



Research article

Determining site-specific background level with geostatistics for remediation of heavy metals in neighborhood soils

Tammy M. Milillo¹, Gaurav Sinha^{2,*} and Joseph A. Gardella Jr.¹

¹ Department of Chemistry, University at Buffalo-SUNY, Buffalo, NY, USA

² Department of Geography, Ohio University, Athens, OH, USA

* **Correspondence:** Email: sinhag@ohio.edu; Tel: +17405930304.

Abstract: The choice of a relevant, uncontaminated site for the determination of site-specific background concentrations for pollutants is critical for planning remediation of a contaminated site. The guidelines used to arrive at concentration levels vary from state to state, complicating this process. The residential neighborhood of Hickory Woods in Buffalo, NY is an area where heavy metal concentrations and spatial distributions were measured to plan remediation. A novel geostatistics based decision making framework that relies on maps generated from indicator kriging (IK) and indicator co-kriging (ICK) of samples from the contaminated site itself is shown to be a viable alternative to the traditional method of choosing a reference site for remediation planning. GIS based IK and ICK, and map based analysis are performed on lead and arsenic surface and subsurface datasets to determine site-specific background concentration levels were determined to be 50 µg/g for lead and 10 µg/g for arsenic. With these results, a remediation plan was proposed which identified regions of interest and maps were created to effectively communicate the results to the environmental agencies, residents and other interested parties.

Keywords: site-specific background value; contaminated; soil contamination; geostatistics; spatial interpolation; indicator kriging; soil remediation

1. Introduction

1.1. Evaluating chemical concentrations in soil

When contamination in urban neighborhoods is examined, there are unique challenges to planning remediation and interpreting results. These challenges arise because of the neighborhoods' proximity to Brownfields and Superfund sites. In these neighborhoods, three questions arise: (i) what types of contaminants are present in the soil; (ii) what concentrations of these contaminants are present; and (iii) where are the contaminants located. A *background level* refers to the concentration of a substance in an environmental media (air, water, or soil) that occurs naturally or is not the result of human activities [1]. *Background level* concentrations are used to establish the lower bound of clean up levels for contaminated soil.

State and federal agencies debate what concentration is characteristic for a given area. *Action levels* are defined as the existence of a contaminant concentration in the environment high enough to warrant action or trigger a response. These may be two to four times the acceptable background level, depending on the contaminant [1]. An action level serves as a clear indicator for immediate cleanup to eliminate the potential risk to human health. The concentration level to which a contaminant is remediated is referred to as the *cleanup level*. It is important to note that, in practice, cleanup levels are generally higher than the background level, but much lower than an *action level*. Understanding and determining the background concentration is essential for establishing accurate cleanup levels, which ultimately minimizes potential health risks in the contaminated neighborhoods.

Table 1 provides an example of terminology describing concepts used by different states as guidelines for contaminant concentration. This suggests that states report concentrations whose definitions vary per state or location (e.g., industrial, residential, playground). For any agency's jurisdiction, background and action values may also be updated over time based on new methods or data. Thus, the impact that background or action values have on determination of risk-based soil clean-up levels varies considerably depending on which agency is consulted. Determining reliable background, cleanup and action levels is, therefore, a complicated and politically fraught task in many remediation planning exercises.

Table 1. Select US State guidelines for Lead and Arsenic, in PPM.

State	Arsenic		Lead		Operational Concentration Definition
	Rural	Industrial/ Commercial	Rural	Industrial/ Commercial	
New York [9]	13	16	63	1000–4300	Action level
New Jersey [10]	20	20	400	600	Soil clean-up criteria
New Hampshire [11]	11	11	51	51	Background level
California [12]	0.067	0.31	80	320	Soil ESLs
Colorado [13,14]	3–14	6–19	400	1200	Background Soil Concentration
Connecticut [15]	10	10	500	1000	Direct exposure criteria
Michigan [16]	5.8	5.8	21	21	Rural background
Minnesota [17]	10	25	400	700	Soil reference values

This paper presents a novel approach to site-specific background concentration for study areas with known heavy metal contaminants in soils. In “rust belt” post-industrial cities of the US, making decisions regarding background levels for remediation do not often use objective criteria. The contribution of this research project is an intuitive, geostatistically informed decision-making framework to choose values for background levels that guide remediation planning in urban areas, especially where residential neighborhoods bordering legacy contaminated sites, such as brownfields and Superfund sites.

1.2. Site-specific background level for remediation planning

Regional background levels cannot be relied on since substantial local variation resulting from natural processes such as type of soil and bedrock [2-6]. The regional background level can lead to incorrect source apportionment, and also leads to unnecessary remediation costs if cleanup is recommended to a much lower, uncharacteristic level. At any specific site, background level concentration can be adversely affected by the cumulative effects of historical land use [3]. A *site-specific background level* is used to produce cost-effective remediation, thus limiting financial responsibility and remediation costs [2,7].

Site-specific background levels are established for a specific contaminant by comparing the study area to a reference site, typically a neighboring area unaffected by that contaminant [4]. These areas are normally similar in physical, chemical, geological, and biological characteristics, and have a similar history of land use. Background level estimation at the nearby reference site is based on stratified sampling to cover all types of soil variations, and must follow certain recommended geochemical and statistical analytical frameworks [7,8]. When no reference site exists, regional background levels are common substitutes, resulting in increased remediation costs. Such costs can impede remediation efforts, causing immeasurable harm to people and their environment and increasing costs of remediation over time.

1.3. Geostatistical modeling and GIS mapping based decision making framework for identifying site-specific background level concentrations and delineating contaminated areas

A fundamental problem for site contamination analyses is that they are often limited by the lack of a reference site to determine a site-specific background level [18]. Although discrete regions of high contamination levels may be identified, they cannot be confidently suggested as remediation sites for lack of reliable site-specific background levels. This work makes an intellectual contribution in this regard specifically by offering an improvement over current remediation planning guidelines. The decision-making framework presented in this paper is based on application of well-established geostatistical interpolation and geographic information system (GIS) based mapping to estimate site-specific background levels based on sampled data *from the contaminated site itself*.

