J.Natn.Sci.Foundation Sri Lanka 2017 45 (2): 179-191
DOI: http://dx.doi.org/10.4038/jnsfsr.v45i2.8035

RESEARCH ARTICLE

Geo-statistical models for determining spatial variation and spatial dependency of soil arsenic in Bangladesh

M.S. Kabir [1], M.A. Salam [2], D.N.R. Paul [3], M.I. Hossain [1], N.M.F. Rahman [1] and M.A. Latif [4*]
[1] *Agricultural Statistics Division, Bangladesh Rice Research Institute, Gazipur, Bangladesh.*
[2] *Department of Statistics, Jahangirnagar University, Savar, Dhaka, Bangladesh.*
[3] *Department of Business Administration, Uttara University, Uttara, Bangladesh.*
[4] *Plant Pathology Division, Bangladesh Rice Research Institute, Gazipur, Bangladesh.*

Revised: 24 May 2016; Accepted: 21 July 2016

Abstract: Arsenic (As) toxicity has become a major problem since a few decades in different parts of Bangladesh mainly due to the massive groundwater extraction for both drinking and irrigation purposes. This study was undertaken to investigate the spatial variation and spatial dependency of soil As at four locations in Bangladesh. Composite soil samples were collected at 0 – 45 cm depth in a grid area of 1 m^2 to conduct the study of a shallow tube well command area (micro level). Based on composite soil sampling in four different study locations, laboratory analysis and statistical modelling, the spatial variation and dependency of soil As was estimated. Four study area maps were digitised using ArcGIS. Arsenic concentration varied widely within the study areas. The extent and propensity of As concentration was higher in areas with a high concentration of As in groundwater and soil. About 15.01 ppm of As is loaded into soils through groundwater irrigation every year in the As affected areas. The semi-variogram model for describing spatial dependency of soil As was found to be scale dependent. At study area scale, the appropriate model was identified as spherical. Kriging method appeared to be more suitable to create an interpolated surface for study area scale. Development of As tolerant rice varieties, promoting cropping patterns that require less irrigation water, and alternate wetting and drying (AWD) method can be remedial measures to reduce As contamination in soil as well as in the food chain.

Keywords: Arsenic, geostatistics, interpolation, kriging, semi-variogram, soil.

INTRODUCTION

Soil pollution is posing a significant threat to human health and environmental quality across the globe. Among the range of pollutants, deposition of different heavy metals (HM) has drawn considerable attention over the last few decades (Franssen *et al*., 1997). Arsenic (As) concentrations in soils in some areas, particularly in agricultural lands, have reached threat levels. In order to develop effective management recommendations, spatial patterns of pollutants must be known. In practice however, it is difficult to characterise this spatial pattern accurately. Polluted areas show complex spatial patterns, high coefficient of variation and occurrence of hot spots of locally contaminated soils (Franssen *et al*., 1997). Predictions of polluted areas are often based on geo-statistical methods, which calculate unbiased estimates of heavy metal concentrations at un sampled locations (Atteia *et al*., 1994; Goovaerts, 1997; Meuli *et al*., 1998; Van Meirvenne & Goovaerts, 2001). These methods provide either an estimated mean value of heavy metal concentration or the probability of exceeding a given threshold level (Cattle *et al*., 2002).

Bangladesh is one of the most severely As polluted countries in the world. The scenario of As pollution has become critical, particularly because of the massive groundwater extraction for both household and agricultural purposes. Majority of the wells in the As prone areas are, however, known to be naturally contaminated by As, which occurs in alluvial soils of Bangladesh (Acharyya *et al*., 1999). About 35 million people in Bangladesh are vulnerable to As pollution through drinking polluted water containing >50 $\mu g\ L^{-1}$ As. In addition, cultivation of staple crops such as rice and

* Corresponding author (alatif1965@yahoo.com)

other cereals and vegetables in large areas of Bangladesh mainly depends on As contaminated groundwater as the source of irrigation (Nickson *et al*., 1998; Panaullah *et al*., 2003). Such agricultural practices increase the level of As pollution in arable soil (Alam & Satter, 2000), which in turn may lead to increased As concentrations in cereals, vegetables and other agricultural consumables (Abedin, 2002; Alam & Rahman, 2003; Duxbury *et al*., 2003; Meharg & Rahman, 2003). It is evident from the above that As is causing a high level of poisoning in Bangladesh through drinking polluted water coupled with consuming As contaminated foodstuff. Therefore, it is of utmost importance to improve our understanding of the As problem and find a way of mitigation by exploring the sediment-water-soil-crop-human system.

The primary requisite to develop an effective pollution-mitigation strategy is to understand the spatial distribution pattern of the particular pollutant. However for As, it is often challenging to accurately characterise the affected areas because of its complex spatial pattern of distribution, high degree of variation among different locations and occurrence of hot spots in locally contaminated soils (Franssen *et al*., 1997). Predictions of polluted areas are often based on geo-statistical methods, which calculate unbiased estimates of heavy metal concentrations at unsampled locations (Atteia *et al*., 1994; Goovaerts, 1997; Meuli *et al*., 1998; Van Meirvenne & Goovaerts, 2001). These methods provide either an estimated mean value of the heavy metal concentration or the probability of exceeding a given threshold level of the pollutant (Cattle *et al*., 2002).

Little is known on the spatial variability of As concentration in irrigated and non irrigated soil and its impact on As accumulation in rice grain and straw. To understand the geo-statistical processes, knowledge on the spatial variability of As in soils is essential. Geo-statistical methods can be utilised as powerful tools to explore As pollution, provided that sufficient data on the spatial variability of As in soils is available. Such methods are generally used for characterising large-scale spatial distributions of soil properties in precision agriculture. It has a number of implications: it can potentially delineate the spatial variability of soil As; it will reveal the real situation of As pollution in Bangladesh, which may assist formulating a future policy on the management of As problem. It can also provide with an advanced methodology to facilitate spatial interpolation and quantification of spatial temporal variability in soil variables (McGrath *et al*., 2004; Robinson & Metternicht, 2006). The geo-statistical approach has gained popularity for analysing spatial structure and spatial distribution of other soil heavy metals (Imperato *et al*., 2003; Ahsan *et al*., 2009; Wu *et al*., 2009; Zhang *et al*., 2009). For example, 'kriging' is a precise estimator for spatial data analysis as it is unbiased and is capable of minimising total uncertainty (Isaaks & Srivastava, 1989). Although Bangladesh is one of the most As affected countries in the world, the spatial variation and spatial dependency of soil As concentration are still mostly unknown. Therefore, this paper is aimed at describing the spatial variation of As in Bangladesh soils and also identifying an appropriate semi-variogram model for describing spatial dependency of soil As.

