Academic Journal of Social Sciences
Vol. 4 Issue 4 (October- December 2020) PP 978-989
ISSN 2521-0149
ISSN 2519-7983

SPATIAL ANALYSIS OF THE INCREASED ARSENIC CONCENTRATION IN GROUNDWATER OF SOUTHERN PUNJAB, PAKISTAN

Anum Iftikhar[1]
Shi Hongbo[2]
Maqsood Ahmad[3]
Nasir Ali[4]
Saddam Hussain[5*]

Abstract: *Extensive pumping and less freshwater recharge increase the arsenic content in it, make it damaging to human health. This research examines the groundwater water quality measures of the arid and populated regions of southern Punjab which experiences substantial pumping for irrigation and drinking. In order to observe the 8 well-known water quality variables, set by the Pakistan Environmental Protection Agency (PEPA) and the World Health Organization (WHO), spatial data have been considered from 550 locations from three cities of Punjab (Multan, Muzaffargarh, and Khanewal). Initially Pearsons' correlation coefficients are estimated along with the descriptive measures. Several regression models are estimated to assess the dependence structure between variables and the goodness of fit of these estimated equations are tested on the basis of coefficient of determination. Three classes of measures which were chemically correlated and displayed competition with regression analysis were indicated by factor analysis. Among the water samples, 27 percent showed undesired amounts of arsenic that exceeded the requirements of PEPA and WHO. On the basis of Kriging method, unmeasured locations are predicted using the measured data and results are spatially mapped using contour plots. These results will help the policy makers in taking the decisions regarding the health of the people living in these regions.*

Keywords: Health Risk, Water Quality, Factor Analysis, Southern Punjab

[1] Author is PhD Scholar in Department of Statistics & Finance, Shanxi university of finance and economics , Taiyuan China
[2] Author is Professor in Department of Statistics, Shanxi University of Finance and Economics, Taiyuan China
[3] Author is Associate Professor in Department of Statistics, Faculty of Science University of Okara, Okara, 56300, Pakistan
[4] Author is Assistant Professor in Department of Humanity and Social Sciences at National Textile University Faisalabad.
[5*] Author is PhD Scholar in Department of Statistics, Shanxi university of finance and economics , Taiyuan China, Email: saddamhussain.stat88@gmail.com

Introduction

Fresh drinking water is a basic ingredient of human life while contaminated drinking water cause numerous chronic diseases(Organization 2004). In developing countries including Pakistan, a lot of human population is facing water related issues due to its impurification and shortage. Contaminated water is a source thousands of deaths annually round the globe and thus considered as a dominant health risk (Mubarak et al. 2015). Kahlown et al. (2007) reported that about 80% of all diseases are due to contaminated water and insufficient cleanliness conditions. Since environmental pollution is caused by many factors in various ways but the study of water pollution is relatively more imperative as it is considered the life blood for the whole environment (Ahmad et al. 2017).Water is defined on the basis of its chemical and physical indicators and there is much literature on this area throughout the world focusing on the water quality and analysis techniques (Hussain et al. 2014). The values of these indicators are compared with the WHO permissible limits and contour maps are used as valid graphical tools to show the results of gauged and un-gauged (Ahmad et al. 2017). Heydari et al. (2013)used regression and correlation analysis for determining the association among water quality indicators collected from 21 wells in Kashan city of Iran. Multivariate statistics like correlation matrix, factor analysis and PCA are usually applied in analysis of multivariate water quality data(Van Hulle and Cristina Ciocci 2012) and (Sanchez-Martos et al. 2001).

Achene et al. (2010) evaluated100 drinking water samples in Italy and evaluated the effect of Arsenic (As). If Arsenic is more than the permissible limits it cause diabetes and damages children brain function (Tseng 1989; Tseng et al. 2002; Wasserman et al. 2004; Von Ehrenstein et al., 2006). It also cause liver disease, kidney stone, skin cancer and lungs faults (Morales et al., 2000; Chen et al., 2004). Approximately, 100 million of peoples are at health risks due to drinking arsenic polluted water. Arsenic in drinking water cause skin, lungs, kidney and bladder cancer (Baig et al., 2009). Arsenic concentrations in drinking water have been observed in several areas of Pakistan including Lahore, Kasur, Jamshoro and Manchar lake (Farooqi et al., 2007; Baig et al., 2009; Arain et al., 2009). Consequently, to eliminate the risks associated with Arsenic concentration, integrated methods are required especially in highly contaminated areas (Bakhat et al., 2017).