This new decision-making framework relies on indicator kriging models of the contaminated site, which incorporates prior information from the US EPA and state guidelines for background levels. Indicator kriging uses spatial autocorrelation between samples to estimate the probability that the (unknown) attribute value at an unsampled location will exceed a selected threshold value. An indicator kriging map is used to determine a site-specific background levels and delineate areas characterized by the probability (>0.7) of exceeding the background level [19,20].

An assumption of this decision making-framework is that contaminant levels for a large proportion of the study area data should exceed threshold values equal to or less than the site-specific background level. A critical parameter of this geostatistical decision making framework is the minimum proportion ($p_{min} > 0.5$ is recommended) of the study area for which contaminant levels must exceed a threshold value for the latter to be considered background level. Thus, if a site is small and/or it cannot be assumed that a majority (>50%) of the site is uncontaminated, this framework should not be used to determine site-specific background levels. In this study, the minimum proportion (p_{min}) was set to a conservative value of 0.8. Once p_{min} is selected for a site, a range of soil concentration thresholds for the chemical of concern will be found to yield study area proportions greater than p_{min} . To minimize cleanup costs, the goal should be to select the highest possible threshold value (v_{max}) as the site-specific background value. Note that there exists an inverse relationship between p_{min} and v_{max} , since if p_{min} increases, v_{max} must decrease to ensure that a larger proportion of the study area supports contaminant levels exceeding v_{max} .

Since remediation planning is not driven by purely scientific criteria, and site-specific background level determination can be influenced by local social, political, and economic context, it is important to offer a simple, intuitive, and flexible decision-making framework for site-specific background level determination. Thus, a major attraction of this new framework is that it does not automatically generate a site-specific background level, but instead empowers decision makers to interactively explore impacts of different combinations of p_{min} and v_{max} parameters through GIS based geostatistical interpolation modeling and mapping of the spatial pattern of the contamination. Remediation specialists can make scientifically defensible decisions based on maps and statistical summaries generated from the application of this framework. However, they may still need to develop new skills, or have access to technically competent personnel for implementing the various components (mapping, spatial analysis, and geostatistical modeling) of the decision-making framework using GIS software. In the rest of this paper, the geographic context of this study, and the design and validation of the new geostatistical decision-making framework is discussed in detail.

2. Study area

Hickory Woods, a residential neighborhood of approximately one square mile in Buffalo, NY. It is one of many Western NY communities struggling with soil contamination and related health problems [21-26]. This neighborhood is adjacent to a New York State Class 2 Superfund site. The area includes Boone Park, which was found to be highly contaminated with arsenic and has already been remediated (see Figure 1). However, maps of samples and ordinary kriging analysis suggest that arsenic contamination levels are high even outside park boundaries, yet only the park was cleaned up [18]. This is hardly an isolated case of ignoring geostatistics in remediation planning and environmental assessment [27]. However, several important questions remain unanswered. These questions, which deal with the source, spread and extent of the contamination, require collaborative effort from researchers, government agencies and residents of the neighborhood.



Figure 1. Aerial imagery showing the study area of Hickory Woods, in Buffalo, NY.

3. Materials and Methods

3.1. Lead and arsenic point soil samples

Surface and sub-surface soil sample datasets were available from previous studies conducted by the EPA, New York State Department of Environmental Conservation (NYS DEC), the City of Buffalo, and the University at Buffalo, State University of New York. Due to confidentiality concerns, governmental agencies could not consolidate and perform or discuss geospatial analysis on the combined data. Therefore, university students were hired to obtain permission directly from residents to use data samples from each residential lot to create an integrated dataset. All sample point concentrations were established in certified laboratories. Although there likely existed some minor differences in the soil sample analysis techniques between different laboratories, all analyses at the various laboratories were done in accordance with the same analysis protocols established by the EPA. It should be noted that the decision making related to the scientific collection and analysis of soil samples is not a part of the decision-making framework presented in this paper. This framework is based on the (tacit) assumption that the soil sample analysis can be expected to yield accurate measurements of soil contamination for every sampled location.

Because spatial or geostatistical analysis and mapping was not deemed necessary at the time by the agencies, samples were originally located only via postal addresses. The addresses were converted to geographic coordinates using ArcGIS[®] geocoding tools and reference street network data; the results were carefully inspected visually against street maps and high resolution orthoimagery. A few addresses with appreciable geocoding error were manually geocoded using satellite imagery and reference street network data. The impacts of minor location errors on background level estimation were expected to be negligible for this case study.

The surface lead dataset contained 308 sample points, including three duplicate samples that were analyzed by EPA contractors to confirm concentration values that were higher than the linear dynamic range of the analysis. The highest concentration value reported for the duplicate points was used to generate conservative concentration estimates. An additional 330 subsurface sample points were utilized for the co-kriging interpolation. The complete arsenic dataset contained 251 surface sample points and 159 subsurface sample points. For interpolations excluding Boone Park lead samples, 239 surface and 314 subsurface sample points were used, while the reduced dataset for arsenic was composed of 182 surface and 150 subsurface sample points.

3.2. Experimental software

ArcGIS[®] (version 9.3), a commercially-available GIS software suite, and widely used in both government and private settings across the globe, was used for all geospatial data management, spatial analysis, geocoding, and mapping for this research project [28]. All indicator kriging and co-kriging interpolation results reported in this study were generated with the ArcGIS[®] Geostatistical Analyst extension. The interpolated results are automatically available as geostatistical layers for mapping and analysis in ArcMap, a component software of ArcGIS.

3.3. Map creation

A base map was first created which contained parcels and street information as well as the sample locations and concentrations for the individual heavy metal contaminants. All interpolated surfaces were overlaid on this base map to provide visual reference to the community members for whom maps were primarily designed. The parcel boundary dataset was downloaded from Erie County, NY Internet Mapping System, while the dataset mapping street lines was extracted from US Census TIGER files [29] and digitally corrected using ArcGIS[®] vector editing tools and high resolution orthoimagery serving as visual references. The base map was also used in geocoding several samples manually. Map color schemes for displaying sample data and interpolated results were chosen with the help of ColorBrewer[®], an online tool that provides color advice for cartography [30].

