Groundwater Quality Indexing Optimization Based on Geographical Information System in Lahore-Pakistan

Khalid Mahmood[1], Iqra Basit[2], Hafiza Qimrah[3], Rida Batool[4], Samia Rubab[5] and Saira Ramzan[6]

Abstract

The study has been intended to find best possible solution for developing continuous spatial distribution of groundwater quality, being used for drinking in the mega city of Lahore, Pakistan. A variety of interpolation methods have been compared to map spatial trends in different indicators of groundwater quality in the most realistic and effective way. For this purpose, concentration of nine groundwater parameters (Alkalinity, Ca, Cl, Conductivity, Hardness, Mg, pH, TDS and Turbidity) for 348 drinking water production wells have been used with their spatial reference. Ultimate accuracy check over different interpolation methods is applied using contour analysis, surface analysis, semivariogram, error plots, average standard error and root mean square error (RMSE). Kriging family, particularly simple kriging, is found to be most suitable for predicting most of the groundwater quality surfaces with the existing distribution of the production wells. The only exception found is for predicting Cl surface that has been best developed using redial basis function. RMSEs of the beast predicted surfaces for Alkalinity, Ca, Cl, Conductivity, Hardness, Mg, pH, TDS and Turbidity are 69.84, 19.04, 19.66, 212.10, 73.65, 10.21, 0.15, 131.70 and 0.86 respectively.

Key Words: Ground Water Quality Index, Spatial Interpolation, Surface Analysis, Contour analysis, Semivariogram.

Introduction

Quality of the groundwater is an important parameter to be checked before its use for drinking, mining, irrigation and livestock production etc. It requires a long term management plan that includes assessment of bio-chemical condition of the water that governs its reliability for the usage (Dixon and Chiswell, 1996; Shrestha and Kazama, 2007). According to Gabriel H. F. (2010) more than 50% of the world's population depends upon ground water resources. Developing countries are suffering from health problems associated with either the lacking quantity or contamination of water. It is estimated that till 2025 the count of sufferers of water scarcity will reach to 3 billion, 83% of which will belongs to the developing world. The situation in Pakistan is not very much different from the global trend, Approximately 30% of all diseases and 40% of all deaths are directly or indirectly relates to the intake of contaminated water (water vision 2025 report, 2000; Haydar et al., 2009). The growing needs of portable

[1] *Department of Space Science and Remote Sensing, GIS and Climatic Research Lab (National Center of GIS and Space Application) University of the Punjab, Lahore, Pakistan,*

Email: khalid.spsc@pu.edu.pk; khalid.m270@yahoo.com

[2] *Remote Sensing, GIS and Climatic Research Lab (National Center of GIS and Space Application) University of the Punjab, Lahore, Pakistan*

[3] *Department of Space Science, University of the Punjab, Lahore, Pakistan*

[4] *Department of Space Science, University of the Punjab, Lahore, Pakistan*

[5] *Department of Space Science, University of the Punjab, Lahore, Pakistan*

[6] *Department of Archaeology, University of the Punjab, Lahore, Pakistan*

water have acknowledged conservation of water resources as a political issue as well for entire globe.

The quality check of groundwater comprises on spatial analysis of many bio-physical and bio-chemical parameters. To make things easier researchers has developed Geographic Information System (GIS) based Water Quality Index (WQI) that gives spatial distribution of an overall pollution level of the groundwater resource. WQI is a numeric expression that interpret mixing ratio of contamination to the ground water (Arias et al., 2012). Related parameters to this assessment method i.e. pH level, water table depth, temperature, electric conductivity etc. are inherently continuous and their absolute measurement is quite impossible. Therefore researchers, all over the world, take limited experimental observations from field and then use spatial interpolation that made data continuous in the study area before it get used as input to WQI.

Spatial interpolation utilizes different algorithms for prediction at un-sampled location in the hope to acquire exactness for different scenarios (Mahmood et al., 2014). Most of the researchers used inverse distance weighted (IDW) technique for the analysis of spatial variability of pollutant concentration in groundwater (Chen and Liu, 2012; Tiengrod and Wongsere, 2013). With the advances in methodologies researchers have start using kriging as a better option for this assessment. Kriging not only relies on the basic law of geography but also provides unbiased predicted values with a control of variance. Spatial autocorrelation and prediction error plots are the most significant features in this regard. Kriging has various types as well, but which one to choose? is still an ignored question to be incorporated in WQI related studies.

Ordinary kriging is the most usual type of kriging. It surmises that the trend is unknown and constant (Erdogan, 2009). The formula of ordinary kriging (Kumar et al., 2007)

$$Z^*(u) = \sum_{\alpha=1}^{n(u)} \lambda_\alpha(u) Z(\mathrm{u}_\alpha) + \left[1 - \sum_{\alpha=1}^{n(u)} \lambda_\alpha(u)\right] m$$

Where *z(u)*is estimated value at *u* point, *n(u)* is number of data points, $z(u_\alpha)$ are the *n* measured values at place u_α close to *u.m* is mean of distribution, $\lambda_\alpha(u)$ is weights. Simple kriging, on the other hand, presumes that trend is known and constant but with a random variable (Zhang and Srinivasan, 2009).

$$Z^*_{SK}(u) = m + \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{SK}(\mathbf{u})\left[Z(\mathbf{u}_\alpha) - m\right].$$

Existence of a prominent trend in the data lead towards use of stochastic method of universal kriging to predict the variation efficiently (Johnston et al., 2001; Wang et al., 2014).

With a variety of perdition algorithms, kriging made use of semivariograms and variograms to further aid the prediction process. Semivariogram and variogram are the diagrammatic display of autocorrelation and similarity as function of distance between data points (Milillo et al., 2006). The evolution of semivariogram is not a big deal but the choice of suitable semivariogram for spatial data is time consuming (Teegavarapu et al., 2012). Experimental semivariogram describe the extent of measuring the average difference between known and unknown sample points. Value of the experimental semivariogram at distance $\boldsymbol{h}$ is defined as half of the average squared difference between the value Z($\boldsymbol{x}$) and the value Z(x + h) (Lark, 2000; Robinson and Metternicht, 2006).

$$\wp(h) = \frac{1}{2N(h)} \sum [z(x) - z(x+h)]^2$$

Where $\wp(h)$ is the semivariogram and $2\wp(h)$ is variogram that has the information of spatial correlation (Gelfand et al., 2012). To point out the spatial variation, the geostatistical parameters are Nugget, Sill, Range and Lag (Su et al., 2009).
Nugget is intersect of the semivariogram with the vertical axis. Ideally Nugget value should be zero, but in real case it exists due of sampling error (Ashiq et al., 2010). Smaller is the weight assign to the nearest point, larger is nugget values (Li and Heap, 2008). Sill is the value where semivariogram flatten out (Zawadzki et al., 2005). Range defines size of neighborhood window used for prediction such that the autocorrelation get established (Eastman, 2001; Li et al., 2006; Li and Heap, 2008). If the correlation exist in data points over a long range, variogram will have least nugget effect and vice versa (Srivastava, 2013). The lag is the separation distance between points. Other parameters that control reliability of experimental variogram are sample size, lag interval, bin width, anisotropy, trend (Oliver and Webster, 2013).
Larger sample size gives better accuracy of interpolation (Oliver and Webster, 2013). In general at least a set of 100 data points are required for semivariogram (pavao et al., 2012). For randomly spread data, the comparison must be grouped by distance as well as direction that requires adjustment of bin width and lag interval. Narrow bin width and short lag interval yields noisy data (Oliver and Webster, 2013). Anisotropy exists when spatial autocorrelation varies in each and every direction (Pokhrel et al., 2013). Trend can be identified by mapping data with a suitable graphic program or by fitting a simple trend surface such as linear or quadratic (Oliver and Webster, 2013).
Another variable in kriging is variography that fits a model of spatial-dependence to data in order to reduce prediction error (Ibrahim, 2011). The available models include spherical, exponential, linear and gaussian which rely on auto correlation (Hoef et al., 2004; Bargaoui and Chebbi, 2009).
Linear semivariogram model is the simplest in this family and it uses a trend of linear increase with distance (Mazzella and Mazzella, 2013).

$$\wp(h) = A^* \frac{h}{r} \qquad for \in (0, r)$$

Spherical semivariogram model is used when there exists a clear range, sill and small value of nugget (Zhang and Srinivasan, 2009). It shows linear behavior close to the origin (Goovaert, 2000; Zhang and Srinivasan, 2009).

$$\wp(h) = \begin{cases} c\left[1.5\frac{h}{a} - 0.5\left(\frac{h}{a}\right)^3\right] & for\ h \le a \\ c & for\ \ h > a \end{cases}$$

When range approach progressively to clear nugget and sill then exponential semivariogram model is favored. It also behaves linearly near the origin (Sunila and Kollo)

$$\wp(h) = C_\circ + A^*\left(1 - e^{\frac{h}{r}}\right) \qquad for\ h > 0$$

Preference of gaussian semivariogram model is associated with small nugget having smooth variation (Zhang and Srinivasan, 2009). It shows parabolic behavior at the origin.

$$\wp(h) = C_\circ + A^*\left(1 - e^{-\left(\frac{h}{r}\right)^2}\right) \qquad for\ h > 0$$

This research has been intended to find best interpolation method for studying water quality parameters such as Alkalinity, Ca, Cl, Conductivity, Hardness, Mg, pH, TDS and Turbidity. It helps in computing water quality index with best predicted inputs to have the true inside of the local groundwater quality.