METHODOLOGY

Study site

The study was conducted in four different locations: Faridpur Sadar (Faridpur District), Tala (Faridpur District), Brahmin Baria (Brahmin Baria District) and Sonargaon (Narayanganj District) in Bangladesh (Figure 1).

Soil sampling and analysis

Soil samples were collected from Boro Rice (dry season rice) fields irrigated with As contaminated water and transferred to airtight polyethylene bags. Composite soil samples were collected from 0 – 45 cm depth in a 1 m^2 area in 2004 and brought back to the Soil Science Laboratory of the Bangladesh Rice Research Institute (BRRI). The collected soil samples were immediately sun-dried and then oven-dried at 60 °C for 72 hrs. The dry soil samples were ground and sieved through a 2.0 mm sieve. Each sample point was geo-referenced using GPS coordinates. For Faridpur command area the sample size was 101 (Figure 1); 63 for Tala (Figure 2); 96 for Brahmin Baria (B. Baria) (Figure 3) and 144 for Sonargaon (Figure 4). Distribution of the sampling points covered the field quite evenly for all command areas. The average sampling interval was 12.35 m grind points over an area of 1.54 ha for Faridpur. It was 12.64 m grind points over an area of 1.01 ha, 13.9 m grind points over an area of 2.43 ha and 18.53 m grind points over an area of 4.95 ha, respectively for Tala, B. Baria and Sonargaon.

Smaller sampling intervals were chosen to resolve any short-scale variation that might be present. The soil samples were air-dried and analysed by Tri Acid method and the results of the soil elements (As, Fe, Mn, P, OC, pH and soil texture) were recorded accordingly.

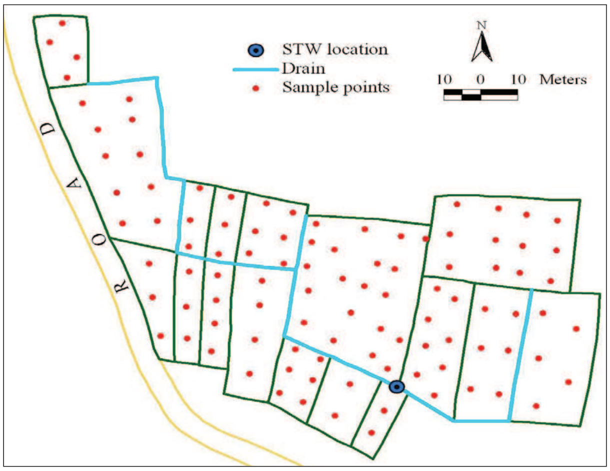

Figure 1: Distribution of the sample points in Faridpur command area, Faridpur

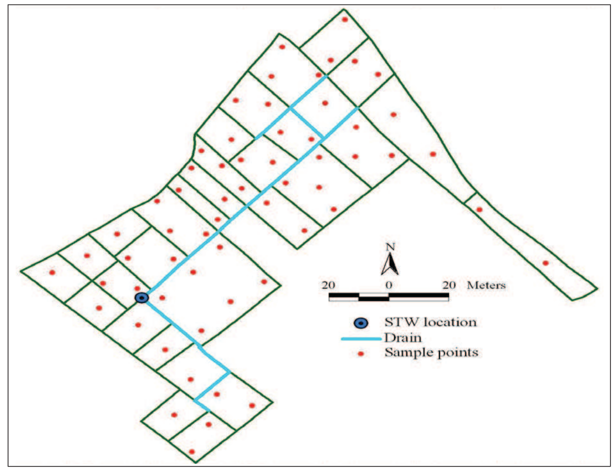

Figure 2: Distribution of the sample points in Tala command area, Satkhira

Figure 3: Distribution of the sample points in Brahmin Baria command area, Brahmin Baria

Figure 4: Distribution of the sample points in Sonargaon command area, Narayanganj

Geo-statistical methods

Descriptive statistical analysis (such as mean, median, mode, minimum, maximum, standard deviation, coefficient of variation and percentages) was used to summarise soil As concentration data. Semi-variograms were calculated to determine the spatial variations in As content. After normalisation of the data, maximum lag distance and lag interval for the semi variance were determined iteratively to best fit the model having the highest R^2 and the lowest residual sum of squares (RSS). Akaike information criterion (AIC) and inverse distance weight (IDW) methods were used for interpolation of the soil As kriging. Finally, the best method of interpolation was selected using the root mean square error (RMSE) and cross validation. A map of soil As distribution was prepared using this interpolation data and geographical information system (GIS).

Semi-variogram models

Normally a variogram model is fitted through the emperical semi-variogram values for the distance classes or lag classes. Variogram properties such as the sill, range and nugget can provide insights on which model will fit best (Cressie 1993; Burrough & McDonnell, 1998). The experimental variogram is fitted by a theoretical model $\gamma(h)$, which can be a spherical, exponential or Gaussian model, and three parameters of the fitted model - the nugget effect, the sill and the range - are determined (Deutsch, 2002). These models can be categorised by the presence or absence of a sill and by the behaviour at the origin, which is either linear or parabolic. Figure 5 shows the experimental variogram.

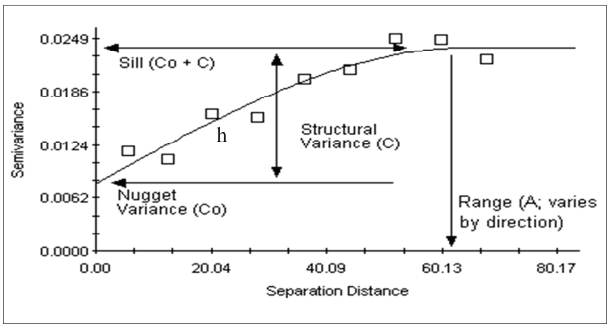

Figure 5: Experimental variogram

Co, which is commonly called the nugget effect, provides the intercept of the semi-variogram with the vertical axis. In an ideal case where there is no measurement error, the nugget value is zero. A, commonly called the range, provides a distance beyond which the variogram value remains essentially constant. Co + C, commonly called the sill, will no longer increase with distance, causing a flat region on the semi-variogram. Sill represents the maximum variance of the variogram and spatially-independent variance. Data locations separated by a distance beyond which semi-variance does not change, i.e. after the model asymptote or sill, are spatially independent of one another. Theoretically the sill is equivalent to sample variance.

h is the distance between the locations, which includes both observed and estimated locations. Proportion of spatial structure or C/(Co+C) provides a measure of the proportion of sample variance (Co+C) that is explained by spatially structured variance C. R^2 or regression coefficient provides an indication of how well the model fits the variogram data. RSS or residual sums of squares provide an exact measure of how well the model fits the variogram data; the lower the reduced sums of squares, the better the model fits. GS+ uses RSS to choose parameters for each of the variogram models by determining the combination of parameter values that minimises RSS for any given model.