3.4. Geostatistical methods

3.4.1. Indicator kriging/co-kriging

Indicator kriging allows the estimation of the probability of unknown attribute values (concentrations) to exceed a certain threshold value at a given location. Indicator kriging provides a flexible interpolation approach that is well suited for datasets where: (1) many observations are below the detection limit, (2) the histogram is strongly skewed, or (3) specific classes of attribute values are better connected in space than others (e.g., low concentrations). To apply indicator kriging at its full potential requires, however, the tedious inference and modeling of multiple indicator semivariograms, as well as the post-processing of the results to retrieve attribute estimates and associated measures of uncertainty [31]. The mathematical basis of indicator kriging procedure is briefly summarized below.

The variable representing original sample values ($z(s)$), where s denotes a specific point location, is transformed into an indicator variable based on a chosen threshold (z_k), such that if the concentration $z(s)$ at any location (s) exceeds z_k , then the indicator is 0, otherwise it is 1:

$$I(s; z_k) \begin{cases} 1, & \text{if } z(s) \leq z_k \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

The expected value of $I(s; z_k | (n))$, conditional on n surrounding sample values for a sampled location (s), is expressed as a probability (Equation 2): [32]

$$E[I(s; z_k | (n))] = \text{Prob}[z(s) \leq z_k | (n)] \quad (2)$$

The hazardous probability that exceeds z_k can then be easily calculated (Equation 3):

$$\text{Prob}[z(s) > z_k | (n)] = 1 - \text{Prob}[z(s) \leq z_k | (n)] \quad (3)$$

At an unsampled location s_0 , the ordinary indicator kriging estimator $I^*(s_0; z_k)$ can be estimated as the probability of a weighted average of n surrounding sample indicator values (Equation 4):

$$I^*(s_0; z_k) = \text{Prob}[z(s) \leq z_k | (n)] = \sum_{j=1}^n \lambda_j I(s_j; z_k) \quad (4)$$

where $I(s_j; z_k)$ represents the indicator values at sampled locations s_j ($j = 1, 2, \dots, n$) surrounding the unsampled location s_0 , and λ_j is the weight assigned to that indicator for the estimation of $I^*(s_0; z_k)$. The estimator must be unbiased (Equation 5), and with minimum estimation error variance. (Equation 6) [33] These conditions can be ensured by computing the weights λ_j from the following specially constructed system of linear equations [19].

$$E[I^*(s_0; z_k) - I(s_0; z_k)] = 0 \quad (5)$$

$$\text{Var}[I^*(s_0; z_k) - I(s_0; z_k)] \text{ is minimum} \quad (6)$$

The IK equation can be easily extended for the addition of a second indicator in Indicator Co-Kriging (ICK) as shown in Equation 7, where $K = 2$: [34]

$$I^*_{\text{coIK}}(s_0; z_{k_0} | (n)) - F(z_{k_0}) - \sum_{k=1}^K \sum_{j=1}^n \lambda_j (s_j; z_k) [i(s; z_k) - F(z_k)] \quad (7)$$

where $F(z_k)$ is simply the expected value of the binary indicator function (Equation 8):

$$F(z_k) = E[I(s; z_k)] \quad (8)$$

ICK is also used to derive conditional cumulative distributions for modeling prediction uncertainty by creating multiple indicator variables at different thresholds (cut-offs) from the same variable [19]. In this study, ICK is used to test if subsurface samples incorporated through a secondary correlated variable can improve results from indicator kriging of only surface samples of lead or arsenic. For a more comprehensive reference for IK and ICK, readers should consult Goovaerts [19].

IK/ICK interpolations produce maps which display the probability that a given location will exceed a threshold concentration value. A summary probability measure, which is the average of all interpolated probabilities, can also be used to characterize the overall probability distribution for an interpolation. This statistic can be interpreted as the probability of exceeding the threshold value for the study area *as a whole*; as the measure increases, more locations are estimated to exceed the threshold value. All kriging methods also yield a probabilistic error associated with each predicted value. Maps of such errors can help assess the quality of the predicted probability for every location.

3.4.2. Cross-validation error analysis

Any interpolation model can be made to yield a set of cross-validation errors for every sample point. This standard method of testing the predictive robustness of a model involves making predictions for every sample by running the model with all other sample points only, and then

comparing the predicted and actual sample values for all samples. ArcGIS® Geostatistical Analyst calculates five cross-validation summary statistics for each interpolation model: *root mean squared error (RMS)*, *average standard error (ASE)* and *RMS standardized error (RMS Std.)* [35]. An ideal interpolation is one that has ASE and RMS values that are nearly identical, signifying good point to point variability, and a RMS Std. value of 1 [35,36]. These error statistics were used in conjunction with graphical plots of measured versus predicted values, interpolated kriging estimates and the original sample values to select the most accurate IK/ICK interpolation model, whose indicator threshold then is deemed the best estimate of the site-specific background level.

3.5. Decision making framework for selecting the site-specific background level

The justification and general approach to selection of the site-specific background level based on indicator kriging has already been described in Section 1.3. In this section, the main principles and all relevant parameters of the proposed decision making framework are explained:

- (a) The fundamental assumption for this decision making-framework is that contaminant levels for a large proportion of the study area data should exceed threshold values that are equal to or less than the site-specific background level. There are two important parameters of this decision-making framework: the minimum proportion of the study area (p_{min}) that must be exceeded by any threshold value that could be considered a background level, and the highest possible threshold value (v_{max}) that can be chosen as the site-specific background value.
- (b) To accurately calculate a maximum threshold value (v_{max}) from IK/ICK interpolations for estimating a reasonable site-specific background level, the minimum proportion of study area p_{min} exceeding v_{max} must be chosen first, keeping in mind that v_{max} and p_{min} are inversely related. In this study, p_{min} was chosen to be 0.8, based on previous soil contamination studies and geostatistical analyses of contamination in the study area. A high p_{min} of 0.8 ensures a lower v_{max} , which in turn would require a smaller area to be remediated, thus reducing remediation costs. Lower projected remediated costs increase the likelihood of cleanup projects being greenlighted. However, in the absence of any background information about a site, lower p_{min} values are necessary. Note that if soil samples and geostatistical analyses suggest that contamination affects more than 50 percent of the site, then this method cannot be used to establish a background value.
- (c) To determine the possible threshold values, the statistical distributions of the samples, and prior knowledge about regional background levels presented in guidelines for remediation planning from various states in the US were used (see Table 1). Table 1 includes both liberal and conservative guidelines to show the wide range of current contaminant concentration regulations across US States. The four potential choices ranged from well below the background level to well above the background level proposed for each contaminant. The selection of candidate thresholds was not based on quartiles or deciles, as is common practice in IK studies [19,20] Thresholds below the lowest known background level in the literature for the contaminant as well as values beyond the 10th and 90th percentiles of sample attribute values were not considered. Based on these criteria, the candidate threshold values chosen in this study for lead were 50, 100, 500 and 1000 $\mu\text{g/g}$, and the candidate threshold values chosen for arsenic were 5, 10, 20 and 30 $\mu\text{g/g}$ [9-17].
- (d) For any given indicator threshold, the interpolation models were validated for different variogram model types (circular, spherical, and Gaussian) and parameters (sill/range/nugget),