Spatial statistical techniques are relatively advanced that are used to handle large spatial data for prediction purpose (Saeed et al., 2015). The use of spatial statistical techniques along with multivariate statistical methods are effective analytical tools for handling multivariate spatial data (Ahmad et al., 2016; Ahmad et al., 2017). Saeed et al., (2015) examined the Arsenic concentration level in groundwater of Lahore-Pakistan using Ordinary Kriging, a well-known geo statistical prediction technique and mapped the prediction estimates using contour plots.

The main objective of this research is to assess the water quality of Khanewal, Muzaffargarh and Multan with special focus on Arsenic concentration. To predict the unmeasured sites, measured data is processed by R software using the latest geo statistical kriging methods along with some multivariate statistical techniques. Suitable distance based correlation structure is evaluated and prediction results are shown by contour plots for the visual display of unmeasured sites to help the policymakers in decision making for the best interest of the community health.

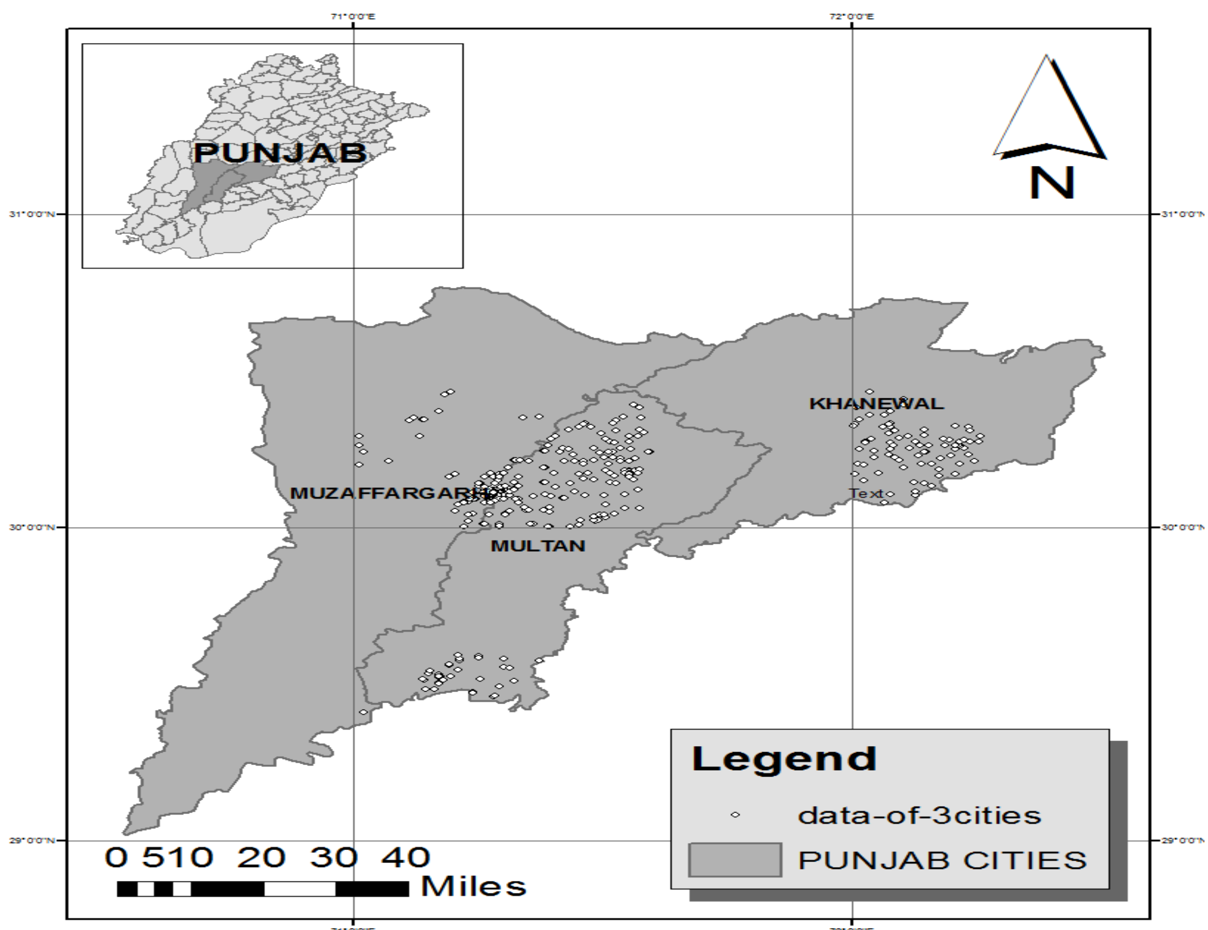
Materials and Methods

Study Area

In this research, we collected data on water samples from Muzaffargarh, Multanand Khanewal, three cities of southern Punjab, Pakistan (Figure 1). The climate of these cities consists of hot summers and cold winter. These cities witness some of the riskiest temperatures in the country where dust storms are very commons. Multiple crops like cotton, wheat, rice, maize, sugarcane, pulses, oilseeds, vegetables and many types of fruits are grown in these cities. For assessing the ground water condition, 550 samples were collected and analyzed to assess the levels of Arsenic, TDS, Magnesium, Bicarbonate, Calcium, Sodium, Sulfate and Potassium.