Materials and Methods

Study Area

Lahore is the provincial capital and second largest city of Pakistan lying between 31°15′-31°45′ N and 74°01′- 74°39′ E. with an area of 1014 km². Hydrogeologically it is a part of inter fluvial Bari Doab which is surrounded by River Ravi to North West (Basharat and Rizvi, 2011). All the domestic needs of water in Lahore are being fulfilled by abstracting groundwater through a great number of production wells. Rapid urbanization, improved living standards and industrialization has put an extra ordinary pressure on local aquifer from last couple of decades. This over usage along with decreasing recharge has results into detrition of both quality and quantity of the groundwater (Gabriel and Khan 2006). Water and Sanitation Agency (WASA), Lahore supervise the provision of water supply, sewerage, disposal and drainage collection in the study area.

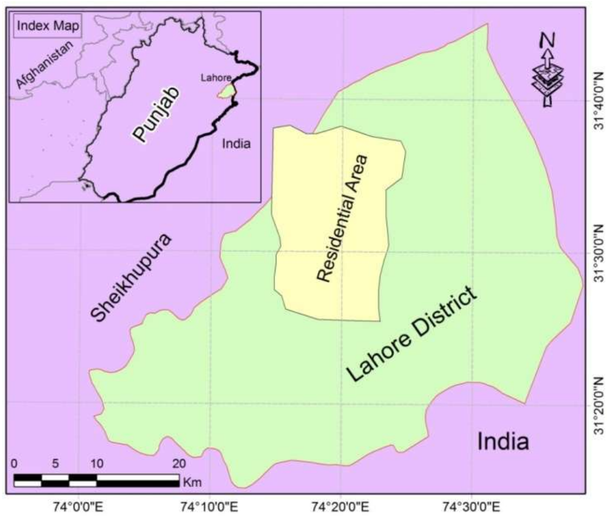


Figure 1: Study Area

Data Collection:

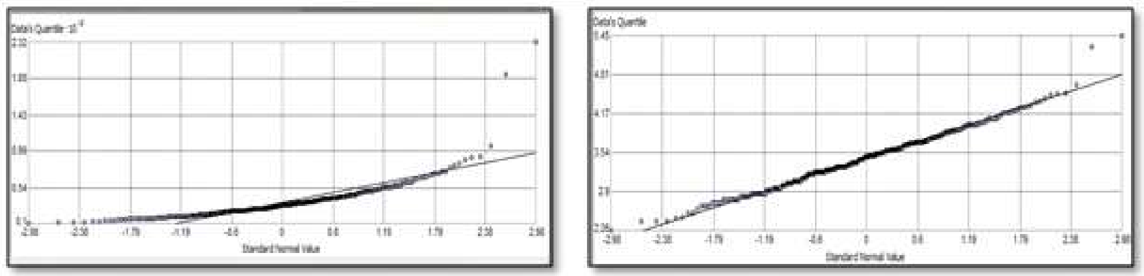
Sample water quality data of 348 production wells has been acquired from Water and Sanitation Agency (WASA), Lahore. The dataset consists of some physical i.e. Turbidity and Hardness and some chemical parameters i.e. Cl, Mg etc. of the year 2013. The location information is collected with an accuracy of ±3m using GPS model Garmin GPSmap 76CSx. All the quality as well as location data was then prepared as excel data sheets. Attribute data exported in GIS software for further processing and analysis. Statistical variations of datasets are given in table 1

Table 1: Initial characteristics of Sample Data

Parameters Properties	Alk.	Ca	Cl	Cond.	Hardness	Mg	pH	TDS	Turb.
Sum	80444.4	12517.1	11758	221991	61424	7299.2	2848	140307.1	224.58
Mean	224.7050	34.9640	32.8436	620.0866	171.5754	20.3888	7.9553	391.9193	0.6271
Std.	83.1095	20.3575	26.2607	288.7539	82.4081	11.2687	0.1558	180.6994	0.8721
Min	22.6	9.6	08	108	60	5.7	7.6	137.7	0
Max	590	232	188	1745	784	102	8.3	1099.2	10.96
Range	567.4	222.4	180	1637	724	96.3	0.7	961.5	10.96

Exploratory data analysis has been performed to check data consistency, removing outliers and identifying statistical distribution. The histograms and normal Q-Q plots have been plotted to

check the normality of the observed data as shown in Figure 2.Histogram analysis helps to remove the outliers and made the data smooth by applying log transformation.
Taking Ca as the explanatory example, its histogram analysis showed that its concentration data was positively skewed, but after applying transformation the skewness reduces down to -3.56 and has achieved the Gaussian distribution. The related Q-Q Plots are shown in figure 2.



Before Applying Transformation (A) After Applying Log Transformation (B)

Figure 2: Transformation Applied

Table 2 elaborates histogram properties for each of the parameters after improving their noramality using suitable transformations.

Table 2: Histogram Analysis of water Quality Parameters

Parameters / Properties	Alk.	Ca	Cl	Cond.	Hardness	Mg	pH	TDS	Turbidity
Transformation	Log	Log	Log	Log	Log	Log	Log	Log	None
Mean	5.3462	3.4487	3.2778	6.3278	5.0515	2.904	2.0735	5.874	0.63658
Median	5.3471	3.4657	3.1781	6.3725	5.0499	2.8792	2.0669	5.906	0.43
Std.	0.37544	0.4556	0.6144	0.4559	0.4275	0.4664	0.0193	0.4388	0.87846
Skewness	-0.7494	0.3783	0.6521	-0.0956	0.2002	0.3606	0.4544	0.1426	5.3923
Kurtosis	6.8408	4.0214	3.202	3.028	3.0948	3.3761	2.272	2.4138	57.563

Accuracy Check Meters:

Comparison of different outputs for their suitability can only be possible through some check meter that may be a visual analysis or some mathematical formulation. For this study following mathematical and visual check meters have been used to compare all possible interpolation results for each of the selected water quality parameters.

Root Mean Square Error:

Root mean square error (RMSE) is the under root of average square difference between measured and estimated value (Gundogdu and Guney, 2007). The mathematical expression of RMSE as described by Siska and Hung, 2001 is as under.

$$\mathrm{RMSE} = \sqrt{\frac{(Z_i - Z)^2}{n}} = \sqrt{\frac{{SSE_i}^2}{n}}$$

Where SSE is the sum of error (observed – predicted) and n is the number of pairs (Su et al., 2009).

Standardized Root Mean Square Error:

Smallest RMSE and standardized RMSE should be as closer to unit value as possible for the best interpolating member of the kriging family (Chang, 2010). The formula for standardized RMSE is (Adhikary et al., 2014).

$$\text{Standardize RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left[{(Z_i - Z)}/{\sigma_i}\right]^2}{n}}$$

Mean Absolute Error:

The MAE results are often preferred over the RMSE as an evaluator because it is less sensitive to extreme values (Chung and Rogers, 2012). Mathematically mean absolute error can be written as (Goovaerts, 2010).

$$\text{Mean absolute error} = \frac{1}{n}\sum_{i=1}^{n}[Z_i - Z]$$

Average Standard Error:

Overestimation of variability in prediction occurs when average standard error is larger than root mean square prediction error and vice versa (Millillo et al., 2006). Mathematically it is given as (Goroi and Kumar, 2013)

$$ASE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\tilde{\sigma}^2(X_i)}$$

Where $\tilde{\sigma}^2$ (x_i) is kriged variance at location Xi

Mean Squared Error:

It is average difference between measured and predicted values.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}[z(x_i) - z^*(x_i)]^2$$

Where *n* is variable (Lopez-Granados, 2005).

Cross Validation:

Cross validation is quick method for equating prediction and measured values (Omran, 2012). It uses all dataset to estimate the trend and autocorrelation models. While performing cross validation, mean error and RMSE should be calculated for each interpolation method.

