186	1					
NO_3^-	0.419	0.515	-0.095	0.17	**0.61**	0.145	1				
F^-	0.234	0.254	-0.182	0.017	0.391	0.132	**0.53**	1			
TH	**0.91**	**0.88**	**-0.6**	**0.82**	**0.63**	**0.603**	0.36	0.26	1		
Na^+	**0.86**	0.171	-0.238	**0.79**	**0.62**	**0.548**	0.248	0.04	**0.6**	1	
Ca^{2+}	**0.65**	**0.66**	-0.255	**0.57**	**0.51**	0.355	0.398	0.17	**0.71**	0.37	1

Table 5. Coefficient of Determination (R^2) between measured and predicted value of various water quality parameters

Parameters	EC	TDS	pH	TH	Na^+	Ca^{2+}	F^-	Cl^-	HCO_3^-	SO_4^{2-}	NO_3^-
R^2	0.4	0.5	0.6	0.2	0.6	0.38	1	0.2	0.47	0.6	0.79

The estimation of EC shows that medium to high levels of EC are found in the southern, north-eastern, northern parts of the Dungarpur tehsil which can be contributed to high levels of salinity and high mineral percentage due to the ionization and solubilization taking place in ground water aquifers in these areas. Strong positive association of EC occurs with TDS (r = 0.988), TH (r = 0.91), Cl^- (r = 0.898), and Na^+ (r = 0.856). EC also moderately positively correlates with levels of SO_4^{2-} (r = 0.695) Ca^{2+} (r = 0.646) and HCO_3^- (r = 0.634).

The distribution surface of TDS reveals that the high concentrations are found in the southern, north-eastern and northern parts of the tehsil. High levels of TDS strongly correlate with occurrence of high levels of Cl^- (r = 0.868), SO_4^{2-} (r = 0.771) and TH (r = 0.881). Moderate positive association occurs with Ca^{2+} (r = 0.662) and $HCO3^-$ (r = 0.583). The highest occurrences predominantly correspond with barren, rocky wasteland region. Thus lithology may be a significant factor for the observed spatial distribution.

Concentration of pH is higher in a major area of eastern and north-eastern zone of the tehsil. However the entire tehsil has highly alkaline ground water with the minimum being 7.8. The distribution of pH shows moderate to low correlation with all other parameters. Strongest negative association occurs with TH (r = -0.599), HCO_3^- (r = -0.53) and Cl^- (r = -0.421).

Distribution of Na^+ and Cl^- have strong positive association (r = 0.786). Distribution of Na^+ is moderately associated with SO_4^{2-} (r = 0.615), HCO_3^- (r = 0.548) and TH (r = 0.595). The concentrations of these parameters are distinctly highest in the north central, north western and north eastern part of the tehsil. Extreme southern part of the tehsil has moderate levels of Na^+, Cl^- and TH while high concentration of HCO_3^- and SO_4^{2-}. Entire tehsil, except a small pocket in the south-east has moderate to high levels of Ca^{2+} being highest in north-east and western half of the region, and moderate in the central and northern part. Distribution of Ca^{2+} has strongest positive correlation with TH (r = -0.71), EC (r = 0.646) and TDS (r = 0.662), and moderately associated with Cl^- (r = 0.573) and SO_4^{2-} (r = 0.51)

The spatial distribution of F^- shows a distinctly concentrated pattern. Extreme eastern and south-western parts of the tehsil have notably high levels, while western part is a low concentration region. The central region has moderate F levels. The distinctive concentration is substantiated by the correlation values which are on the lower end of the scale for all parameters except NO_3 (r = 0.533), which also shows highest concentration in the eastern half of the tehsil. . The southern pocket of high concentration corresponds with the predominantly forest and agricultural area. The eastern part of the tehsil is predominantly agricultural and settlement area. The NO_3 distribution in the tehsil may be accounted for by geological conditions and anthropogenic causes like use of agricultural fertilizers and coal combustion as well.
Broadly it is observed that most of the red patches of the geo-chemical parameters (high predicted values) are falling in the areas occupied by gneisses particularly the extreme east and northern part. Minimum values of EC, TDS and TH (blue) are in an elongated zone falling in areas occupied by phyllites. However, the areas in the extreme west having moderate values (yellow) are also occupied by phyllites. The high values of F^- are in the extreme east and extreme south-west suggesting relation of fluoride concentration with Gneissic and Schistose rocks.