To carry out the analysis, we have considered two software's: Minitab (summary statistics, regression and correlation analysis and R software for spatial analysis based on kriging (Ribeiro and Diggle 2006)

Fig. 1. Site Map of water samples collected from three cities of Punjab (Muzaffargarh, Multan, Khanewal)..



Water Quality Parameters

The water quality parameters describe that either that water is suitable for human use or not. These parameters have been described with summary statistics in Table 1. Among these parameters, Arsenic (As) takes place in many minerals like sulfur and metals. It is one of the risky elements that exists in the water (Saeed et al. 2015).

High level of TDS has a main role in contaminating the fresh groundwater that disturb the groundwater quality(Ahmad and Chand 2015). Organic matters that include household dirt, mud, leaves and wastages of industries and other inorganic matters such as bicarbonates, carbonates, chlorides, nitrates and sulfates are effective sources that cause excessive level of TDS in water. Excessive TDS in groundwater have ecological impacts both on human health and aquatic. Sulfate is another chemical parameter that naturally exist in drinking water and its excessive amount in drinking water cause lasting diseases such as stomach disorder, diarrhea, laxative effects and food poisoning infants and adults(Ahmad et al. 2016). Potassium with symbol Kisvery reactive with water, producing sufficient heat to burn the hydrogen. Sodium (Na) is exactly solvable in water and is clarified from the earth to surface and groundwater. Bicarbonate is used to increase the pH level in water. Calcium (Ca) is another water quality parameter and its suitable and permissible intake is vital for achieving highest bone mass. Magnesium has also great significance to human welfare and its low intake of either nutrient can harm the human health(Muhammad et al. 2010).

Statistical techniques

Statistical techniques are used in all types of data analysis either it is qualitative or quantitative. Usually descriptive and inferential statistical techniques are applied for univariate data analysis. Multivariate statistical methods are used to handle the data having more than two variables. In environmental researches, usually data sets are of multivariate nature with high dimensions thus for such data sets classification, modeling and dimension reduction is required (Ahmad et al. 2017); (Muhammad et al. 2010).To assess the dependence between the water quality parameters, we use an empirical regression equation described as:

$$Y_i = \beta_0 + \beta_1 X_i \quad (1)$$

Where, β_0 is the intercept that shows the value of response when explanatory variable is zero whileβ_1 is the slope of the estimated regression equation. Moreover, Y is dependent and X is explanatory variable. The normal equations can be deduced from equations (1) as:

$$\sum_{i=1}^{n} Y_i = n\beta_0 + \beta_1 \sum_{i=1}^{n} X_i \quad (2)$$

$$\sum_{i=1}^{n} X_i Y_i = \beta_0 \sum_{i=1}^{n} X_i + \beta_1 \sum_{i=1}^{n} X_i^2 \quad (3)$$

Where β_0 and β_1 are theregression parameters which can be estimated by solving the two normal equations (2)- (3).

Geostatistics and Kriging model

Geostatistics is a branch of Spatial Statistics that deal the geographical referenced spatial and spatio-temporal data for prediction and interpolation. Kriging is just like the regression model with a difference that it deals the spatial data for predicting the unmeasured locations on the basis of measured locations. It is a family of estimators applied to predict the spatial or spatio-temporal data. Simple Kriging, ordinary kriging and cokriging are usually used in GIS for handling such data. In kriging methods, response variable is usually referred as regionalized variable. Let

$Y(s_i): Y(s_1), Y(s_2), \dots, Y(s_n)$ are n spatial realizations of the regionalized variable than the unknown value $Y(s_0)$ is predicted on the basis of ordinary kriging equation described as:

$$\hat{Y}(s_0) = \sum_{i=1}^{n} \lambda_i Y(s_i) \quad (4)$$

Where λ_i are the weights related to the distances among measured observations and sum of theses weights always equal to unity such as:

$$\sum_{i=1}^{n} \lambda_i = 1 , \quad (5)$$

which is a basic prerequisite for an unbiased prediction of unmeasured locations.

Results and discussion

In this section, we have described the results deduced from water quality indicators using different statistical methods.