Akaike information criterion or AIC is a measure of the relative goodness-of-fit of a statistical model. For fitting, the likelihood is given by

$$AIC = n \ln \left(\frac{RSS}{n} \right) + 2k + C$$

where k is the number of parameters in the model and C is a constant independent of the model used, which can be ignored in model comparisons. Given a set of candidate models for the data, the preferred model is the one with the minimum AIC value.

Thus, the best model for fitting an experimental variogram was selected based on the highest R^2, the lowest RSS and the proportion of spatial structure close to unity. The quality of the semi-variogram fit to the data was indicated using R^2, and an F-test calculated as follows (Wang *et al*., 2001):

$$F = \frac{R^2}{1 - R^2} \text{ x } \frac{n - k}{k - 1}$$

where k is the number of parameters in the regression model, and n is the number of samples with (k-1), (n-k) degrees of freedom.

Spherical model

The most commonly used variogram model is the spherical model. When the nugget variance is important but not large and there is a clear range and sill, a curve known as the spherical model often fits the variogram well.

$$\gamma(h) = C_0 + C_1 \left\{ \frac{3}{2}\left(\frac{h}{a}\right) - \frac{1}{2}\left(\frac{h}{a}\right)^3 \right\} \quad \text{for } h \in 0\text{, and}$$

$$= C_0 + C_1 \quad \text{for } h > a$$

where a is the range, h is the lag or distance and C_0+C_1 is the sill.

Exponential model

If there is a clear nugget and sill but only a gradual approach to the range, the exponential model is often preferred.

$$\gamma(h) = C_0 + C_1 \left\{ 1 - e^{-\left(\frac{h}{a}\right)} \right\} \quad \text{for } h > 0$$

Gaussian model

The Gaussian model is a transition model that is often used to model extremely continuous phenomena. If the variation is very smooth and the nugget variance is very small compared to the spatial random variation, then the variogram can often be best fitted by a curve having an inflection such as the Gaussian model:

$$\gamma(h) = C_0 + C_1\left\{1 - e^{-\left(\frac{h}{a}\right)^2}\right\} \text{ for } h > 0$$

Linear model

Non-transitive variograms have no sill within the sampled area and may be represented by the linear model:

$$\gamma(h) = C_0 + bh$$

Spatial prediction methods

Kriging

Kriging is the estimation procedure used in geo-statistics using known values and a semi-variogram to determine unknown values. The experimental variogram measures the average degree of dissimilarity between unsampled values and a nearby data value (Deutsch & Journel, 1998), and thus can depict autocorrelation at various distances. The value of the experimental variogram for a separation distance of h (referred to as the lag) is half the average squared difference between the value at $z(x_i)$ and the value at $z(x_i + h)$ (Lark, 2000; Robinson & Metternicht, 2006):

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} [z(x_i) - z(x_i + h)]^2$$

where $2n(h)$ is the number of data pairs within a given class of distance and direction. If the values at $z(x_i)$ and $z(x_i + h)$ are auto correlated, the result of the equation will be small relative to an uncorrelated pair of points. From analysing the experimental variogram, a suitable model (e.g. spherical, exponential, Gaussian and linear to sill models) is then fitted. Cross validation was done to validate the accuracy of the fitted models for prediction. The models were compared using R^2 and RSS and the best fit was selected for the kriging phase. The model with the highest R^2 and the lowest RSS value was selected. The spatial dependency of As in soil was evaluated by weighted least squares and the parameters (e.g. range, nugget and sill) were then used in the kriging procedure.

Inverse distance weight (IDW)

In interpolation with the IDW method, a weight is attributed to the point to be measured. The amount of this weight depends on the distance of the point to another unknown point. These weights are controlled on the bases of power of ten. With increase of the power of ten, the effect of the points further diminishes. Lesser power distributes the weights more uniformly between neighbouring points. In this method, the distance between the points counts, so the points of equal distance have equal weights (Burrough & McDonnell, 1998). In this method, the weight factor is calculated with the use of the following formula:

$$\lambda_i = \frac{D_i^{-\alpha}}{\sum_{i=1}^{n} D_i^{-\alpha}}$$

λ_i is the weight of point, D_i is the distance between point i and the unknown point, and α is the power ten of weight.

Comparison between the different methods

Finally, the RMSE was used to evaluate model performances in cross-validation mode. The smallest RMSE indicates the most accurate predictions. The RMSE was derived from the following equation

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

where $z(x_i)$ is the observed value at point x_i, $z^*(x_i)]^2$ is the predicted value at point x_i, and N is the number of samples.

Data analysis

Data were analysed using different statistical software packages. The descriptive statistical parameters were calculated using SPSS (version 16). Maps were produced using ArcGIS (version 9.2). The geo-statistical analyses and the probability calculations were performed using GS+ (for windows version 5).