and sizes and shapes of the search neighborhoods. Model cross-validation error statistics (mean, root mean square) were used to select the most accurate model for a given threshold. Cross-validation error statistics were generated by estimating the error for every sampled location by developing a geostatistical model without that sample, and then comparing the original (true) sample value with the model predicted value for that sampled location.

- (e) IK was performed with surface points for both lead and arsenic distributions. ICK for both contaminants was performed using surface samples as the primary variable and subsurface samples as a secondary variable. A total of 96 interpolations were generated. A few selected well-performing IK and ICK models were also tested with and without Boone Park samples for both lead and arsenic to test the impact of a confirmed cluster of highly contaminated sites on the site-specific background level. Table 2 presents summary cross-validation statistics for the best arsenic and lead IK/ICK models, based on which the site-specific background models were chosen for the two contaminants in Hickory Woods neighborhood.
- (f) The thresholds that did not yield enough (80% of the total area) high probability (>0.7) locations above the chosen background level threshold were eliminated from consideration. If multiple thresholds were found to meet the criterion, the highest threshold was selected as the site-specific background level to minimize potential remediation costs.
- (g) Finally, the documented regional background level is used to make a final assessment of the information yielded from the IK/ICK analyses. If the site-specific background level determined above is higher than the regional background level, then one of three conclusions can be drawn: (i) it is an indication that the site is widely contaminated, (ii) samples are preferentially located in higher contamination areas, or (iii) the background level percentage was set too conservatively, indicating that an analysis with a higher percentage is necessary.

Table 2. Cross-validation summary statistics for Arsenic and Lead.

Contaminant	Lead	Lead	Arsenic	Arsenic	Arsenic
Interpolation Type	Indicator	Indicator	Indicator	Indicator	Indicator
	Kriging	Co-Kriging	Kriging	Co-Kriging	Co-Kriging
Threshold ($\mu\text{g/g}$)	50	50	10	10	30
Boone Park	No	No	No	No	Yes
Playground					
Variogram Model	Gaussian	Gaussian	Gaussian	Gaussian	Spherical
Anisotropic	No	No	Yes	Yes	Yes
Correction					
Mean	0.01009	0.00997	-0.00569	-0.00424	0.00558
RMS	0.39	0.39	0.48	0.46	0.26
ASE	0.41	0.41	0.47	0.46	0.39
M STD	0.02	0.02	-0.01	-0.01	0.01
RMS STD	0.94	0.95	1.02	1.01	0.66
Probability	0.84	0.84	0.39	0.44	0
Error	0.41	0.41	0.46	0.45	0.39

3.6. Region of interest (ROI)

Since GIS based IK/ICK modeling creates maps of the spatial model of contamination, with contextual site features overlaid on the map, it becomes possible to engage in further visual spatial analysis of the site. In a previous publication [18], the authors have shown the viability of delineating region(s) of interest (ROI) to make decisions about which areas to focus on for soil remediation. ROI delineation is also strongly recommended for making decisions about site-specific background level determination. Candidate ROIs are first determined by delineating individual areas within the site, where the probability of exceeding the chosen site-specific background level is higher than 0.7. In addition to this primary criterion, density of the sample points and sample values, associated estimation errors, parcel boundaries, past cleanup efforts, and remediation costs should all be considered. ROIs may also be ranked based on multicriteria analysis methods, if remediation resources must be prioritized across the site for incremental cleanup stages.

4. Results and Discussion

4.1. Point sample mapping

Figures 2a–b show the lead and arsenic soil surface and subsurface samples mapped on the basemap of the study area. Seven classes proved to be optimal for distinguishing different ranges of concentrations for both lead and arsenic. The concentration intervals were chosen by comparison of soil background levels, contamination action levels, and NY state soil cleanup standards [9]. These interval values were also used for choosing the possible threshold concentrations for the IK and ICK interpolated surfaces. As is displayed in Figures 2a–b, there are known sample values that dictate an immediate need for remediation. Due to extremely high arsenic contamination levels in Boone Park, that area has already been remediated. Every lead and arsenic IK/ICK interpolation model for this study was tested with and without the Boone Park samples to show that a large cluster of high concentration samples substantially affects the site-specific background level.

4.2. Indicator kriging and co-kriging interpolations

4.2.1. Site-specific background level for lead

In a previous study, we have shown that spatial correlation exists between surface and subsurface lead distributions, as well as between surface and subsurface arsenic distributions [18]. For both IK and ICK interpolation methods, the threshold values mentioned in section 3.5(c) were used. As expected, because fewer sample values were above the threshold value, the probability of exceeding the threshold decreased as threshold value increased. For lead indicator kriging, the summary probability measure is 0.84 for a threshold of 50 $\mu\text{g/g}$, and decreases significantly to 0.094 for a threshold of 500 $\mu\text{g/g}$. This demonstrates how strongly estimates are influenced by the number of points used to predict a given probability measure. A low summary probability measure indicates that most of the study area is unlikely to exceed 500 $\mu\text{g/g}$ and a lower threshold value needs to be explored to find the site-specific background value representative of the site at uncontaminated locations.