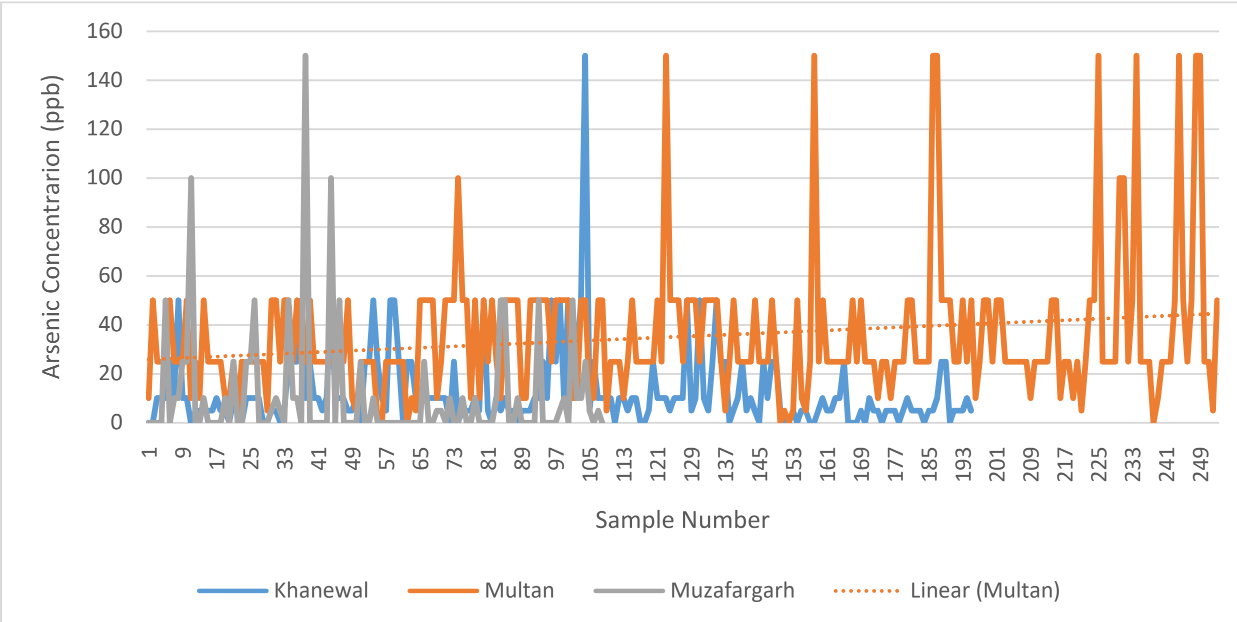
Summary statistics

Descriptive statistical measures are usually calculated to view the center, dispersion and shape of the variables. In Table 1, minimum, maximum, average and SD of water quality parameters along with permissible limits of WHO have been calculated. The average amount of every water quality parameter has been equated with the standard permissible limit WHO. It has been observed that Arsenic (ppb) concentration for the underlying data ranges from 0 to150 with an average of 22.70 ppb which was alarming.TDS concentration for the underlying data ranges from133 to 3424mg/l, Sulfate concentration ranges from 13 to 2133 mg/l, Calcium ranges from 12 to 464 mg/l, Sodium ranges from 7 to 930 mg/l, Potassium ranges from 1 to 37 mg/l, Magnesium ranges from 0 to 145 mg/l and Bicarbonate ranges from 73 to 1244 mg/l. Most of the water quality parameter crossed the permissible limit of the water quality parameter.

Table 1. Summary Statistics (Minimum, Maximum, Average and SD) of water quality parameters along with permissible limits of WHO.

Water Quality Parameters	Min.	Mean	Max.	S.D	WHO limits
Arsenic (ppb)	0.00	22.70	150	26.15	≤12
TDS (Mg/l)	133	722	3424	521.3	≤1000
Sulfate (Mg/l)	13	197.30	2133	206.99	≤500
Calcium (Mg/l)	12	69.17	464	40.88	-
Sodium (Mg/l)	7	144.50	930	138.86	≤200
Potassium(Mg/l)	1	6.155	37	4.316	≤12
Magnesium (Mg/l)	0	28.569	145	19.602	≤150
Bicarbonate(Mg/l)	73	336.66	1244	141.65	-

Fig 2: Arsenic concentration (ppb) at Khanewal (195 samples), Multan (253 samples and Muzafargarh (108 samples)



In Fig 2, Arsenic concentration has been compared city wise. It has been noted that Multan has comparatively high in Arsenic concentration (ppb) in drinking water while Khanewal is at lower level.

Pearsons' correlation coefficients

To evaluate the pair wise association among the water quality parameters, Pearson's correlations coefficients have been computed. Fig 3represented that that Arsenic has negative association with all water quality parameters except Calcium (r=0.035). TDS have high pair wise correlation with all water quality parameters expect Calcium and Potassium(r=0.38, 0.43 respectively). Sulfate represented maximum association with sodium (r=0.85), while magnesium showed positive correlation with all water quality parameters expect Arsenic (r=-0.172) having maximum correlation with TDS (r=0.64).