RESULTS AND DISCUSSION

The descriptive statistics of soil As at the four study sites are shown in Table 1. Results revealed that the mean concentrations of soil As were 9.42, 13.15, 20.40 and 8.82 ppm within the range of 3.21 to 24.43, 6.99 to 18.99, 14.26 to 25.19 and 4.31 to 15.79 ppm, respectively in B. Baria, Faridpur, Tala and Sonargaon. For agricultural soil the acceptable As limit is 20.0 ppm as recommended by the European Community (Smith, 1998; Abedin, 2002; Rahman *et al*., 2007). According to WHO recommendation the permissible As limit in rice grain is 1.0 ppm (Smith, 1998; Abedin, 2002; Rahman *et al*., 2007). Therefore, soil As concentration has exceeded the critical level (>20 ppm) only in Tala. The mean concentration of As in rice grain from the command areas was found to vary from 0.23 to 0.42 ppm (ranging from 0.08 to 1.07 ppm) which did not exceed the permissible limit (1.0 ppm) according to WHO recommendation (Rahman *et al*., 2007; Abedin *et al*., 2002; Smith, 1998). The remarkable shielding of rice grains from contaminated soil As is consistent with several previous studies (Abedin, 2002; Rahman *et al*., 2007). The median value of soil As in Tala (20.46 ppm) exceeded the critical value (20 ppm), however, all the other descriptive values were lower than the critical value. The higher mean value than the median value indicated the influence of a few samples with high As levels. Comparing the standard deviations with means, the distributions were not found to be symmetrical (Atteia *et al.*, 1994). The standard deviation of soil As at Sonargaon was close to 1/3 of the mean. All the other values of standard deviation were not equal to 1/3 of the mean indicating the non-normality of the distribution of As data. According to Carvalho *et al*. (2002) values of skewness and kurtosis were close to 0 and 3, respectively and are indicative of a normal distribution. Skewness of soil As in Faridpur (0.09 ppm) was close to 0 and other data on As was strongly skewed. In case of kurtosis, none of the coefficient of kurtosis values was close to 3. Non-normality of the distribution of As data in terms of skewness and kurtosis is an indication of the need for transformation of data for further analyses. The required transformation was determined using GS+ software based on the criteria of having skewness close to 0 and kurtosis close to 3. In B. Baria and Sonargaon soil As data required log normal (LN) transformation; however, in Faridpur and Tala no transformation was required. The coefficient of variation (CV) was classified as: ≤10 % low; 10 – 20 % medium; 20 – 30 % high and ≥30 % very high (Gomes & Garcia, 2002). A very high variability was observed in soil As in B. Baria ranging from 32.89 to 69.36 % except for that in Faridpur, Tala and Sonargaon. Soil management may be contributing to the great variability of data in the study area because distribution of irrigation water was not homogeneous due to the lack of plot levelling. This fact might have caused a modification of the natural spatial variability of soil elements.

Concentration of As in soil samples in four command areas

The concentration of As in soil samples obtained from the four command areas is presented in Table 2. In the case of soil sampled from B. Baria, 65 % contained less than 10 ppm As, 29 % contained between 10 – 20 ppm and 6 % contained more than 20 ppm. In Faridpur, 8 % of the samples contained less than 10 ppm As and 92 % contained between 10 – 20 ppm. In Tala, the As concentration was less than 10 ppm in 47 % of the samples and was between 10 – 20 ppm in 53 % of the samples. In Sonargaon, 73 % soil samples contained As less than 10 ppm and 27 % contained between 10 – 20 ppm.

Table 1: Descriptive statistics of soil As in Faridpur, Tala, B. Baria and Sonargaon

Parameters	Soil As (ppm)			
	Faridpur	Tala	B. Baria	Sonargaon
Mean	13.15	20.40	9.42	8.82
Median	12.90	20.46	4.85	8.42
Mode	13.26	19.41	3.21	11.31
Min	6.99	14.26	3.21	4.31
Max	18.99	25.19	24.43	15.79
Sd	2.51	2.48	5.38	2.63
Skewness	0.09	- 0.11	0.89	0.78
Kurtosis	- 0.18	- 0.25	- 0.10	- 0.12
CV (%)	19.10	12.13	54.05	29.80
Transform	None	None	LN	LN

Table 2: Percentage distribution of As concentration in soil in Faridpur, Tala, B. Baria and Sonargaon

As concentration	Percentage			
	Faridpur	Tala	B. Baria	Sonargaon
< 10 ppm	8	47	65	73
10 – 20 ppm	92	53	29	27
> 20 ppm	0	0	6	0

Table 3: Soil As and textural parameters of surface soil under different command areas

Location	Sample size (no.)	Soil As (ppm)	OM	Soil textural parameters		
				Sand %	Clay %	Silt %
Faridpur	100	13.15	1.23	24.50	34.42	38.08
Tala	60	20.40	1.07	4.98	35.40	56.62
B. Baria	96	9.42	2.10	20.33	26.43	53.24
Sonargaon	144	8.82	1.99	4.69	85.31	4.00

Relationship between soil As and soil parameters

The relationship of soil As, organic matter (OM) and soil texture is presented in Table 3. It was observed that the soil As content decreased with the increase of soil OM. Besides, soil As content was higher for soils with higher silt and lower with higher clay contents.

Semi-variogram models

Selection of the best semi-variogram model

From Tables 4a to 4d, it is evident that the range of spatial dependency of soil As within the study areas of B. Baria, Faridpur, Tala and Sonargaon was about 250.6, 31.1, 64.1 and 66.8 m, respectively. The lag distance was 207, 80, 90 and 222 m, respectively in B. Baria, Faridpur, Tala and Sonargaon. Such findings indicate that the As content in Sonargaon soil is more dispersed and weakly correlated with increasing distance followed by B. Baria, Tala and Faridpur. The lag interval in B. Baria, Faridpur, Tala and Sonargaon was 20.7, 8.0, 9.0 and 22.2, respectively. The nugget effect of soil As was 0.001, 0.29, 1.1 and 0.06, respectively in B. Baria, Faridpur, Tala and Sonargaon indicating a small analytical error and less variability in B. Baria and Sonargaon. A high analytical error and a high variability of soil As was found in Faridpur and Tala within the lag intervals. The sill value of soil As was 0.573, 5.55, 7.9 and 0.116, respectively in B. Baria, Faridpur, Tala and Sonargaon. The proportion of spatial structure or $C/(C_0+C)$ statistic provides a measure of the proportion of sample variance (C_0+C) that is explained by spatially structured variance C. The proportion of spatial structure to sampling variance was close to unity for soil As at B. Baria (100 %), Faridpur (95 %) and Tala (86 %), indicating less variability in As within the lag intervals, and the semi-variogram model explained most of the sampling variation. For soil As at Sonargaon, only 52 % of the sampling variance within the lag interval could be attributed to spatial structure, which indicated that some factors other than the distance contributed to the variation of As levels within the command areas. Regression coefficient (R^2) provides an indication of how well the variogram model fits the data. The value of R^2 for soil As was 0.98, 0.98, 0.98 and 0.93, respectively in B. Baria, Faridpur, Tala and Sonargaon. All the regression coefficients were highly significant at $F_{0.01, k-1, n-k}$ indicating that in almost all the cases the spherical model fits very well with the variogram data, where k is the number of parameters in the model. Residual sums of squares (RSS) provide an exact measure of how well the variogram model fits the data; the lower the residual sum of squares, the better the model fit. The minimum value of RSS for soil As was 0.0076, 0.388, 0.891 and 0.0002, respectively in B. Baria, Faridpur, Tala and Sonargaon. Akaike information criterion (AIC) values provide a means for model selection. The lowest values of AIC for soil As were -900.62, -555.75, -262.29 and -1936.13, respectively in B. Baria, Faridpur, Tala and Sonargaon. For soil As, the neighbourhoods were 10, 10, 10 and 6, respectively in B. Baria, Faridpur, Tala and Sonargaon. Finally the results showed that based on the criteria used to judge the competence of a model, the spherical model, in general, fitted well in almost all the cases. In many previous studies, the spherical model was found to be the most used for computing semi-variograms of soil (Cambardella *et al*., 1994; Grego & Vieira, 2005; Siqueira *et al*., 2008). The semi-variograms of soil As for all the command areas were therefore computed using the spherical model and are presented in Figure 6.