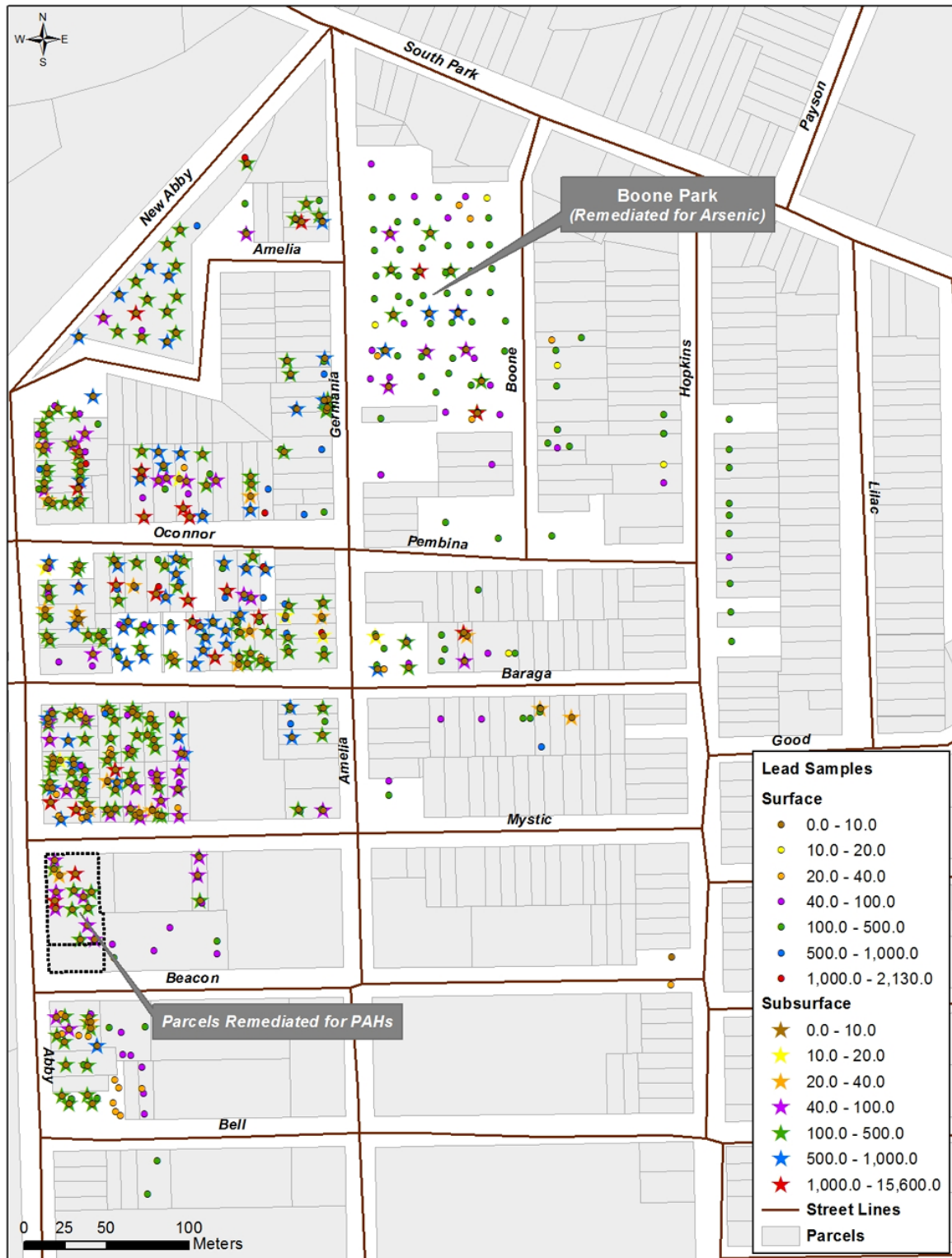


Figure 2a. Surface lead sample points are shown with colored dots, and subsurface lead sample points with colored stars. Areas previously remediated are labeled.