Error Plots:

The bias between the predicted and measured value is used to evaluate mean error, root mean square error, mean standardize error, root mean square standardized error (Chung and Rogers, 2012).

Contours:

Contours contribute to join adjacent points of equal data values. Quality, reliability and distribution of contours helpful to find the optimal method (Trochu, 1993). Number of neighbors controls texture of contours the (large no of neighbors produce smoot pattern of contours and vice versa) leading to the substantial variation in estimates (Varouchakis and Hristopulos, 2013).

Correlation Coefficient:

Correlation coefficient is the sum up the correlation between measured and estimated value (Sun et al., 2009). High correlation coefficient defines the superior model.

$$R = \left\{\frac{\sum_{i=1}^{N}(Z_i^{\circ} - Z^{\circ})(Z_i^{p} - Z^{p})}{\left[\sum_{i=1}^{N}(Z_i^{\circ} - Z^{\circ})^2\right]^{0.5}\left[\sum_{i=1}^{N}(Z_i^{p} - Z^{p})^2\right]^{0.5}}\right\}$$

Where Z_i^p the estimated value is Z_i° is the observed value (Zang and Srinivasan, 2009).

Selection of Interpolation Method:

The present study is based on the comparison of different interpolation methods for generating accurate distribution surfaces to aid water quality indexing. Spatial interpolation techniques are mainly of two types deterministic and Stochastic Methods.

Deterministic Methods

Power factor that control influence of neighbors depending upon distance is the key element IDW. So forth the concentration values of Ca are estimated by using optimize power 1 corresponding to the four sector with the directional angle 46° and 12 neighbors at the most. Global and Local interpolators interpolates at different order of polynomials. Both global and local polynomials have used the same order of 1 for interpolating alkalinity. For Calcium, completely regularized spline has been used corresponding to the anisotropic factor equals of 2.35. Neighborhood are selected by using four sector with 45° ellipse in such a way that each sector includes at least 8 neighbors with the directional influence of 46° from north, which is demonstrated in figure 3.

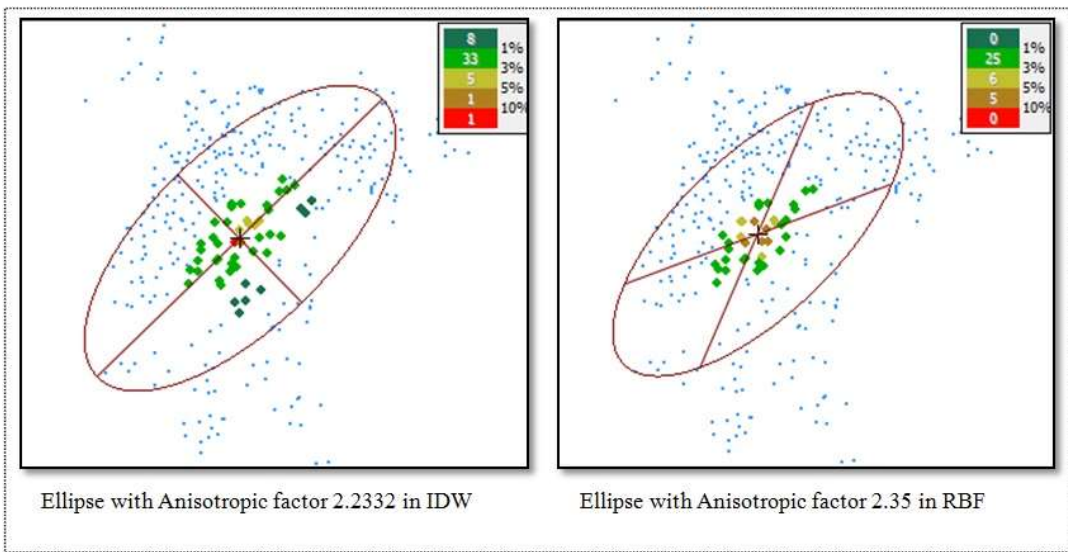


Figure 3: Directional Distribution of Ellipse for Ca

Similarly all these methods are applied to other water quality parameters (Alkalinity, Cl, Conductivity, Hardness, Mg, pH, TDS and Turbidity). Table 2 illustrates the resultant statistics of the best deterministic method for each water quality parameter:

Stochastic Methods:

Normality of the data plays a key role in kriging analysis. Kriging operates different algorithmic models of semivariogram (Spherical, Exponential, Gaussian and Circular) correspond to transformation and trend. All the parameters including nugget, sill, partial sill, lag size and directional angle have been calculated for each of the quality parameter. The data in hand was almost homogeneously distributed along northeastern and southwestern directions. A clear indication of equal scale variability along the directional angle of 45° to up to 49° is showing presence of the recharge sources in northeastern and southwestern edges of the study area as shown in figure 4.

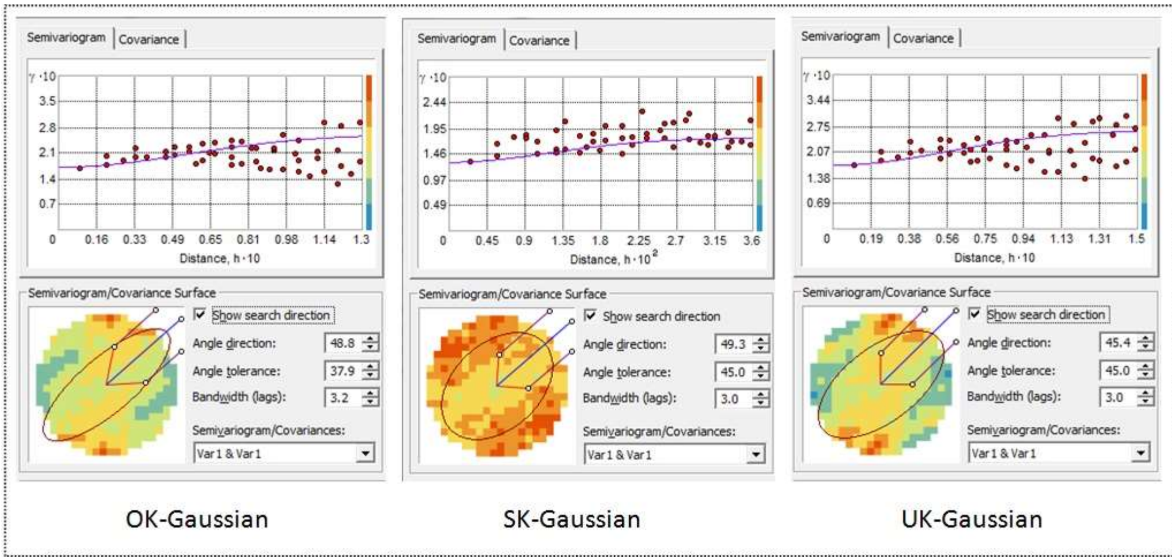


Figure 4: Anisotropic behavior in different types of Kriging Methods at different angles for Ca

Anisotropic models fitted upon the surface to indicate the directionality influence as well as selection of neighborhood behavior. Similarly different semivariogram models have been fitted over the water quality parameters as per influencing statistical parameters. A common overriding trend of estimated values lying along the northwestern direction (River Ravi) and southwestern direction (Lahore Canal). In the following table, all the parameters associated to the best fitted models of variography are mentioned for selected best kriging member for the ground water quality assessment.

Table 3: Parameters of Probabilistic Interpolation Methods

Properties / Parameters	Method	Lag Size	Partial Sill	Nugget	Minor Range(m)	Major Range(m)	Angle Direction (Degree)
Alkalinity	SK_S	0.003	0.03	0.06	0.01855	0.04	51
Ca	UK_C	0.0135	0.024	0.165	0.0383	0.124	52.9
Cl	OK_G	0.0031	0.3194	0.25	0.043	0.03999	155
Conductivity	OK_E	0.0027	0.1000	0.3999	0.0220	0.0973	55.4
Hardness	UK_S	0.0020	0.037	0.1389	0.03	0.0391	54.3
Mg	SK_G	0.0050	0.0999	0.1586	0.1	0.0223	348.5
pH	SK_S	0.0022	0.0001	0.0002	0.0105	0.0193	79
TDS	UK_E	0.0033	0.4	0.0622	1.1	0.2	320.1
Turbidity	SK_S	0.0046	0.1999	0.9008	0.0973	0.0560	11.9

Results and Discussions:

Spatial Analysis of ground water quality parameters generates distinguishable surfaces with certain prominent features. So, it's inappropriate to suspect best interpolated surface for a certain water quality parameter just on the bases of surface analysis. The following table 4 illustrates the statistical analysis of the estimated errors corresponding to the deterministic methods. Maximum proportion of parameters shows best results in completely regularized spline with least RMSE value. Whereas in comparison of deterministic methods the only exception is found for turbidity, where the best interpolation method is found to be the IDW.