Table 4a: Comparison of variogram models for soil As in B. Baria

Parameters	Spherical	Exponential	Linear	Linear to sill	Gaussian
Range (m)	250.6	1354.8	196.4	413.7	189.5
Lag distance	207	207	207	207	207
Lag interval	20.7	20.7	20.7	20.7	20.7
Nugget (C_0)	0.001	0.001	0.018	0.020	0.043
Sill (C_0 + C)	0.573	1.581	0.574	1.186	0.546
Proportion of structural variance to total sampling variance: Q = $C/(C+C_0)$	1.00	1.00	0.97	0.98	0.92
C_0 / (C+C_0) *100	0.0	0.0	3.0	2.0	8.0
R^2	0.99	0.98	0.95	0.95	0.96
RSS	0.0076	0.0076	1.150	0.013	0.025
AIC	- 900.62	- 900.62	- 418.76	- 849.09	- 786.31
Neighbourhood	10	10	10	10	10

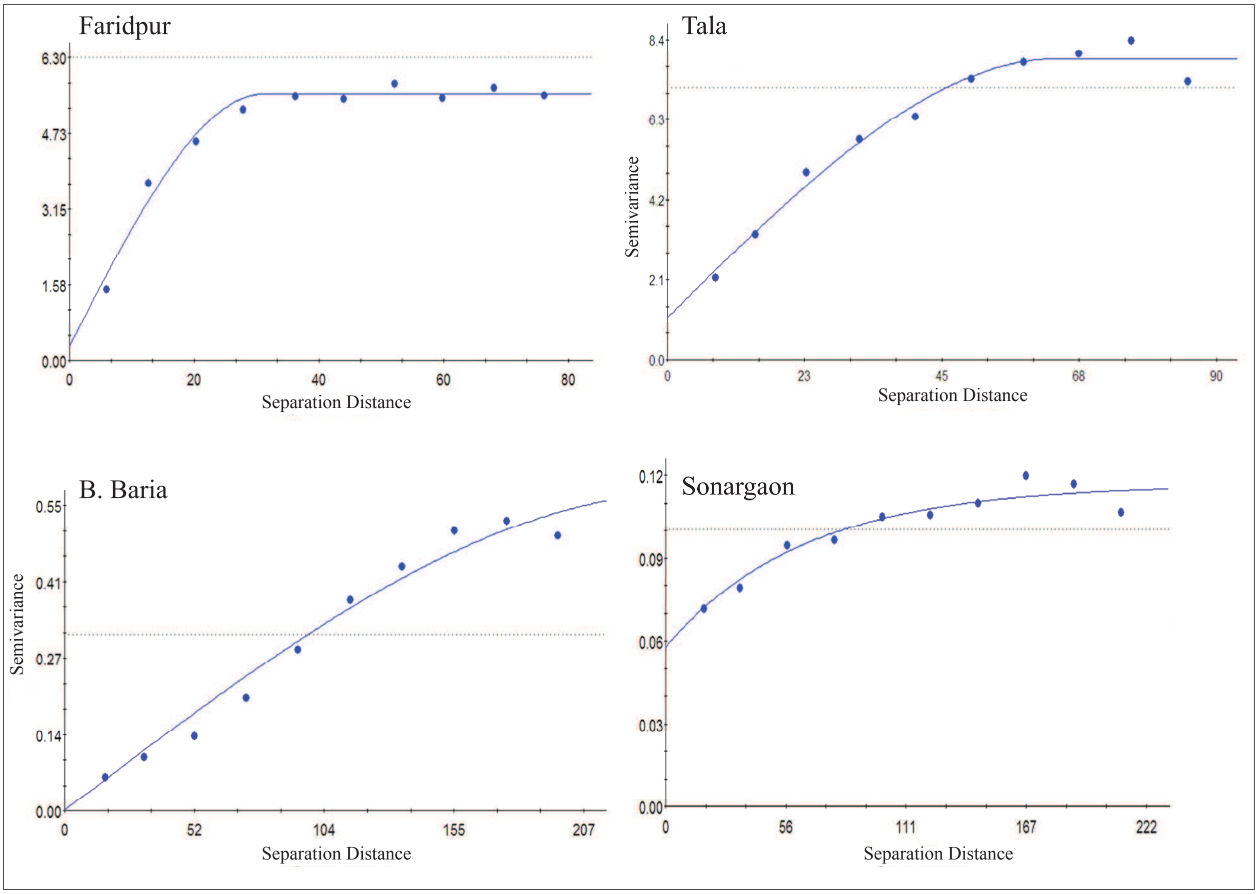

Figure 6: Semi-variogram models for soil As in the command areas

Spatial dependency of the semi-variogram models

The ratio of nugget variance to sill is presented in Tables 4a to 4d, showing that for soil As this ratio was < 25 % in B. Baria, Faridpur and Tala, which exhibited strong spatial dependence among sample points. However, in Sonargaon, it showed moderate spatial dependence (48.7 %). Therefore, there was no spatial dependency beyond the above ranges of the command areas. However, within a study area, the nature and extent of dependency were not uniform in all the sampling sites. In B. Baria, the extent of spatial dependency spread up to long distances in all cases, whereas in Tala it was true for most of the cases.

Selection of interpolation methods

For the determination of the most suitable interpolation method, root mean square error (RMSE) technique was used. The results of RMSE are presented in Table 5. Irrespective of the location and variates the RMSE value of kriging was less than that of inverse distance weight (IDW). The results indicated that the geo-statistic method (kriging) was relatively more precise than the IDW method. This result was in line with the previous findings of Safari (2002), Nazari *et al.* (2006), Ahmed (2002) and Barca and Passarella (2007).