Estimation of Regression Equations

Correlation coefficients only show the pair wise association while regression equation depicts the dependence structure of the variables. Thus, regression equations along with coefficient of determination (R^2), have been estimated. R^2assists in evaluating the goodness of fit of the regression models(Ahmad et al. 2017). Among all equations shown in Table 2, it has been observed that the effect of sulfate on TDS is the best fit with R^2=98.4%.

 ISSN 2521-0149
ISSN 2519-7983

Fig 3: A graphical display of Pearsons' correlation coefficient of eightwater quality parameter

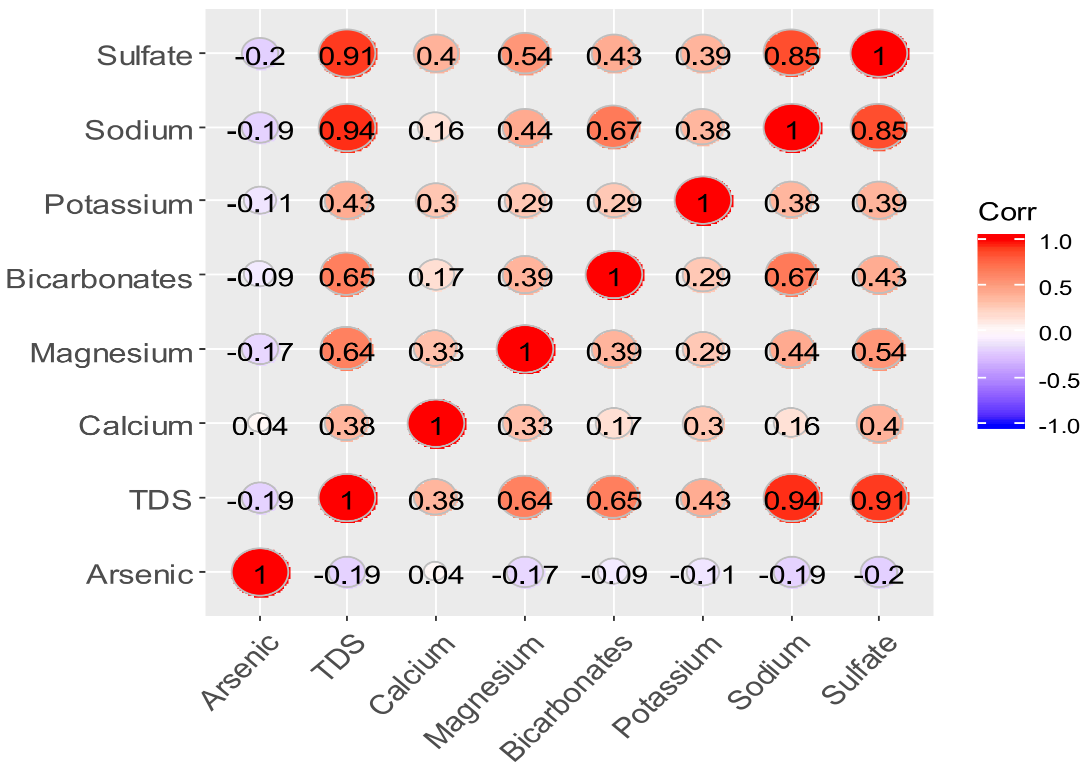


Table 2: Best Regression models for the observed water quality indicators

Dependent Variable	Independent Variable	$\beta_0 + \beta_1$	R^2
TDS	Calcium	-42 + 12.0	87.7%
TDS	Magnesium	28 + 30.4	85.5%
TDS	Sulfate	277 + 1.91	98.4%
Magnesium	Sulfate	24.9 + 0.0830	83.9%
Magnesium	Sodium	35.4 + 0.193	56.7%
Potassium	Sodium	3.18 + 0.00953	70.4%
Potassium	TDS	2.78 + 0.00173	67.8%

It has also been noted that the dependence of TDS over Calcium with R^2=87.7%, Magnesium over Sulfate R2=83.9%, Calcium over Sulfate R^2=86.6% and Potassium over Sodium R^2=87.6% were also high.

3.4 Multivariate Factor Analysis

Since, multivariate statistical analysis is a consistent method for analyzing the environmental and natural sources (Panda et al. 2006).Factor analysis, a well-known multivariate technique has been used to assess the linearly correlated variables(Ahmad et al. 2017). Scree plot suggest the three uncorrelated factors to carry out the analysis rather than dealing all water quality parameters. In Table 3, three factors constructed using varimax rotation method are shown. Following relations can be derived between these three factors and original variables:

Factor1=0.972[SO_4]+0.960[TDS]+0.916[Na] +0.894[Ca]+0.866[Mg]+0.8[K].