Table 4: Error Statistics of Deterministic Interpolation Methods

Water Quality Parameters	Method	RMS	Mean	Sum	Std.	Minimum	Maximum	Range
Alkalinity	RBF	71.12	-0.0187	-6.5132	71.1206	-394.1389	267.9173	662.0562
Ca	RBF	19.27	-0.1659	-57.7480	19.2684	-199.0822	31.0106	230.0927
Cl	RBF	19.66	-0.0229	-7.9763	19.6598	-116.8634	49.4343	166.2977
Conductivity	RBF	217.4	2.1758	757.1760	217.3635	-961.3808	565.8659	1527.2467
Hardness	RBF	74.22	0.7197	250.5144	74.2172	-601.2333	175.3401	776.5734
Mg	RBF	10.44	0.2551	88.7622	10.4348	-80.8697	18.2709	99.1405
pH	RBF	0.1481	-0.0007	-0.2540	0.1481	-0.4234	0.3468	0.7702
TDS	RBF	137	1.6330	568.2752	136.9940	-606.0618	355.1954	961.2572
Turbidity	IDW	0.8508	0.0036	1.2597	0.8508	-10.0289	1.9016	11.9304

As compare to deterministic interpolations, kriging family is more productive due to its ability to incorporate both statistical as well as mathematical concepts for estimating continuous surface for the water quality parameters. Best optimized interpolation method for each of the quality parameter along with the numerical check meter results are given in table 5.

Table 5: Best suited interpolation methods with check meter results

Water Quality Parameters	Method	RMSE	Mean	Std.	Range	ASE	MSE	RMSS
Alkalinity	SK_S	69.84	-0.6299	69.8402	648.2687	64.23	-0.0079	1.073
Ca	UK_C	19.04	-0.2973	19.0438	228.5188	16.08	-0.0258	1.203
Cl	RBF	19.66	-0.0229	19.6598	166.2977	-	-	-
Conductivity	OK_E	212.1	-4.196	212.1030	1604.9210	193.8	-0.0318	1.124
Hardness	UK_S	73.65	0.1275	73.6694	782.2191	75.42	0.0003	0.9796
Mg	SK_G	10.21	0.01132	10.2070	94.8999	9.248	-0.003349	1.094
pH	SK_S	0.1474	-0.0019	0.1474	0.767104	0.1424	-0.0119	1.033
TDS	UK_E	131.7	-2.731	131.7110	1089.7020	185.6	-0.02618	0.718
Turbidity	SK_S	0.8258	-0.0073	0.8257	10.9597	0.8258	-0.0049	0.7812

Error-Plots:

A prediction error plot gives the relationship between measured (x-axis) and prediction (y-axis) values. Y-intercept explains overestimation and underestimated of data values. Comparative analysis of different methods based on error plots for Ca is given in Figure 5. Large difference between dotted and regression line deduce small slope. Local Polynomial has maximum value of Y-intercept that shows overestimation. Ordinary kriging with gaussian model and universal kriging with exponential model have maximum values of slope. Whereas simple kriging and universal kriging both with spherical model result into high value of slope as well as y-intercept. From the figure 5 it clears that ordinary kriging with circular model, ordinary kriging with exponential model and IDW are showing overestimation as depicted by y-intercept and slop. Some interpolation methods such as universal kriging with Gaussian model and simple kriging with circular model both are showing relatively same behavior with a moderate slope. Simple kriging with Gaussian model and RBF have the same value of slope. Universal kriging with circular model gives the best estimation for the Ca with an intermediate value of y-intercept. It can be concluded here that stochastic methods have better prediction results than deterministic methods. Error plots of best suited interpolation for each of the water quality parameter are shown in figure 6.