Spatial variability of soil As in command areas

The interpolated maps of soil As are presented in Figure 7. As observed 40 % of the area of B. Baria exhibited an As level >10 ppm, while more than 5 % of the area fell under As level >20 ppm. On the other hand, in Faridpur, more than 90 % of the area fell under As level >10 ppm. In Tala, it was 100 % and in Sonargaon it was more than 30 %. Among the command areas the As concentration in the soil of B. Baria showed a decreasing trend with the increase of distance from the tube well. Such a trend, however, was not consistent in the other study areas, where As was distributed in patches spread over the study areas. However, in Tala the As concentration was high along the irrigation drains

Table 4b: Comparison of variogram models for soil As in Faridpur

Parameters	Spherical	Exponential	Linear	Linear to sill	Gaussian
Range (m)	31.1	39.6	76.1	144.7	25.8
Lag distance	80	80	80	80	80
Lag interval	8	8	8	8	8
Nugget (C_0)	0.290	0.010	3.125	3.140	1.060
Sill (C_0 + C)	5.548	5.732	6.379	9.287	5.541
Proportion of structural variance to total sampling variance: Q = $C/(C+C_0)$	0.95	1.00	0.51	0.66	0.81
C_0 / $(C+C_0)$ *100	5.0	0.0	49.0	34.0	19.0
R^2	0.98	0.97	0.58	0.58	0.98
RSS	0.388	0.531	36.7	6.74	0.388
AIC	- 555.75	- 524.06	- 96.25	- 267.41	- 555.75
Neighbourhood	10	10	10	10	10

Table 4c: Comparison of variogram models for soil As in Tala

Parameters	Spherical	Exponential	Linear	Linear to sill	Gaussian
Range (m)	64.1	80.4	85.3	125.0	53.35
Lag distance	90.0	90.0	90.0	90.0	90.0
Lag interval	9.0	9.0	9.0	9.0	9.0
Nugget (C_0)	1.10	0.010	2.894	2.920	2.060
Sill (C_0 + C)	4.897	8.487	9.024	11.849	4.881
Proportion of structural variance to total sampling variance: Q = $C/(C+C_0)$	0.86	1.00	0.68	0.75	0.74
C_0 / $(C+C_0)$ *100	14.0	0.0	32.0	25.0	26.0
R^2	0.98	0.97	0.81	0.81	0.97
RSS	0.891	1.16	114.0	4.43	1.14
AIC	- 262.29	- 245.67	43.36	- 161.25	- 246.76
Neighbourhood	10	10	10	10	10

Table 4d: Comparison of variogram models for soil As in Sonargaon

Parameters	Spherical	Exponential	Linear	Linear to sill	Gaussian
Range (m)	498.0	192.6	210.1	434.4	484.1
Lag distance	222	222	222	222	222
Lag interval	22.2	22.2	22.2	22.2	22.2
Nugget (C_0)	0.075	0.057	0.0773	0.0774	0.085
Sill (C_0 + C)	0.152	0.116	0.120	0.167	0.171
Proportion of structural variance to total sampling variance: Q = $C/(C+C_0)$	0.50	0.50	0.36	0.54	0.50
C_0 / $(C+C_0)$ *100	0.50	0.50	64.0	46.0	50
R^2	0.79	0.93	0.77	0.77	0.64
RSS	0.0005	0. 0002	0.0074	0.0005	0.0008
AIC	- 1804.18	- 1936.13	- 1416.16	- 1804.18	- 1736.50
Neighbourhood	6	6	6	6	6

Table 5: Best interpolation method according to RMSE

Command area	Soil kriging	IDW
B. Baria	3.458	3.550
Faridpur	6.6529	6.6872
Tala	3.458	3.572
Sonargaon	4.4413	4.7120

and the low zone of the command area. This may be due to more As deposition in the lowest areas compared to the elevated areas. Besides, due to the variability in soil characteristics a safe level in one soil may be unsafe in a different soil.

Cross-validation analysis of interpolated surface of As concentration

The results of cross-validation analysis of interpolated surface of As concentration based on scatter diagram and the estimated regression lines are displayed in Figure 8, and the results are summarised in Table 6 for all study areas to satisfy the hypothesis testing regression coefficient (β) of estimated interpolated values on actual values, $\beta = 1$, with alternative hypothesis $\beta \neq 1$, showed that the regression coefficient in all study areas were close to 1. This indicates a fairly good agreement of the estimated regression line with the 45- degree line on the graph.

CONCLUSION

Detailed geo-statistical analysis of As distribution helps to gather more information on the variability of As concentrations in soil. The spatial dependency of soil As gives an indication that longer irrigation channels prior to inlets into rice fields, and prevention of standing water for long periods in rice plots could be positive measures in reducing As loading in irrigated rice soils. Moderate to strong spatial dependence was observed at micro level. At all levels of command areas the As concentrations in soil varied widely. The concentration of As in soil decreased with the distance from the tube well and with increase of soil OM. Besides, soil As content was