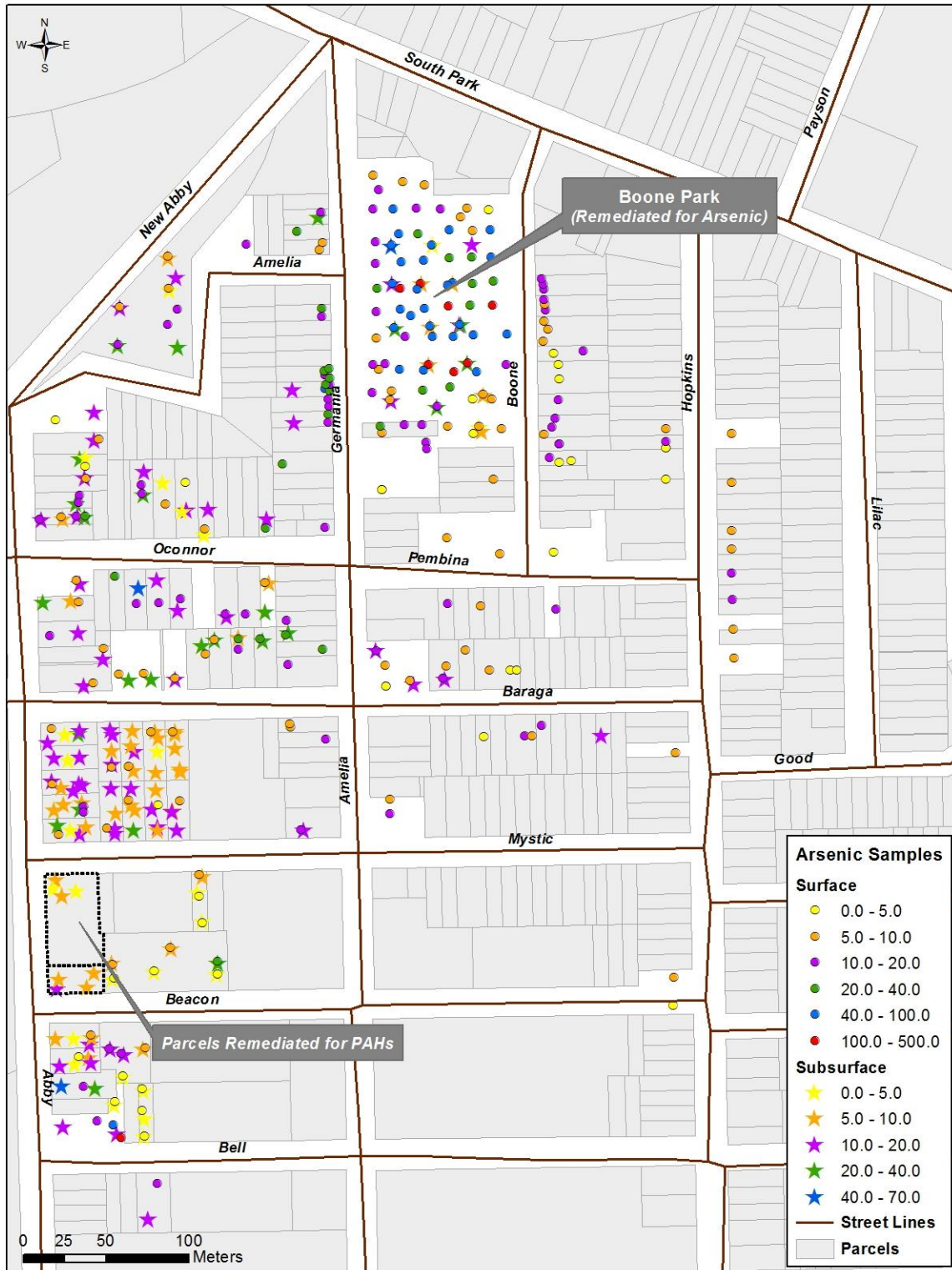


Figure 2b. Surface arsenic sample points are shown with colored dots, and subsurface arsenic sample points with colored stars. Areas previously remediated are labeled.

Interpolation model parameters and summary error statistics were quite similar for IK and ICK interpolations with similar threshold values, which suggests that subsurface samples were not able to improve accuracy. A Gaussian variogram model, with samples from Boone Park excluded, no anisotropic correction, and a threshold of 50 $\mu\text{g/g}$ provided the most accurate estimates for both IK and ICK kriging models. Figures 3a–b shows screenshots (from ArcGIS® Geostatistical Analyst) of the lead variogram models for the chosen IK and ICK models, respectively. Since NY State’s rural background level for lead is also 50 $\mu\text{g/g}$, that provided secondary corroboration for this choice of site-specific background level—and for its geostatistical determination from site samples. The ICK model exhibited marginally better point-to-point variability, which helps accurately predict even small changes between sample locations. Thus, the ICK model was selected to delineate the ROIs for lead remediation.

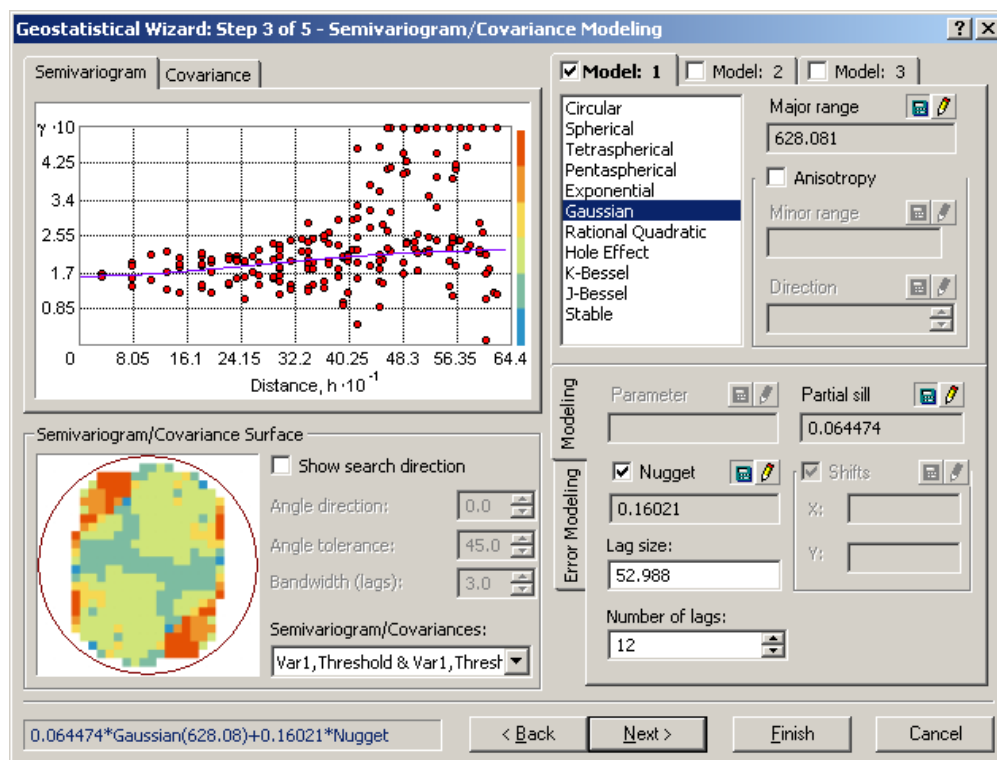


Figure 3a. Gaussian variogram for Indicator Kriging of Lead samples, without anisotropic correction, a threshold of 50 $\mu\text{g/g}$ and without Boone Park samples.