The observed patterns of various parameters derived as interpolated surfaces relate logically to the probable natural, geological and anthropogenic factors to a significant extent. The R^2 values between the measured and predicted values of various parameters range from approx. 20 percent for TH, Cl^- to approx. 80 percent for F and NO_3. Approximately 50 percent variability is explained for distribution of SO_4, TDS, Na and pH.

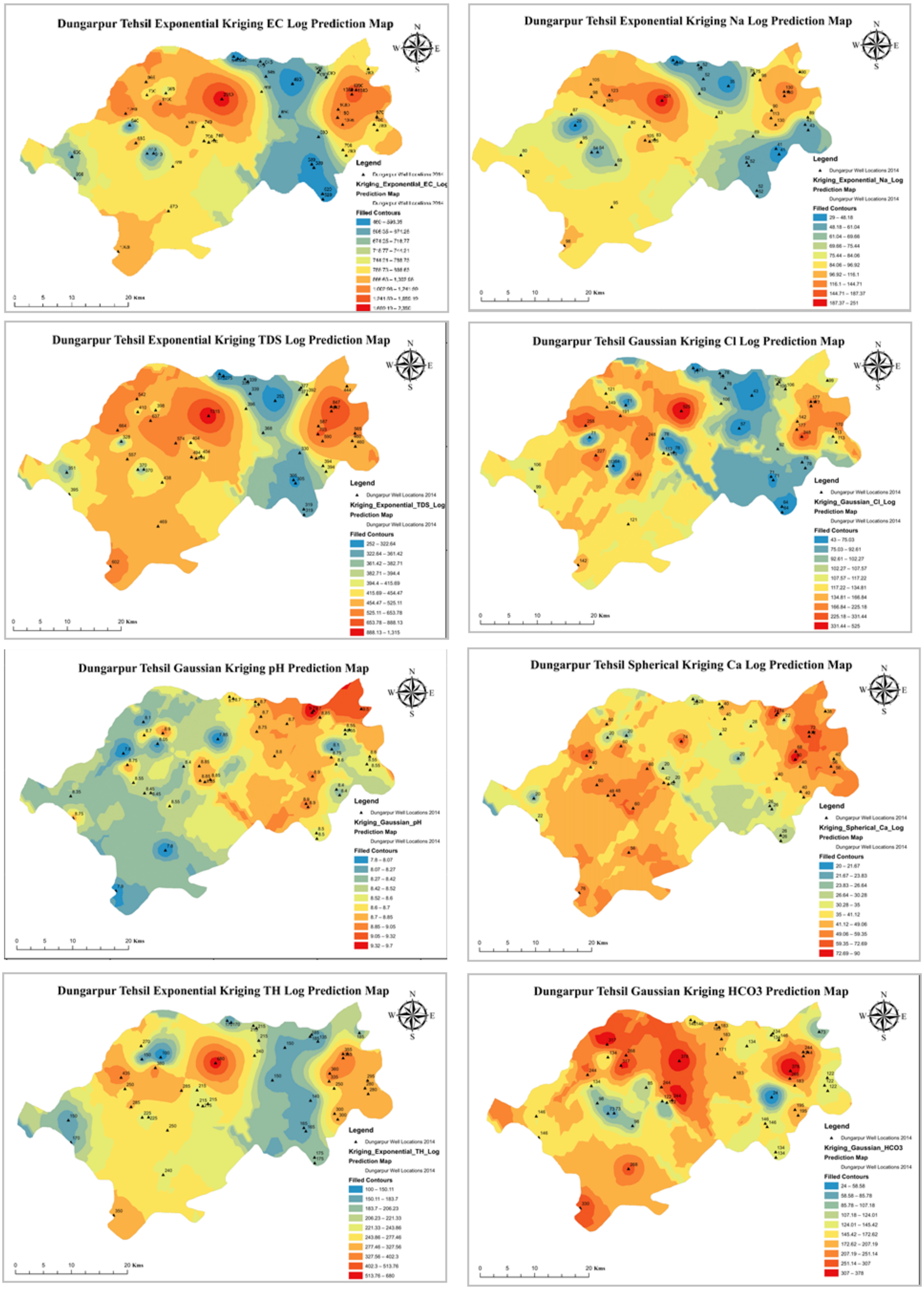

Figure 5a. The interpolation surfaces generated by Kriging with optimal parameters models for EC, TDS, pH, TH, Na, Ca,, Cl and HCO₃.