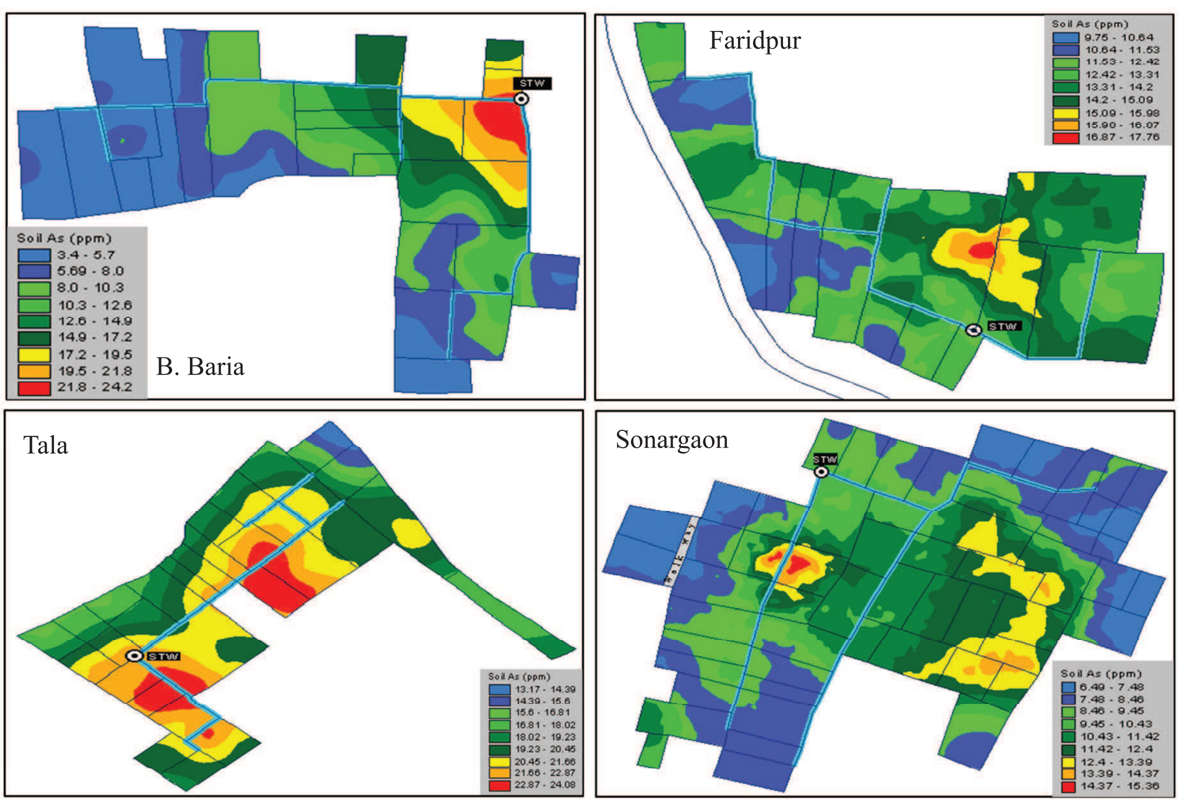

Figure 7: Map of spatial variation of soil As in four locations

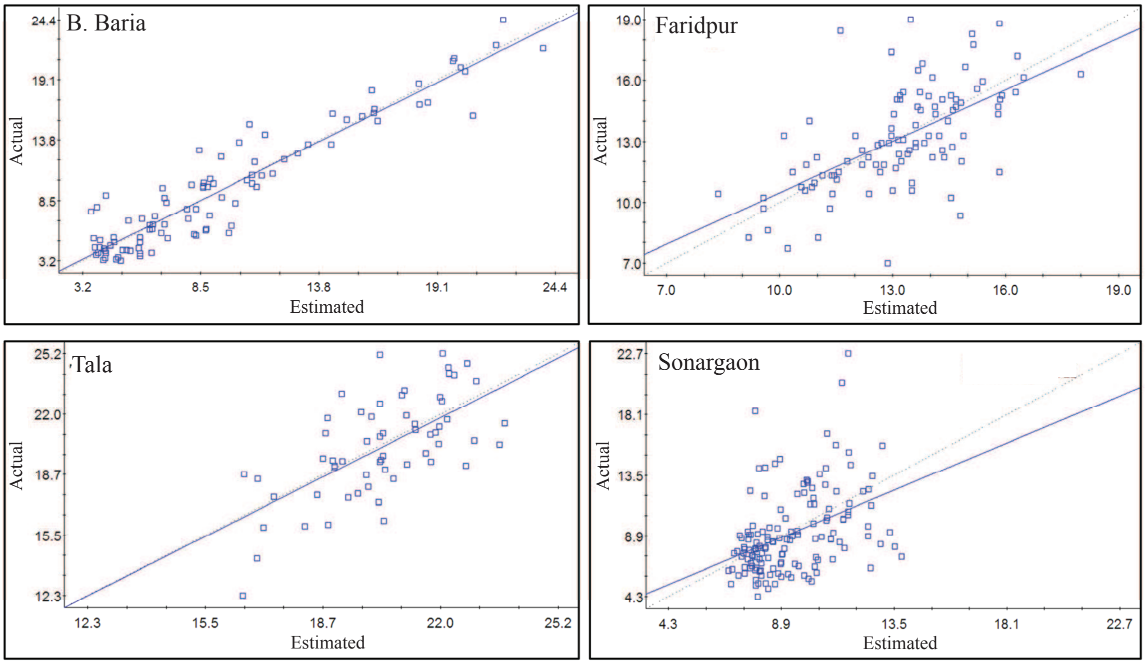

Figure 8: Cross validation interpolated surface of soil As in four command areas

Table 6: Regression coefficient for cross validation of interpolated surface

Command area	Interpolated surface Soil As	
	Regression coefficient (b)	SE
B. Baria	1.00 ns	0.073
Faridpur	0.85 ns	0.109
Tala	0.98 ns	0.153
Sonargaon	0.76 ns	0.129

higher for soils with higher silt and lower with higher clay content. Within the command area, the nature as well as the extent of dependency was not the same in all locations. The spherical model, in general, fitted well to almost all the cases in command areas. Based on the results the geo-statistic method (kriging) was relatively more precise than the IDW method for micro level studies.

REFERENCES

1. Abedin M.J. (2002). As uptake, metabolism and toxicity in paddy rice (*Oryza sativa* L.). *PhD thesis*, pp. 311 – 319, University of Aberdeen, Aberdeen, UK.
2. Acharyya S.K., Chakraborty P., Lahiri S., Raymahasay B.C., Guha S. & Bhowmik A. (1999). As poisoning in the Ganges delta. *Nature* **401**: 545.
DOI: https://doi.org/10.1038/44052
3. Ahmed S. (2002). Groundwater monitoring network design: application of geostatistics with a few case studies from a granitic aquifer in a semiarid region. *Groundwater Hydrology* **2**: 37 – 57.
4. Ahsan D.A., Delvalls T.A. & Blasco J. (2009). Distribution of arsenic and trace metals in the floodplain agricultural soils of Bangladesh. *Bulletin of Environmental Contamination and Taxicology* **82**: 11 – 15.
5. Alam M.B. & Satter M.A. (2000). Assessment of As contamination in soils and water in some areas of Bangladesh. *Water Science and Technology* **42**: 185 – 193.
6. Alam M.Z. & Rahman M.M. (2003). Accumulation of As in rice plant from As contaminated irrigation water and effect on nutrient content. *Proceedings of the BUET-UNU International Symposium on Fate of As in the Environment*, 5 – 6 February, Dhaka, Bangladesh, pp. 131 – 136.
7. Atteia O., Dubois J.P. & Webster R. (1994). Geostatistical analysis of soil contamination in the Swiss Jura. *Environmental Pollution* **86**: 315 – 327.
8. Barca E. & Passarella G. (2007). Spatial evaluation of the risk of groundwater quality degradation. A comparison between disjunctive kriging and geostatistical simulation. *Environmental Monitoring and Assessment* **137**: 261 – 273.
DOI: https://doi.org/10.1007/s10661-007-9758-3
9. Burrough P.A. & McDonnell R.A. (1998). *Principles of*

Geographical Information Systems. Oxford University Press, New York, USA.