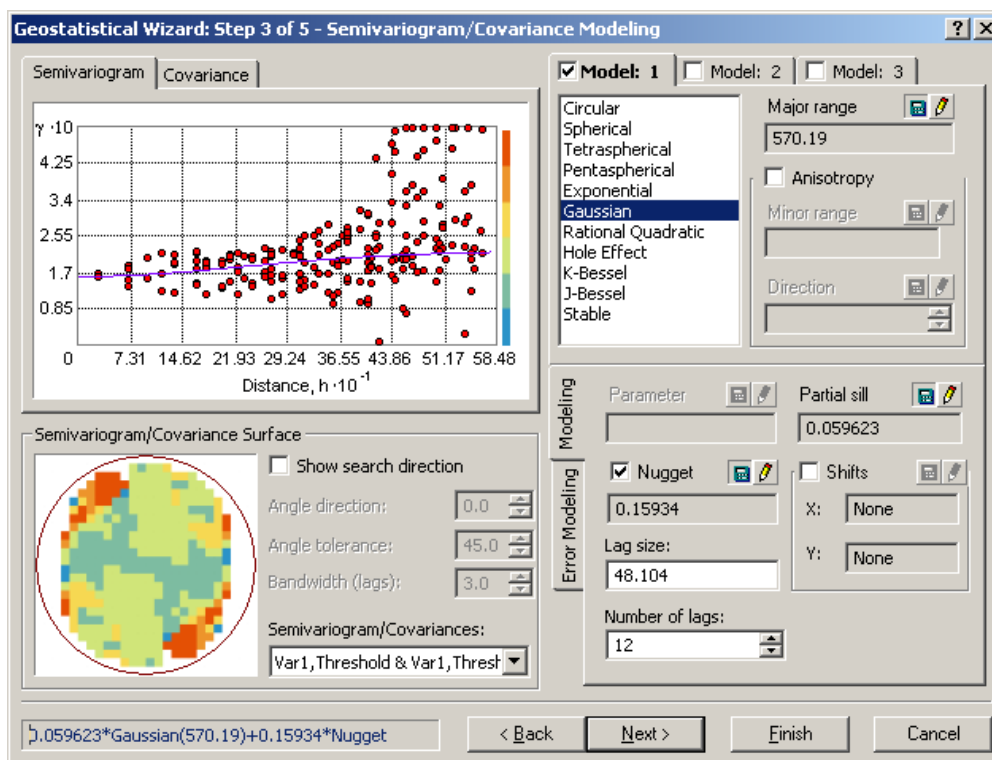


Figure 3b. Gaussian variogram for Indicator Co-Kriging of Lead samples, without anisotropic correction, a threshold of 50 $\mu\text{g/g}$ and without Boone Park samples.

Figure 4a maps the interpolated surface resulting from ICK modeling of lead contamination. A darker shade indicates a higher probability of exceeding the chosen threshold on the map. By comparing the distribution of high probabilities and high sample values, four ROIs were delineated for lead remediation planning where the probability of exceeding the site-specific background level is approximately higher than 0.7. These are labeled L1–L4 in Figure 4a. ROIs serve as additional tools for visual comparison of specific regions across different interpolated surfaces resulting from various methods for each contaminant. In theory, ROIs can be generated and compared for several thresholds to assess risks at different contamination levels. This may help form a multistep remediation plan beginning with locations that are characterized by highest risks (e.g., high probability of exceeding much higher than background levels or even action levels). This kind of spatially explicit remediation planning is not possible without using indicator kriging mapping and ROI delineation, the traditional method is to either remediate the site wholly or often partially based on arbitrary administrative and parcel boundaries. The probability of exceeding the threshold of 50 $\mu\text{g/g}$ across all ROIs is above 0.5 and ranges from 0.59 to 0.98. For ROIs L3 and L4, the probabilities range from 0.79–0.98. Although 50 $\mu\text{g/g}$ is typical of rural areas (see Table 1), the only regions predicted to contain concentrations above this level are located only within ROIs that contained points which exceed the 400 $\mu\text{g/g}$ action level set by the EPA. The inclusion of Boone Park samples caused a minor change in some variogram parameters to maintain the lowest cross-validation errors. However, the base variogram type of the model (Gaussian) and the highest background level meeting the background level percentage remained unchanged.



Figure 4a. Indicator Co-kriging results from lead surface and subsurface sample analysis. The model used a Gaussian variogram, threshold of 10 $\mu\text{g/g}$, and no anisotropic correction. As the probability of exceeding the threshold value increases, the surface color darkens. Labeled outlined represent ROIs.