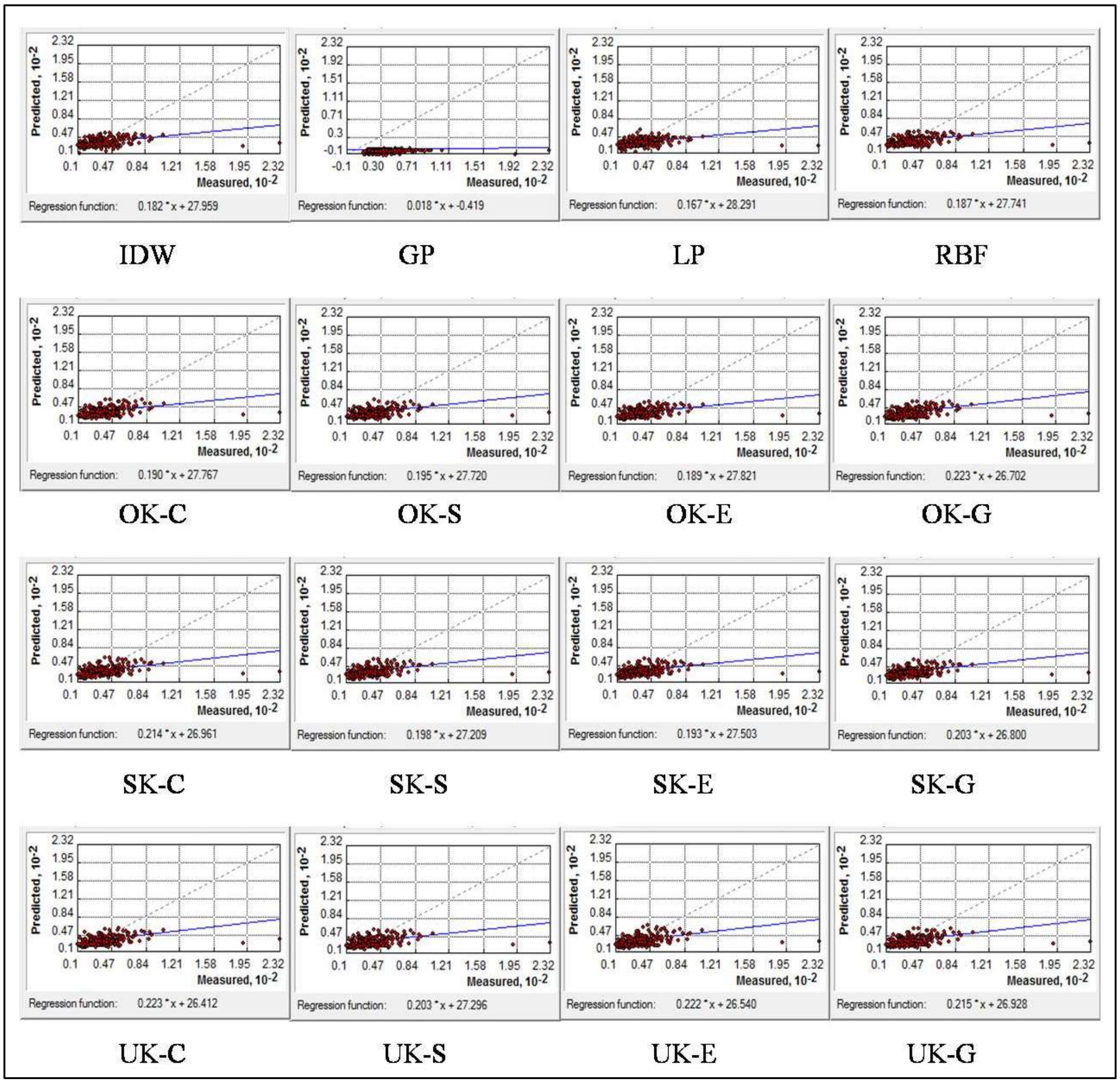


Figure 5: Error Plots of Ca

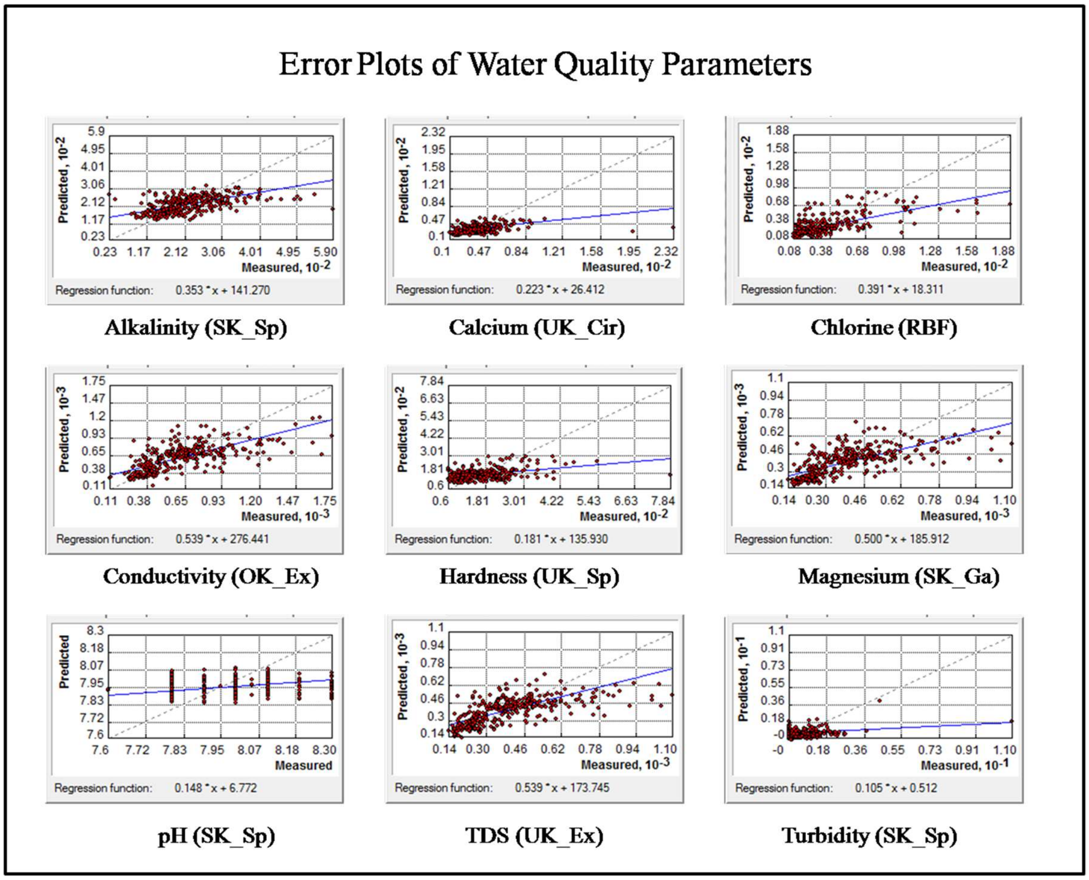


Figure 6: Error Plots of Water Quality Parameters

Semivariogram Analysis:

Ideally trend line should passes diagonally from the points and no nugget value exists in semivariogram but it's nonviable. Taking example of alkalinity, almost all the models, when fit to universal kriging have produced good results. But the best correlation of semivariogram values to distance has been explained by spherical model fitted to simple kriging. In case of Ca, small variation exists in data spread and value of nugget in ordinary and universal kriging models. Whereas simple kriging behaves differently in each semivariogram model corresponding to the nugget, sill and partial sill, as shown in the figure 7. Simple kriging with gaussian model illustrates the best fitted trend line.

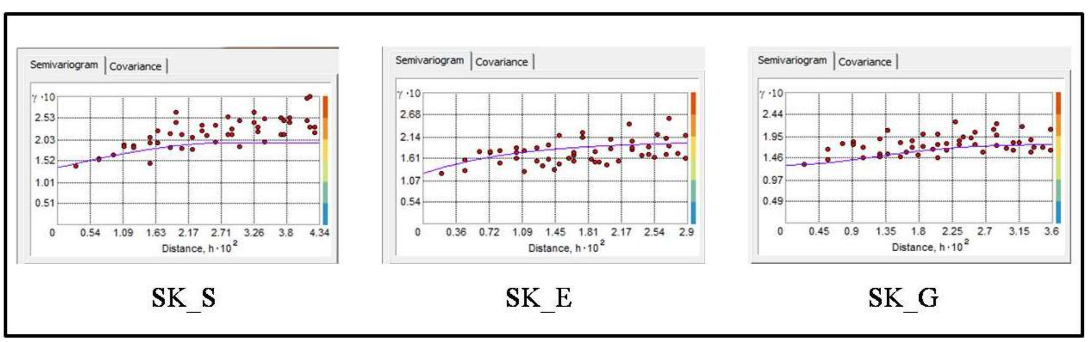


Figure 7: Comparison of the Semivariogram of Simple Kriging for Ca

For Conductivity, universal kriging with almost all the models, except gaussian, display smoothly increasing trend that tends auto correlation decrease and covariance increase slickly. However, exponential ordinary kriging, gaussian simple kriging, spherical universal kriging shows that a sharp trend exist in data point and an obvious range exists in concentration values of Conductivity. Semivariogram analysis for the Hardness dipicts that the all the models of ordinary kriging represents all most flat semivariogram with the least correlation factor. Whereas simple kriging and universal kriging signify correlation upto maximum extent particularly in spherical and exponential models. Best auto correlation has been observed in simple kriging prominently in with spherical model for the pH. The best results for TDS with the high correlation factor estimated by simple kriging with circular model and universal kriging with exponential model. Semivariogram analysis of Mg and Turbidity has shown that simple kriging explained well the trendy behavior of this variable. In case of Mg simple kriging gaussian model explains unique trend that asymptotically approaches the range and spherical simple kriging exhibits a sharp change of trend line. Resultant best semivariogram for each of the quality parameters is shown in figure 8.

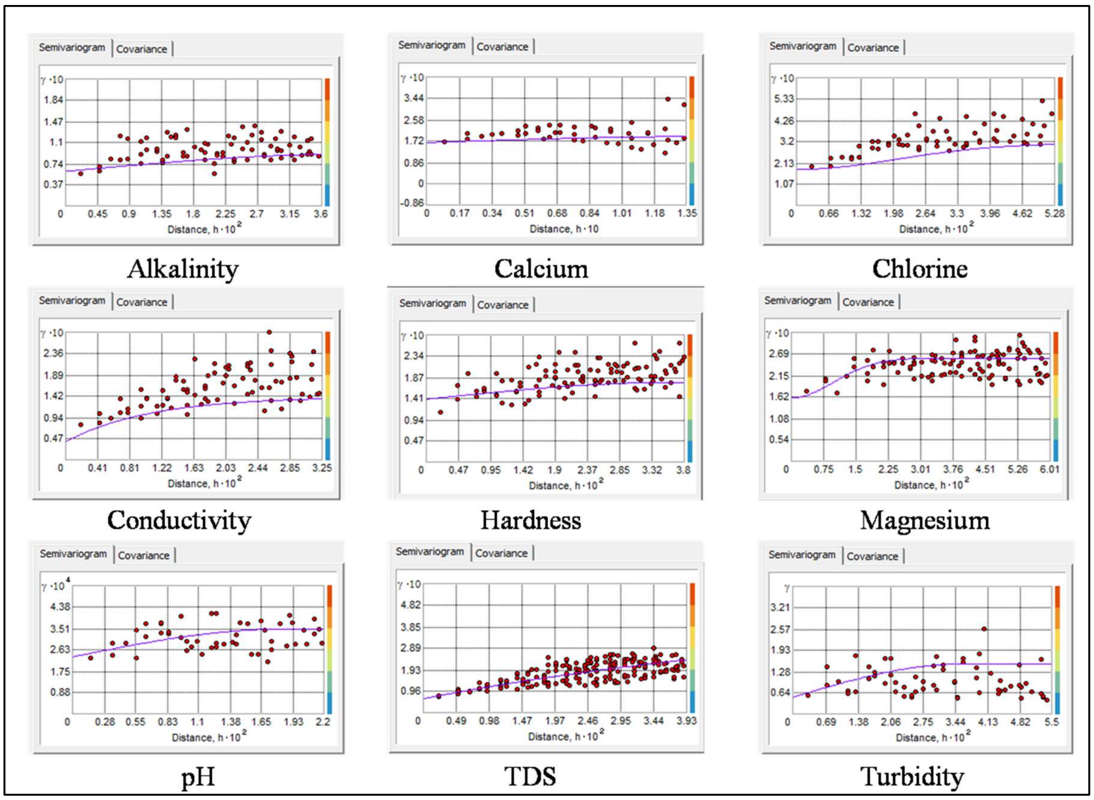


Figure 8: Best Semivariogram of Quality Parameters

Contour Analysis:

Aquifer of Lahore consists of unconsolidated alluvial deposits of quaternary age surrounded by river Ravi to North West (Basharat and Rizvi, 2011). Groundwater quality of Lahore is mainly control by industrial effluents in river Ravi and central canal of the city. Contour maps assist to analyze the variation and trend of concentration values along the depression zone lies in the North western direction (Mahmood et al., 2013; Mahmood et al., 2016). GPI failed to define high or low values because of irregular contours. Whereas IDW and LPI underrated high scale variability not only along the depression zone but also throughout the study area for all the studied quality parameters. RBF has also elaborated the central depression zone with small scale variability and sharp edges. In case of Cl contours appered edgy and located on the uper left of the study area. Stochastic methods have given visuals of not only the central hot spot but also the spatial trends in the entire region. Comparison of different stochastic methods for Ca has been given in the figure 9. Though universal kriging with circular model estimated with more precise RMSE, but contour generated by simple and ordinary kriging are smoother and regular. A small spot of Ca concentration has been found to be intruded from northeastern edges of the study area showing existence of Ca source their.