10. Cambardella C.A., Moorman T.B., Novak J.M., Parkin T.B., Karlen D.L., Turco R.F. & Konopka A.E. (1994). Field-scale variability of soil properties in central Iowa soils. *Soil Science Society of America Journal* **58**: 1501 – 1511.
11. Carvalho J.R.P., Silveira P.M. & Vieira S.R. (2002). Geoestatística na determinação da variabilidade espacial de características químicas do solo sob diferentes preparos. *Pesquisa Agropecuária Brasileira* **37**: 1151 – 1159. DOI: https://doi.org/10.1590/S0100-204X2002000800013
12. Cattle A.J., McBratney A.B. & Minasny B. (2002). Kriging methods evaluation for assessing the spatial distribution of urban soil lead contamination. *Journal of Environmental Quality* **319**: 1576 – 1588. DOI: https://doi.org/10.2134/jeq2002.1576
13. Cressie N.O.A. (1993). *Statistics for Spatial Data*, 2nd edition, pp. 900. John Willey and Sons, New York, USA.
14. Deutsch C.V. & Journel A.G. (1998). *GSLIB: Geostatistical Software Library and User's Guide*. Oxford University Press, Oxford, UK.
15. Deutsch C.V. (2002). *Geostatistical Reservoir Modelling*, pp. 124 – 152. Oxford University Press, New York, USA.
16. Duxbury J.M., Mayer A.B., Lauren J.G. & Hassan N. (2003). Food chain aspects of As contamination in Bangladesh: effects on quality and productivity of rice. *Journal of Environmental Science and Health Part A-Toxic Hazardous Substances and Environmental Engineering* **38**: 61 – 69. DOI: https://doi.org/10.1081/ESE-120016881
17. Franssen H.J.W.M.H., van Eijnsbergen A.C. & Stein A. (1997). Use of spatial prediction techniques and fuzzy classification for mapping soil pollutants. *Geoderma* **77**: 243 – 262.
18. Gomes F.P. & Garcia C.H. (2002). *Estatística Aplicada a Experimentos Agronômicos e Florestais: Exposição Com Exemplos e Orientações Para Uso De Aplicativos*, p. 309. Piracicaba: FEALQ
19. Goovaerts P. (1997). *Geostatistics for Natural Resource Evaluation*. Oxford University Press, New York, USA.
20. Gorsich David J. & Genton M.G. (2000). Variogram model selection *via* nonparametric derivative estimation. *Mathematical Geology* **32**(3): 249 – 269. DOI: https://doi.org/10.1023/A:1007563809463
21. Grego C.R. & Vieira S.R. (2005). Variabilidade espacial de propriedades físicas do solo em uma parcela experimental. *Revista Brasileira de Ciência do Solo* **29**: 169 – 177. DOI: https://doi.org/10.1590/S0100-06832005000200002
22. Imperato M., Adamo P., Naimo D., Arienzo M., Stanzione D. & Violante P. (2003). Spatial distribution in heavy metals in urban soils of Neples cities (Italy). *Environmental Pollution* **124**: 247 – 256.
23. Isaaks E.H. & Srivastava R.M. (1989). *An Introduction to Applied Geostatistics*. Oxford University Press, New York, USA.
24. Lark R.M. (2000). Estimating variograms of soil properties by the method-of-moments and maximum likelihood. *European Journal of Soil Science* **51**: 717 – 728. DOI: https://doi.org/10.1046/j.1365-2389.2000.00345.x
25. McGrath D., Zhang C.S. & Carton O.T. (2004). Geostatistical analyses and hazard assessment on soil lead in Silvermines area, Ireland. *Environmental Pollution* **127**: 239 – 248.
26. Meharg A.A. & Rahman M.M. (2003). As contamination of Bangladesh paddy field soil: implication for rice contribution to As consumption. *Environmental Science and Technology* **37**(2): 224 – 234. DOI: https://doi.org/10.1021/es0259842
27. Meuli R., Schulin R. & Webster R. (1998). Experience with the replication of regional survey of soil pollution. *Environmental Pollution* **101**: 311 – 320.
28. Nazari Zade F., Arshadiyan F.Z. & Behnaz Vakily K. (2006). Study of spatial variability of groundwater quality of Balarood Plain in Khuzestan province. *Proceedings of the 1st Congress of Optimized Exploitation from Water Source of Karoon And Zayanderood Plain*. University of Shahre Kord, pp. 1236 – 1240.
29. Nickson R., McArthur J., Burgess W., Ahmed K.M., Ravenscroft P. & Rahman M. (1998). As poisoning in Bangladesh ground water. *Nature* **395**: 338 – 338. DOI: https://doi.org/10.1038/26387
30. Panaullah G.M. *et al*. (11 authors) (2003). The As hazard in the irrigation water-soil-plant system in Bangladesh: a preliminary assessment. *Proceedings of the 7th International Conference on the Biogeochemistry of Trace Elements*, Uppsala, Sweden, June 15 – 19.
31. Rahman M.A., Hasegawa H., Rahman M.M., Rahman M.A. & Miah M.A.M. (2007). Accumulation of As in tissues of rice plant (*Oryza sativa* L.) and its distribution in fractions of rice grain. *Chemosphere* **69**: 942 – 948. DOI: https://doi.org/10.1016/j.chemosphere.2007.05.044
32. Robinson T.P. & Metternicht G. (2006). Testing the performance of spatial interpolation techniques for mapping soil properties. *Computer and Electronics in Agriculture* **50**: 97 – 108. DOI: https://doi.org/10.1016/j.compag.2005.07.003
33. Safari M. (2002). Determination filtration network of Groundwater using geostatistic method. *MSc thesis* (Persian Version), Tarbiyat Modares University, Tehran, Iran.
34. Siqueira G.M., Vieira S.R. & Ceddia M.B. (2008). Variabilidade espacial de atributos físicos do solo determinados por métodos diversos. *Bragantia* **67**: 203 – 211. DOI: https://doi.org/10.1590/S0006-87052008000100025
35. Smith E. (1998). As in the environment: a review. *Advances in Agronomy* **64**: 149 – 195.
36. Van Meirvenne M. & Goovaerts P. (2001). Evaluating the probability of exceeding a site specific soil cadmium contamination threshold. *Geoderma* **102**: 63 – 88.
37. Wang J., Fu B.J., Qiu Y., Chen L.D. & Wang Z. (2001). Geostatistical analysis of soil moisture variability on Da Nangou catchments of the loess plateau, China. *Environmental Geology* **41**(1): 113 – 116.

DOI: https://doi.org/10.1007/s002540100350
38. Wu W., Xie D.T. & Liu H.B. (2009). Spatial variability of soil heavy metals in the three gorges area: multivariate and geostatistical analyses. *Environmental Monitoring and Assessment* **157**: 63 – 71.
39. Zhang X.Y., Lin F.F., Wong M.T.F., Feng X.L. & Wang K. (2009). Identification of soil heavy metal sources from anthropogenic activities and pollution assessment of Fuyang County, China. *Environmental Monitoring and Assessment* **154**: 439 – 449.