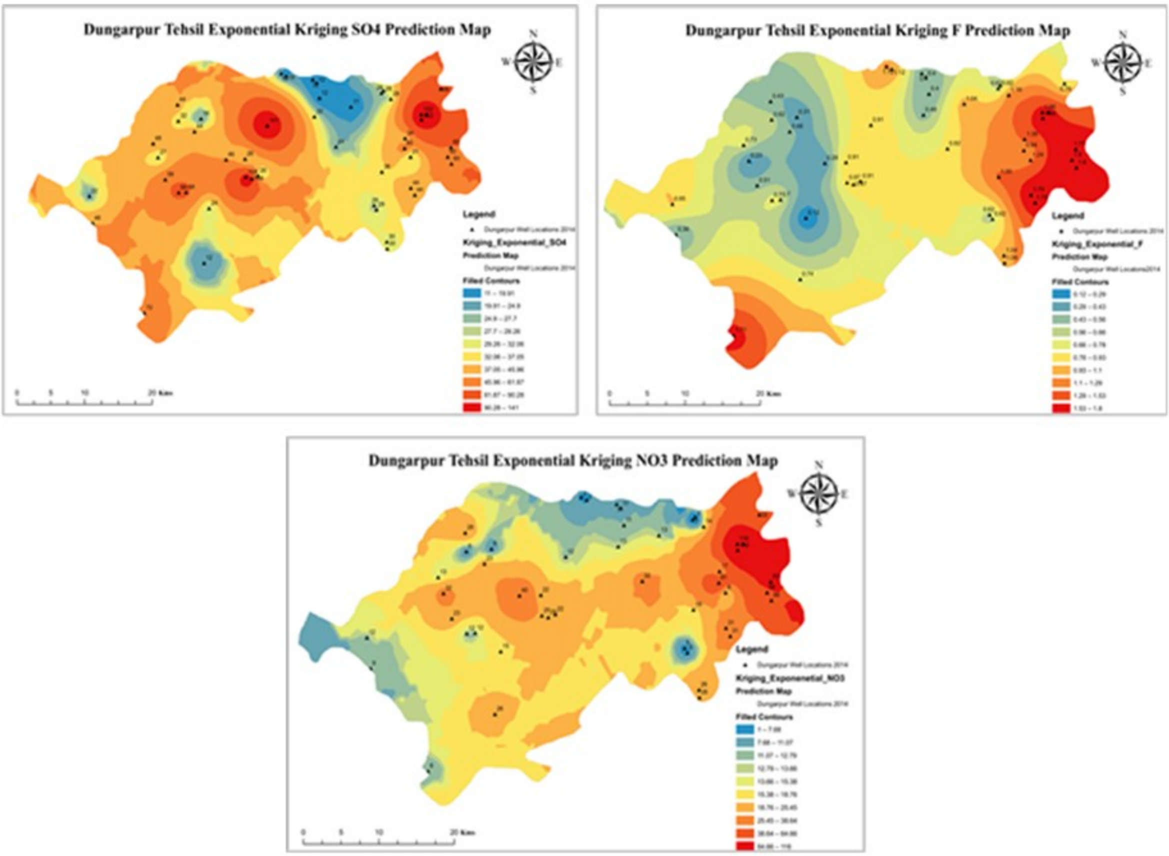

Figure 5b. The interpolation surfaces generated by Kringing with optimal parameters models for F, SO_4 and NO_3

4. Conclusions

It can be concluded that the most of the northern, north-eastern, central and extreme southern regions of the Dungarpur tehsil are affected by higher concentrations of various chemical parameters such as EC, TDS, pH, TH, Na^+, Ca^{+2}, F^-, Cl^-, HCO_3^-, SO_4^{-2} and NO_3^-. Kriging estimation can be considered as a reliable method for generation of interpolated surfaces of ground water quality based on point sample data and selection of best suitable semi-variance model among Gaussian, Spherical and Exponential over the log normal transformed data of water quality parameters. The nugget effect can be employed as an efficient index to illustrate the degree of spatial correlation based on semi-variable function. The performance of the models varies for different parameters. In the present study the distribution of geochemical parameters like Flouride, Nitrate, Alkalinity, TDS, Sulphate and Sodium has been effectively modelled using the Krigging tool. However, the prediction capability for TH and Chlorine has been relatively low.