Factor2=0.722[HCO_3]

Factor3=0.847[As]

Table 3: Organized rotated factor loadings and communalities

Table 3: Significant loadings

Variable	**Factor_1**	**Factor_2**	**Factor_3**	**Communality**
Sulfate (SO_4)	0.972*	0.157	-0.104	0.981
TDS	0.960*	0.215	-0.156	0.995
Sodium (Na)	0.916*	0.012	-0.255	0.929
Calcium (Ca)	0.894*	0.358	-0.021	0.936
Magnesium (Mg)	0.866*	0.425	-0.039	0.934
Potassium (K)	0.800*	0.061	-0.311	0.770
Bicarbonate	0.334	0.722*	-0.519	0.934
Arsenic	-0.158	-0.135	0.847*	0.785
% Explained Variance	0.481	0.146	0.090	0.898

Table 3 defines the results of factor analysis with a total variance of 89.8% for 3 factors. First factor added 48.1% to the total variance with maximum loading on SO_4, TDS, Na, Ca, Mg and K. Second factor added 14.6% to the total variance with maximum loading on HCO_3 while third factor contributed 9% to the total variance with maximum loading on Arsenic. Thus TDS, Magnesium Calcium, Sulfate, Potassium and Sodium falls in first factor therefore these water quality parameters are linearly associated with each other. Arsenic did not represent strong collinearity with any water quality parameter thus it falls distinctly in a third factor.

3.5 Spatial mapping

Initially, we estimated the three parameters of the variogram (sill, range and nugget)for all water quality parameters. Exponential and spherical covariance functions best fitted the variogram model for all parameters. These are estimated using OLS and WLS estimation method where WLS showed better results as compared to OLS. We use the geo R package (Ribeiro and Diggle 2006) of R statistical software (R Core Team 2014) to spatially model and predict these water parameters. We plug-in the estimated parameters values in kriging equations and compared two geostatistical prediction techniques (ordinary kriging and simple kriging) through mean square error (MSE) and coefficient of determination (R^2). As table 4 shows that the MSE and R^2 for ordinary kriging are showing its superiority. Thus we selected the ordinary kriging and estimated its parameters for each water quality indicator. Arsenic was found to be alarming at the east latitude 29.5^o-30^o, north longitude 71^o-71.5^{o}and similarly at the east latitude 28^o-28.5^o, north longitude 71^o -72^o.

 ISSN 2521-0149
ISSN 2519-7983

Table 4: Comparison of ordinary kriging with simple kriging through mean square error (MSE) and coefficient of determination (R^2).

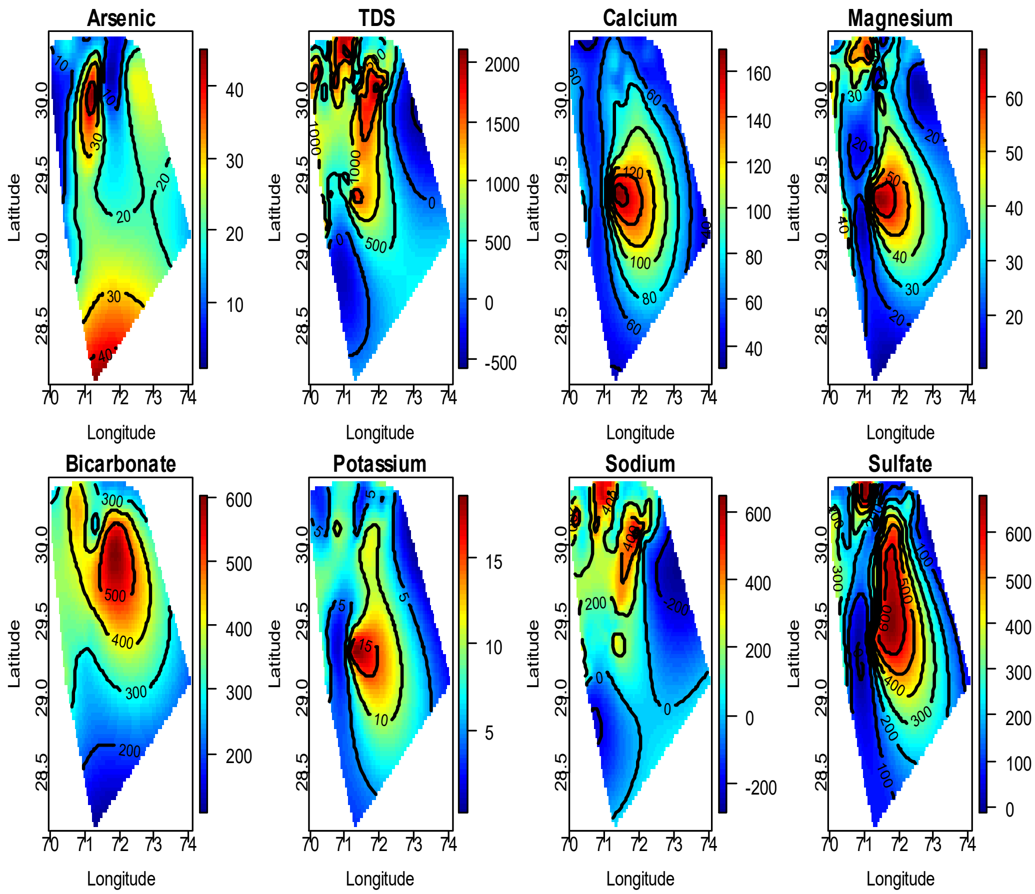
Parameter	Unit	Ordinary Kriging		Simple Kriging	
		MSE	R^2	MSE	R^2
Arsenic	ppb	19.97	0.64	26.33	0.34
TDS	Mg/l	1921.3	0.73	2249	0.44
Calcium	Mg/l	12.881	0.69	1482	0.63
Sulfate	Mg/l	1206.5	0.81	1304.9	0.59
Potassium	Mg/l	88.71	0.64	98.5	0.47
Sodium	Mg/l	2812	0.74	3221	0.60
Bicarbonate	Mg/l	1416	0.65	1843	0.58
Magnesium	Mg/l	98.12	0.71	149.7	0.49