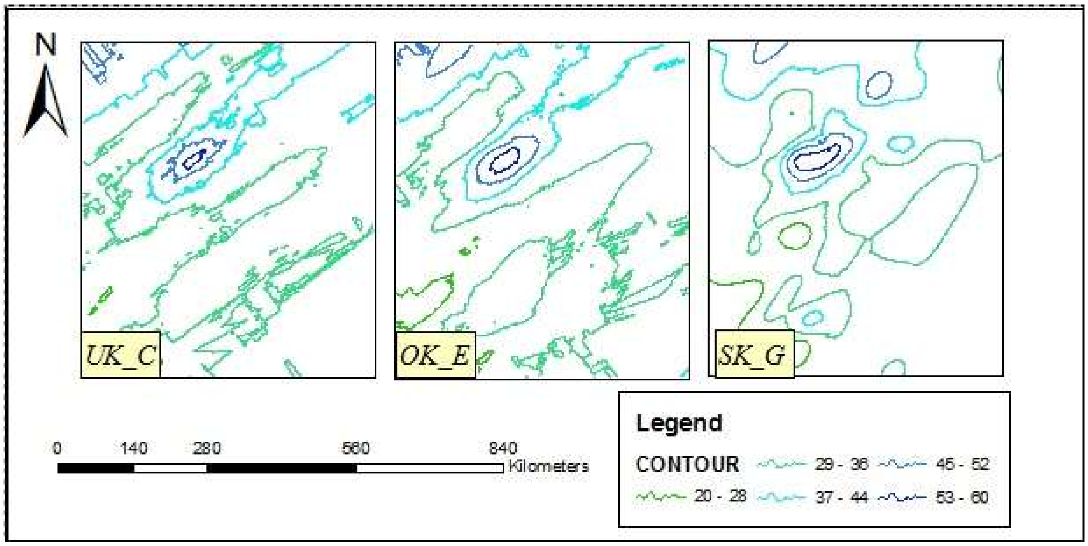


Figure 9: Comparison of Stochastic Methods for Ca

In Conductivity, the contours generated by the universal kriging were very much sharp and intermixed that provide crude information. Whereas the contours analysis of simple and ordinary kriging with exponential model demonstrate well the upper part with sharp variations. Due to shorter range of pH concentration values, contours are widely separated for it, giving almost no useful information. Hardness and TDS demonstrate the central depletion zone more precisely in the universal kriging with smooth variation and diverse sort of contours. Visuals of best suited contour generated for individual quality parameters are show in figure 10.

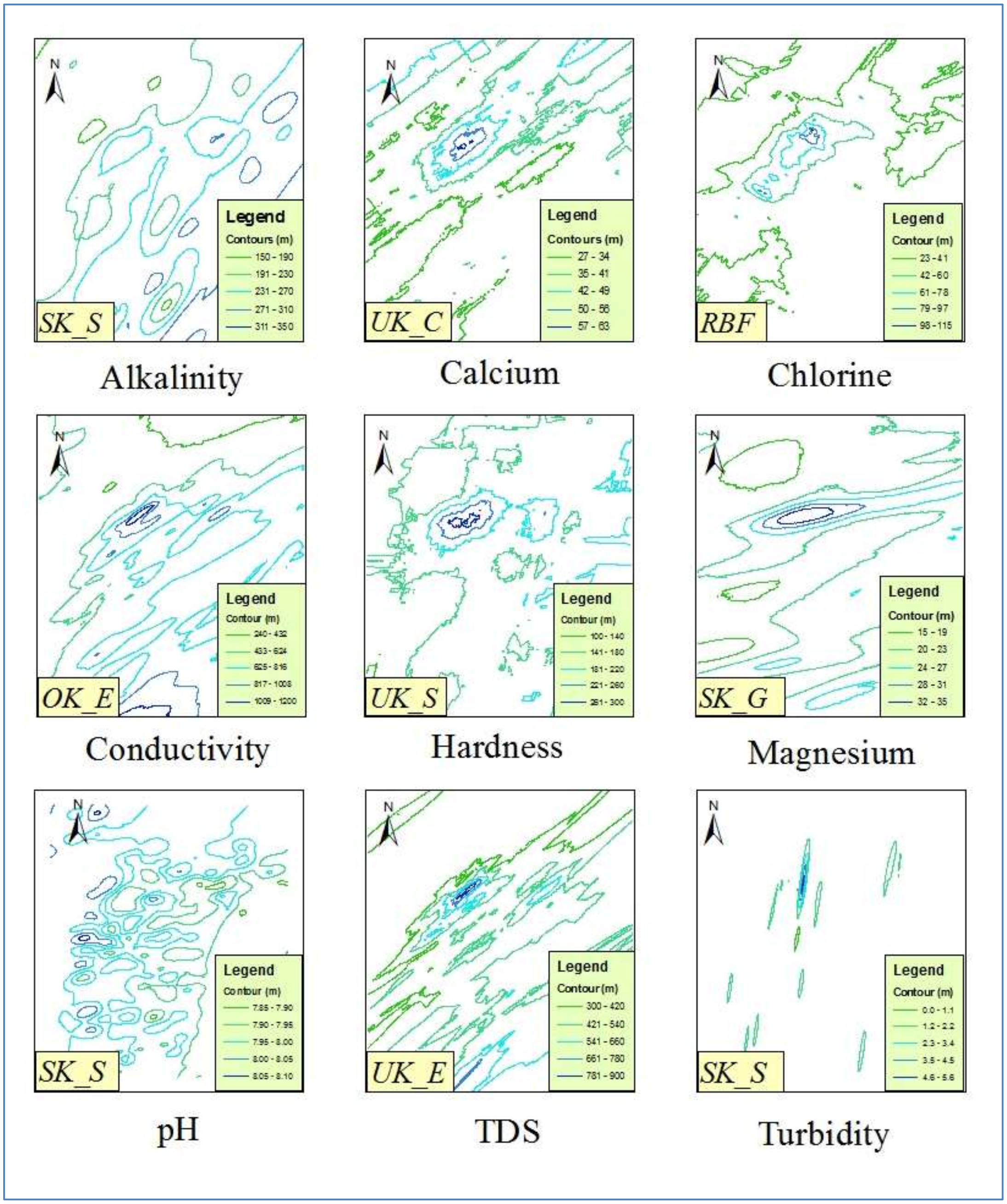


Figure 10: Contour Analysis of Different Water Quality Parameters

Surface Analysis

Surface analysis helps to reveal trends of low or high concentration of quality parameters across the study area. GPI and LPI have turnout to be week candidates to elaborate variation throughout the surface analysis of water quality parameters. Surfaces furnished by IDW and RBF for Ca have provided better results but best were generated by stochastic class, shown in figure 11.

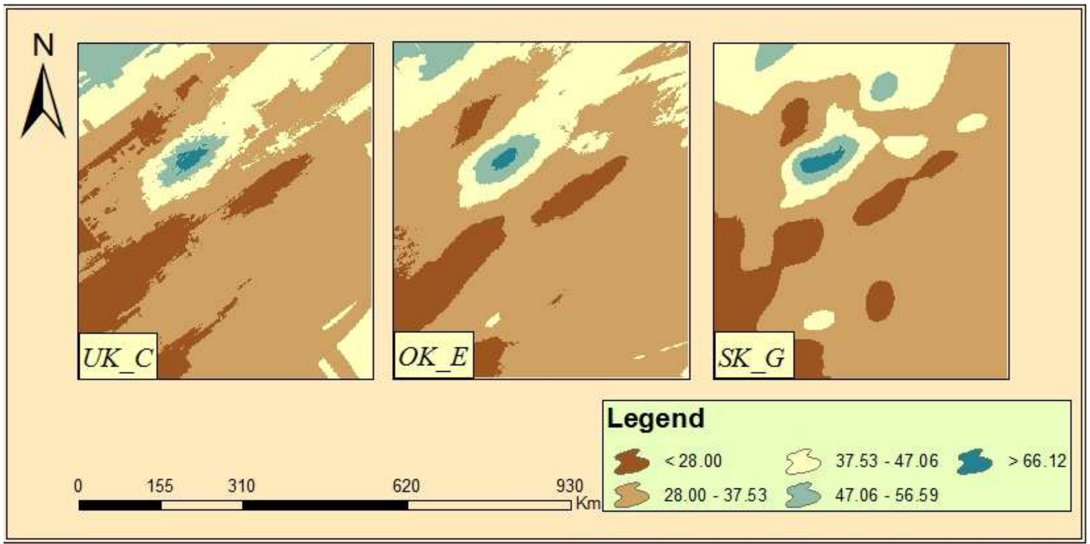


Figure 11: Comparison of Stochastic Methods for Ca

Best developed surfaces along with their suitable interpolation methods for each of the quality parameters are shown in figure 12.