Acknowledgements

This paper is a part of the study being carried out under major research project funded by Ministry of Human Resource Development, Government of India (MHRD) under Rashtriya Uchhtar Shiksha Abhiyan (RUSA) 2.0 scheme. The authors are thankful to MHRD and State Project Directorate, Government of Rajasthan for funding the project.

References

Ali S.A. and A. Ahmad (2019). Analysing water-borne diseases susceptibility in Kolkata Municipal Corporation using WQI and GIS based Kriging interpolation. *GeoJournal, 85*, 1151-1174.

Bárdossy A. (1997). Introduction to geostatistics. *Institute of Hydraulic Engineering, University of Stuttgart*, 28-32

Burrough P. A. and R. A. McDonnell (1998). Creating continuous surfaces from point data. *Principles of Geographic Information Systems. Oxford University Press, Oxford, UK*.

Garson G. D. (2012). Testing statistical assumptions. *Asheboro, NC: Statistical Associates Publishing*.

Goovaerts P. (1999). Geostatistics in soil science: state-of-the-art and perspectives. *Geoderma*, *89*(1-2), 1-45.

Hu K., H. Chen, Y. Zhang, B. Li and D. Chen (2009). Spatial distribution of shallow groundwater depth, total dissolved solid and nitrate pollution. *Transactions of the Chinese Society of Agricultural Engineering*, *25*(Supp. 1), 21-25.

Chen J, Z. Hanting, Q. Hui, W. Jianhua and Z. Xuedi (2013, June). Selecting proper method for groundwater interpolation based on spatial correlation. In *2013 Fourth International Conference on Digital Manufacturing & Automation* (pp. 1192-1195). IEEE.

Johnston K. (2004): ArcGis 9. Using ArcGIS geostatistical analyst. Redlands, California. *Esri Press (GIS by ESRI).*

Katipoğlu O.M. (2021). Spatial Analysis of Seasonal Precipitation Using Various Interpolation Methods in the Euphrates Basin, Turkey.

McNeely R. N., V. P. Neimanis and L. Dwyer (1979). A guide to water quality parameters. *Water quality source book. Ottawa: Inland waters directorate, water quality branch*, *88*.

Mehrjardi R. T., M. Z. Jahromi, S. Mahmodi and A. Heidari (2008). Spatial distribution of groundwater quality with geostatistics (case study Yazd-Ardakan Plain). *World Applied Science Journal* 4(1):9-17.

Mutuma E. (2017), Innovative approaches of predicting soil properties and soil classes in the eastern slope of Mt. Kenya. *Thesis of Hungarian University of Agriculture and Life Sciences.*

Nas B. (2009). Geostatistical Approach to Assessment of Spatial Distribution of Groundwater Quality. *Polish Journal of Environmental Studies, 18*, 1073-1082.

Nayanaka V. G. D., W. A. U. Vitharana, and R. B. Mapa (2010). Geostatistical analysis of soil properties to support spatial sampling in a paddy growing alfisol.

Olea R. A. (2012). *Geostatistics for engineers and earth scientists*. Springer Science & Business Media.

Osbourne J. W. (2002). Notes on the Use of Data Transformation. *Practical Assessment, Research & Evaluation*, *8*(6), n6.

Shankar B.S. and Shobha (2018). A Comprehensive Assessment of Spatial Interpolation Method Using IDW Technique for the Groundwater Quality Evaluation of an Industrial Area in Bangalore, India.

Tatalovich Z., J. P. Wilson and M. Cockburn (2006). A comparison of thiessen polygon, kriging, and spline models of potential UV exposure. *Cartography and Geographic Information Science*, *33*(3), 217-231.

Voltz M. and R. Webster (1990). A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. *Journal of Soil Sciences. Wiley Online Library.*

Zeybek M. and I. Şanlıoğlu (2020). Investigation of landslide detection using radial basis functions: a case study of the Taşkent landslide, Turkey. *Environmental Monitoring and Assessment, 192*, 1-19.