These maps showed clear picture of their concentration level at the study area. The red shaded areas showed the most alarming zones where all parameters extensively crossed the WHO permissible limits (Table 1).

Conclusions

In this paper, we studied eight water quality indicators (Arsenic, TDS, Calcium, Magnesium, Bicarbonate, Potassium, Sodium and Sulfate) in the groundwater of Southern Punjab. It was found that majority of the water samples within the sampling region were beyond the permissible limits set by WHO. Multivariate correlation matrix suggested pair wise association among the water quality indicators while factor analysis grouped the most correlated and seemingly related parameters which were chemically behaving alike. Thus factor analysis proved to be useful in assessing the highly significant water quality parameters. Using the geostatistical ordinary kriging technique, we predicted the unmeasured locations and mapped the resulted prediction values of these variables using contour plot. Arsenic was found to be alarming at the east latitude 29.5^o-30^o, north longitude 71^o -$71.5^o$and similarly at the east latitude 28^o-28.5^o, north longitude 71^o -72^o.

Fig. 4: Contour plots of Arsenic, TDS, Calcium, Magnesium, Bicarbonate, Pottasium, Sodium and Sulfate.



On the basis of these results, it is recommended that the residents of these regions must be aware of that water especially for drinking purpose and it should not be used without desalination at the alarming locations shown in prediction maps for preserving good health.

Acknowledgement

Authors are thankful to Ministry of Science and Technology Pakistan council of research in water resources (PCRWR) for providing data on water quality parameters.

References

Achene L, Ferretti E, Lucentini L, et al (2010) Arsenic content in drinking-water supplies of an important volcanic aquifer in central Italy. Toxicol Environ Chem 92:509–520

Ahmad M, Chand S (2015) Spatial distribution of TDS in drinking water of Tehsil Jampur using Ordinary and Bayesian Kriging. Pakistan J Stat Oper Res 11:377–386. https://doi.org/10.18187/pjsor.v11i3.894

Ahmad M, Chand S, Rafique HM (2017) Geostatistical cokriging and multivariate statistical methods to evaluate groundwater salinization in Faisalabad, Pakistan. Desalin Water Treat 84:. https://doi.org/10.5004/dwt.2017.20911

Ahmad M, Chand S, Rafique HM (2016) Predicting the spatial distribution of sulfate concentration in groundwater of Jampur-Pakistan using geostatistical methods. Desalin Water Treat 57:28195–28204. https://doi.org/10.1080/19443994.2016.1182076

Arain MB, Kazi TG, Baig JA, et al (2009) Determination of arsenic levels in lake water, sediment, and foodstuff from selected area of Sindh, Pakistan: Estimation of daily dietary intake. Food Chem Toxicol 47:242–248. https://doi.org/10.1016/J.FCT.2008.11.009

Baig JA, Kazi TG, Arain MB, et al (2009) Evaluation of arsenic and other physico-chemical parameters of surface and ground water of Jamshoro, Pakistan. J Hazard Mater 166:662–669. https://doi.org/10.1016/J.JHAZMAT.2008.11.069