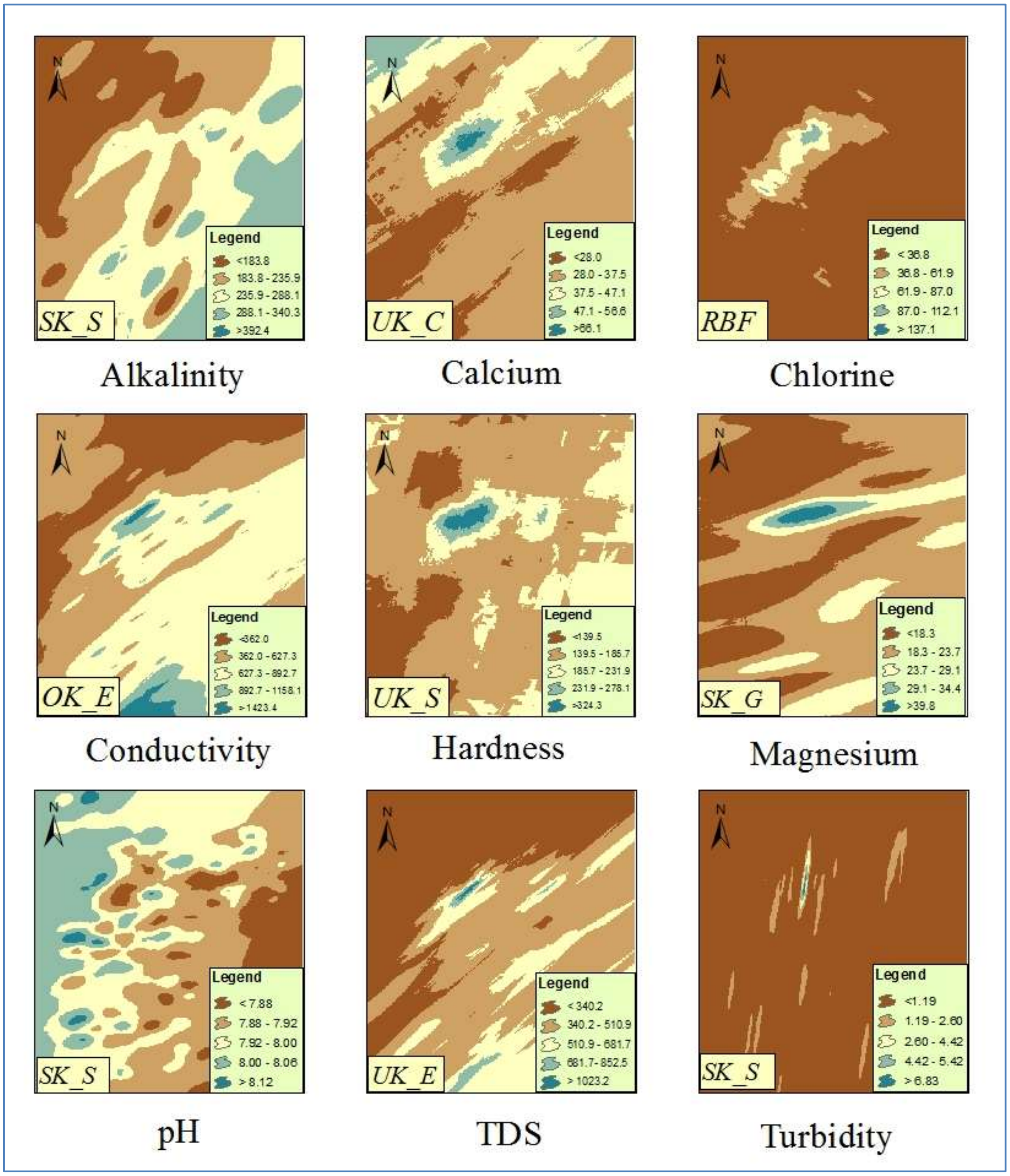


Figure 12: Surface Analysis of the Different Water Quality Parameters

Conclusion

Proper selection of the interpolation method to generate input parametric layers is essential to compute meaningful water quality index. The criteria for the selection of a suitable interpolation method rely on characteristics of sample points, their mutual relationship, concept resemblance between spatial characteristics of measured water quality parameter and formulation of interpolation method. More precise prediction surface for Alkalinity distribution has been acquired using simple kriging with spherical model. For Ca the most precise RMSE of 19.04 has been achieved using universal kriging with circular model. Only Cl is found to be predicted best using deterministic method of RBF. Similarly best methods of interpolation as found for Conductivity, Hardness, pH, TDS and Turbidity are ordinary kriging with exponential model, universal kriging with spherical model, simple kriging with spherical model, universal kriging with exponential model and simple kriging with spherical model with RMSEs as 1.124, 73.65, 0.147, 131.7 and 0.826 respectively. So, the present study conclude that simple kriging belonging, a stochastic method of interpolation is relatively more reliable for the ground water quality analyses of the study area.

References

- Adhikary, S.K., Sharif, A.A., Das, S.K., Saha, G.C. (2014). Geostatistical Analysis of Groundwater Level Fluctuations in the Shallow Aquifer of Northwestern Bangladesh. Proceedings of the 2nd International Conference on Civil Engineering for Sustainable Development (pp. 391-399).
- Arias, H.R., Caraveo, M.C., Quintana, R.M., Teran, R.A., Munguia, A.P. (2012). An Overall Water Quality Index (WQI) for a Man-Made Aquatic Reservoir in Mexico. International Journal of Environmental Research and Public Health, 9: 168-169.
- Armah, F.A., Luginaah, I., Ason, B. (2012). Water Quality Index in the Tarkwa Gold Mining Area in Ghana. The Journal of Transdisciplinary Environmental Studies, 11(2): 2-15.
- Ashiq, M.W., Zhao, C., Ni, J., Akhtar, M. (2010). GIS-based high-resolution spatial interpolation of precipitation in mountain–plain areas of Upper Pakistan for regional climate change impact studies. Theoretical and Applied Climatology, 99: 239-253.
- Azpurua, M., Ramos, K. D. (2010). A comparison of spatial interpolation methods for estimation of average electromagnetic field magnitude. Progress In Electromagnetics Research , 14: 135-145.
- Bargaoui, Z.K., Chebbi, A. (2009). Comparison of two kriging interpolation methods applied to spatiotemporal rainfall. Journal of Hydrology, 365: 66-73.
- Basharat, M., Rizvi, S.A. (2011), "Groundwater Extraction and Waste Water Disposal Regulation – Is Lahore Aquifer At Stake With As Usual Approach", Pakistan Engg. Congress; World Water Day-April 2011: 112-113.
- Caruso, C., Quarta, F. (1998). Interpolation Methods Comparison. *Computers Math. Applic., 35*(12): 109-126.
- Chang, K.T. (2010). Introduction to Geographic Information System. New York, America: McGraw-Hill.
- Chung, J.W., Rogers, J.D. (2012). Interpolations of Groundwater Table Elevation in Dissected Uplands. Ground Water, 50 (4):598-607.
- Dayani, M., Mohammadi, J. (2010). Geostatistical Assessment of Pb, Zn and Cd Contamination in Near-Surface Soils of the Urban-Mining Transitional Region of Isfahan, Iran. Pedosphere, 20 (5): 568-577.
- Dil, A.S., Qazi, I.A., Baig, M.A., Khan, E.A., Tahir, A. (2008). National Standards for Drinking Water Quality. Government of Pakistan Pakistan Environmental Protection Agency.
- Erdogan, S. (2009). A comparison of interpolation method for producing digital elevation models at the field scale. Earth Surface Processes and Landforms, *34* (3): 366-376.
- Gabriel F. H., (2010). An appraisal of climate responsive urban groundwater management options in stressed aquifer system. International Centre of Water for Food Security, School of Environmental Sciences, Faculty of Science, Charles Sturt University.
- Gelfand, A., Diggle, P., Guttorp, P., Fuentes, M. (2012). Handbook of Spatial Statistics. Boston: Taylor and Francis Group, pp. 41-43.
- Goovaerts, P. (2000). Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. Journal of hydrology, 228(1-2): 113-129.