Bakhat HF, Zia Z, Fahad S, et al (2017) Arsenic uptake, accumulation and toxicity in rice plants: Possible remedies for its detoxification: A review. Environ Sci Pollut Res 24:9142–9158. https://doi.org/10.1007/s11356-017-8462-2

Chen C-L, Hsu L-I, Chiou H-Y, et al (2004) Ingested Arsenic, Cigarette Smoking, and Lung Cancer Risk. JAMA 292:2984. https://doi.org/10.1001/jama.292.24.2984

Farooqi A, Masuda H, Firdous N (2007) Toxic fluoride and arsenic contaminated groundwater in the Lahore and Kasur districts, Punjab, Pakistan and possible contaminant sources. Environ Pollut Elsevier 145:839–849

Gupta A, Chauhan VS, Sankararamakrishnan N (2009) Preparation and evaluation of iron–chitosan composites for removal of As(III) and As(V) from arsenic contaminated real life groundwater. Water Res 43:3862–3870. https://doi.org/10.1016/J.WATRES.2009.05.040

Halim MA, Majumder RK, Nessa SA, et al (2009) Groundwater contamination with arsenic in Sherajdikhan, Bangladesh: geochemical and hydrological implications. Environ Geol 58:73–84. https://doi.org/10.1007/s00254-008-1493-8

Heydari MM, Abbasi A, Rohani SM, Hosseini SMA (2013) Correlation study and regression analysis of drinking water quality in Kashan City, Iran. Walailak J Sci Technol 10:315–324

Hussain I, Shakeel M, Faisal M, et al (2014) Distribution of Total Dissolved Solids in Drinking Water by Means of Bayesian Kriging and Gaussian Spatial Predictive Process. Water Qual Expo Heal. https://doi.org/DOI 10.1007/s12403-014-0123-9.

Kahlown MA, Tahir MA, Rasheed H (2007) National Water Quality Monitoring Programme,Fifth Technical Report. In: Pakistan Counc. Res. Water Resour. vol. 5, Islam. http://www.pcrwr.gov.pk/Index.html

Morales KH, Ryan L, Kuo TL, et al (2000) Risk of internal cancers from arsenic in drinking water. Environ Health Perspect 108:655–61. https://doi.org/10.1289/ehp.00108655

Mubarak N, Hussain I, Faisal M, et al (2015) Spatial Distribution of Sulfate Concentration in Groundwater of South-Punjab, Pakistan. Water Qual Expo Heal

7:503–513

Muhammad S, Shah M, Khan S (2010) Arsenic health risk assessment in drinking water and source apportionment using multivariate statistical techniques in Kohistan region, northern Pakistan. Food Chem Toxicol 48:2855–2864

Organization WH (2004) Guidelines for drinking-water quality: recommendations. World Health Organization

R Core Team (2014) R, A Language and Environment for Statistical Computing. In: R Found. Stat. Comput. Vienna, Austria,. http://www.r-project.org/

Ribeiro PJ, Diggle PJ (2006) geoR: A package for geostatistical analysis. R-NEWS 1:1–28

Saeed S, Javed Z, Chand S, et al (2015) Spatial Distribution of Arsenic Concentration in Drinking. Sci Int Lahore 27:949–954

Sanchez-Martos F, Jimenez-Espinosa R, Pulido-Bosch A (2001) Mapping groundwater quality variables using PCA and geostatistics: a case study of Bajo Andarax, southeastern Spain. Hydrol Sci J 46:227–242

Tseng C-H, Tseng C-P, Chiou H-Y, et al (2002) Epidemiologic evidence of diabetogenic effect of arsenic. Toxicol Lett 133:69–76

Tseng W-P (1989) Blackfoot disease in Taiwan: a 30-year follow-up study. Angiology 40:547–558

Van Hulle SWH, Cristina Ciocci M (2012) Statistical evaluation and comparison of the chemical quality of bottled water and flemish tap water. Desalin Water Treat 40:183–193. https://doi.org/10.1080/19443994.2012.671166

Von Ehrenstein OS, Guha Mazumder DN, Hira-Smith M, et al (2006) Pregnancy outcomes, infant mortality, and arsenic in drinking water in West Bengal, India. Am J Epidemiol 163:662–669

Wasserman GA, Liu X, Parvez F, et al (2004) Water arsenic exposure and children's intellectual function in Araihazar, Bangladesh. Environ Health Perspect 112:1329

Xie X, Wang Y, Duan M, Xie Z (2009) Geochemical and environmental magnetic characteristics of high arsenic aquifer sediments from Datong Basin, northern China. Environ Geol 58:45–52. https://doi.org/10.1007/s00254-008-1489-4