- Goroi, A.K., Kumar, S. (2013). Spatial Distribution Analysis of Groundwater Quality Index Using GIS: A Case Study of Ranchi Municipal Corporation (RMC) Area. Geoinformatics & Geostatistics: An Overview, 1(2): 1-11.
- Gundogdu, K.S., Guney, I. (2007). Spatial analyses of groundwater levels using universal kriging. Journal of earth system science, 116(1): 49-55.
- Haydar, S., Arshad, M., Aziz, J.A. (2009). Evaluation of Drinking Water Quality in Urban Areas of Pakistan: A Case Study of Southern Lahore. Pak. J. Engg. & Appl. Sci., 5: 16-23.
- Hengl, T. (2007). A Practical Guide to Geostatistical Mapping of Environmental Variables (p.13).
- Hoef, J.M., Cressie, N., Barry, R.P. (2004). Flexible Spatial Models for Kriging and Cokriging Using Moving Averages and the Fast Fourier Transform (FFT). Journal of Computational and Graphical Statistics, 13 (2), 265-282.
- Hutchinson, M.F., Gessler, P.E. (1994). Splines- more than just a smooth interpolator. Geoderma, 62: 45-67.
- Ibrahim, A.A. (2011). Production of Digital Climatic Maps Using Geostatistical Techniques (Ordinary Kriging) Case Study from Libya. International Journal of Water Resources and Arid Environments, 1 (4): 239-250.
- Johnston, K., Hoef, J., Krivoruchko, K., Lucas, N. (2001). Using ArcGIS Geostatistical Analyst. United States of America.: Esri Press, pp. 33-116.
- Kumar, A., Maroju, S., Bhat, A. (2007). Application of ArcGIS Geostatistical Analyst for Interpolating Environmental Data from Observations. Environmental Progress, 26 (3).
- Li, B.G., Cao, J., Liu, W.X., Shen, W.R., Wang, X.J., Tao, S. (2006). Geostatistical analysis and kriging of Hexachlorocyclohexane residues in topsoil from Tianjin, China. Environmental Pollution, 142 (3): 567-575.
- Li, j., Heap, A.D. (2008). A Review of Spatial Interpolation Methods for Environmental Scientists. Australia: Commonwealth of Australia.
- Lopez-Granados, F., Jurado-Exposito, M., Pena-Barragan, J.M., Garcia-Torres, L. (2005). Using geostatistical and remote sensing approaches for mapping soil properties. European Journal Of Agronomy, 23: 279-289.
- Mahmood K., A.D. Rana, S. Tariq, S. Kanwal, S.R. Ali, A.H. Ali, T. Tahsin (2013). Groundwater Levels Susceptibility to Degradation in Lahore Metropolitan. Science International, 25:123-126.
- Mahmood K., Ul-Haq Z., Batool S.A., Rana A.D., Tariq S. (2016). Application of Temporal GIS to Track Areas of Significant Concern Regarding Groundwater Contamination. Environmental Earth Sciences, 75:33.
- Mahmood, K., Ali, S. R., Haider, A., Tehseen, T., & Kanwal, S. (2014). Selection of the Optimal Interpolation Method for Groundwater Observations in Lahore, Pakistan. Pakistan Journal of Scientific and Industrial Research , 57 (3): 154-166.
- Mazzella, A., Mazzella, A. (2013). The Importance of the Model Choice for Experimental Semivariogram Modeling and Its Consequence in Evaluation Process. Journal of Engineering (Article ID 960105): 1-10.
- Mehdi, S.M., Mian, S.M., Ghani, S., Khalid, M., Sheikh, A.A., Rasheed, S. A. (2013). Modeling of Surface Soil, pH Using Geostatistical Methods In Punjab Province, Pakistan. International Journal of Scientific & Engineering Research, 4 (11), 359-377.
- Milillo, T.M., Gardella Jr, J.A. (2006). Spatial statistics and interpolation methods for TOF SIMS imaging. Applied surface science, 252 (19): 6883-6890.
- National Standards for Drinking Water Quality (NSDWQ) (2010). Pakistan Environmental Protection Agency, Ministry of Environment, Govt. of Pakistan.

- Ninyerola, M., Pons, X., Roure, J.M. (2007). Objective air temperature mapping for the Iberian Peninsula using spatial interpolation and GIS. International Journal Of Climatology, 27 (9): 1231–1242.
- Oliver, M.A., Webster, R. (2014). A tutorial guide to geostatistics: Computing and modelling variograms and kriging. Catena, 113: 56–69.
- Omran, E.S.E. (2012). Improving the Prediction Auccuracy of Soil Mapping through Geostatistics. International Journal of Geosciences, 3: 574-590.
- Pavao, C.G., Franca, G.S., Marotta, G.S., Menezes, P.H., Neto, G.B., Roig, H.L. (2012). Spatial Interpolation Applied a Crustal Thickness in Brazil. Journal of Geographic Information System, 4: 142-152.
- Pokhrel, R.M., Kuwano, J., Tachibana, S. (2013). A kriging method of interpolation used to map liquefaction potential over alluvial ground. Engineering Geology, 152 (1): 26-37.
- Robinson, T.P., Metternicht, G. (2006). Testing the performance of spatial interpolation techniques for mapping soil properties. Computers and electronics in agriculture, 50 (2): 97-108.
- Shrestha, S., Kazama, F. (2007). Assessment of surface water quality using multivariate. ELSEVIER , 22: 464-475.
- Siska, P.P., Hung, I.K. (2001). Assessment of Kriging Accuracy in the GIS Environment. the 21st Annual Esri International User conference. San Diego.
- Srivastava, R.M. (2013). Geostatistics: A toolkit for data analysis, spatial prediction and risk management in the coal industry. International Journal of Coal Geology, 112 (1): 2-13.
- Su, P., Ting-xuan, L., Yong-dong, W., Hai-ying, Y., Xi, L. (2009). Spatial Interpolation and Sample Size Optimization for Soil Copper (Cu) Investigation in Cropland Soil at County Scale Using Cokriging. Agricultural Sciences in China, 8 (11): 1369-1377.
- Sun, Y., Kang, S., Li, F., Zhang, L. (2009). Comparison of interpolation methods for depth to groundwater and its temporal and spatial variations in the Minqin oasis of northwest China. Environmental Modelling & Software, 24 (10): 1163–1170.
- Tahir, M.A., Rasheed, H., Imran, S. (2010). WATER QUALITY STATUS IN RURAL AREAS OF PAKISTAN. Islamabad: (PCRWR). Retrieved from www.pcrwr.gov.pk on 12 October, 2015.
- Teegavarapu, R.S., Meskele, T., Pathak, C.S. (2012). Geo-spatial grid-based transformations of precipitation estimates using spatial interpolation methods. Computers & Geosciences, 40: 28-39.
- Tiengrod, P., Wongseree, W. (2013). A Comparison of Spatial Interpolation Methods for Surface Temperature in Thailand. Computer Science and Engineering Conference (ICSEC), 2013 International, pp. 174-178.
- Trochu, F. (1993). A contouring program based on dual kriging interpolation. 9 (3): 160-177.
- Varouchakis, E.A., Hristopulos, D.T. (2013). Comparison of stochastic and deterministic methods for mapping groundwater level spatial variability in sparsely monitored basins. Environmental Monitoring and Assessment, 185 (1): 1-19.
- Wang, F. (2006). Quantitative Methods and Applications in GIS. New York: Taylor and Francis group: pp. 43.
- Wang, S., Huang, G.H., Lin, Q.G., Li, Z., Zhang, H., Fana, Y.R. (2014). Comparison of interpolation methods for estimating spatial distribution of precipitation in Ontario, Canada. International Journal of Climatology, 34 (14): 3745-3751 .

- Zawadzki, J., Cieszewski, C.J., Lowe, R.C. (2005). Applying Geostatistics for Investigations of Forest Ecosystems Using Remote Sensing Imagery. Silva Fennica, 39 (4): 599-617.
- Zhang, X., Srinivasan, R. (2009). GIS-Based Spatial Precipitation Estimation: a Comparison of Geostatistical Approaches. Journal of the American Water Resources Association, 45 (4): 894-906.