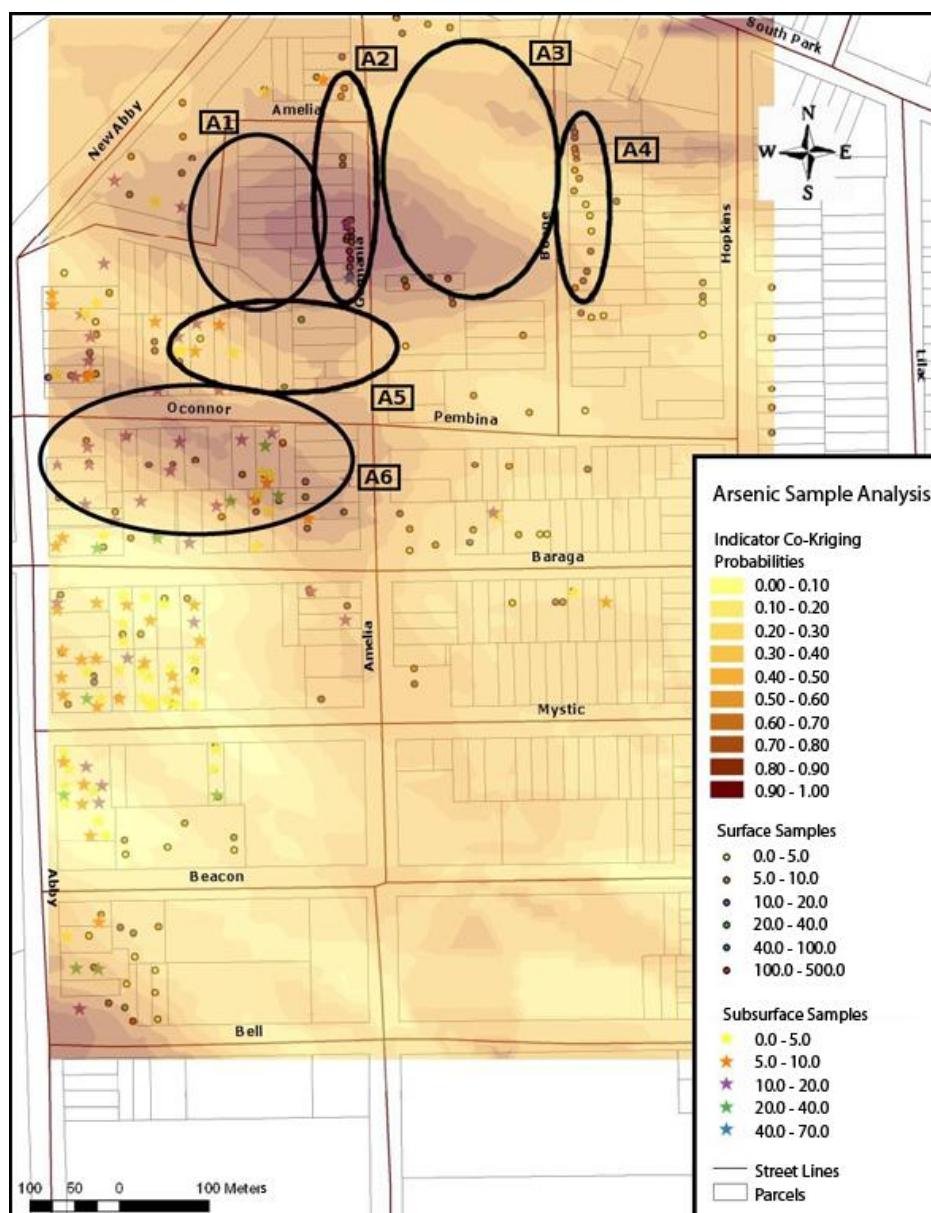


Figure 4b. Indicator Co-kriging results from arsenic surface and subsurface sample analysis. The model used a Gaussian variogram, threshold of 10 $\mu\text{g/g}$, and anisotropic correction. As the probability of exceeding the threshold value increases, the surface color darkens. Labeled outlines represent regions of interest (ROIs).

4.2.2. Site-specific background level for arsenic

Both IK and ICK were also tested for arsenic. Threshold values of 5, 10, 20, and 30 $\mu\text{g/g}$ were chosen based on reasons discussed above. As with the lead distribution, similarly parameterized IK and ICK models for arsenic yielded similar results: the summary probability decreased to approximately zero as the threshold value increased to 30 $\mu\text{g/g}$. This suggests that the background value should be considerably lower. The most accurate IK and ICK models were identical: no samples from Boone Park, a threshold of 10 $\mu\text{g/g}$, and a Gaussian variogram with anisotropic

correction to account for errors associated with unequal sampling (Figures 5a–b). Again, the ICK model had marginally better point-to-point variability and the least overall error, as indicated by the RMS-Std value of 1.023 (compared to 1.010 for the most accurate IK model; see Table 2). The indicator threshold of 10 $\mu\text{g/g}$ used for this ICK model became the estimate of the site-specific background value for arsenic. Figure 4b maps the chosen ICK interpolation surface for arsenic. The color of the interpolation surface darkens as the probability of exceeding the threshold increases. After analysis, six ROIs (A1–A6) were demarcated for identifying areas severely contaminated by arsenic. The probabilities within the six identified ROIs range from 0.58 to 0.96.

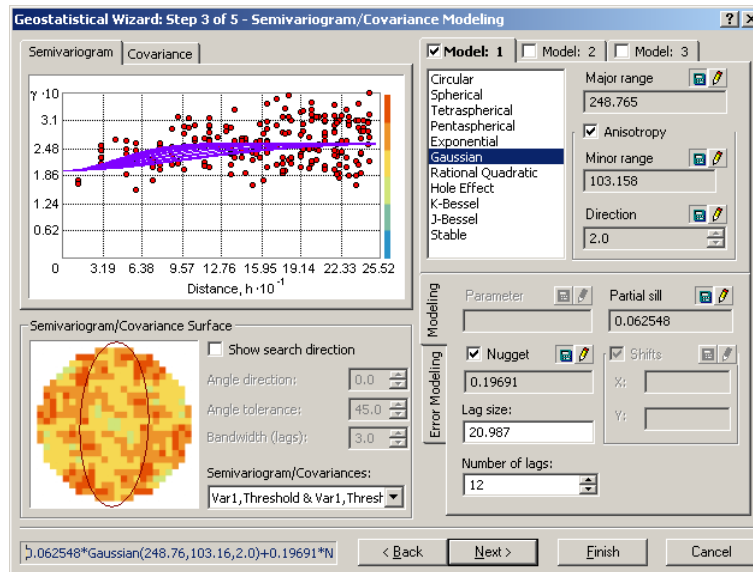


Figure 5a. Gaussian variogram for Indicator Kriging of Arsenic samples, with anisotropic correction, a threshold of 10 $\mu\text{g/g}$ and without Boone Park samples.

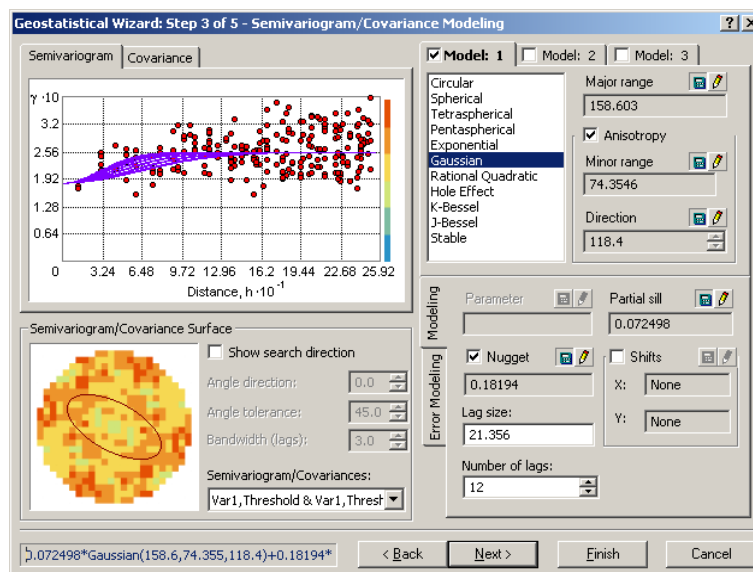


Figure 5b. Gaussian variogram for Indicator Co-Kriging of Arsenic samples, with anisotropic correction, a threshold of 10 $\mu\text{g/g}$ and without Boone Park samples.

This dataset exhibited characteristics of point source pollution. As can be seen in Figure 6, when the surface and subsurface Boone Park samples are included in ROI A3, the model that produces the most accurate interpolation used a threshold of 30 $\mu\text{g/g}$ and a Spherical variogram model. When the interpolation is examined, the only locations with high probabilities of exceeding the threshold value are centered on Boone Park (ROI A3) and its immediate environs (ROIs A2 and A4). These results support previous conclusions that Boone Park and its immediate environs form a separate statistical distribution, due to localized contamination effects. This also serves as validation of this methodology of using IK/ICK and demarcating ROIs, because Boone Park has indeed been determined to be a clear case of arsenic contamination and has already been remediated.

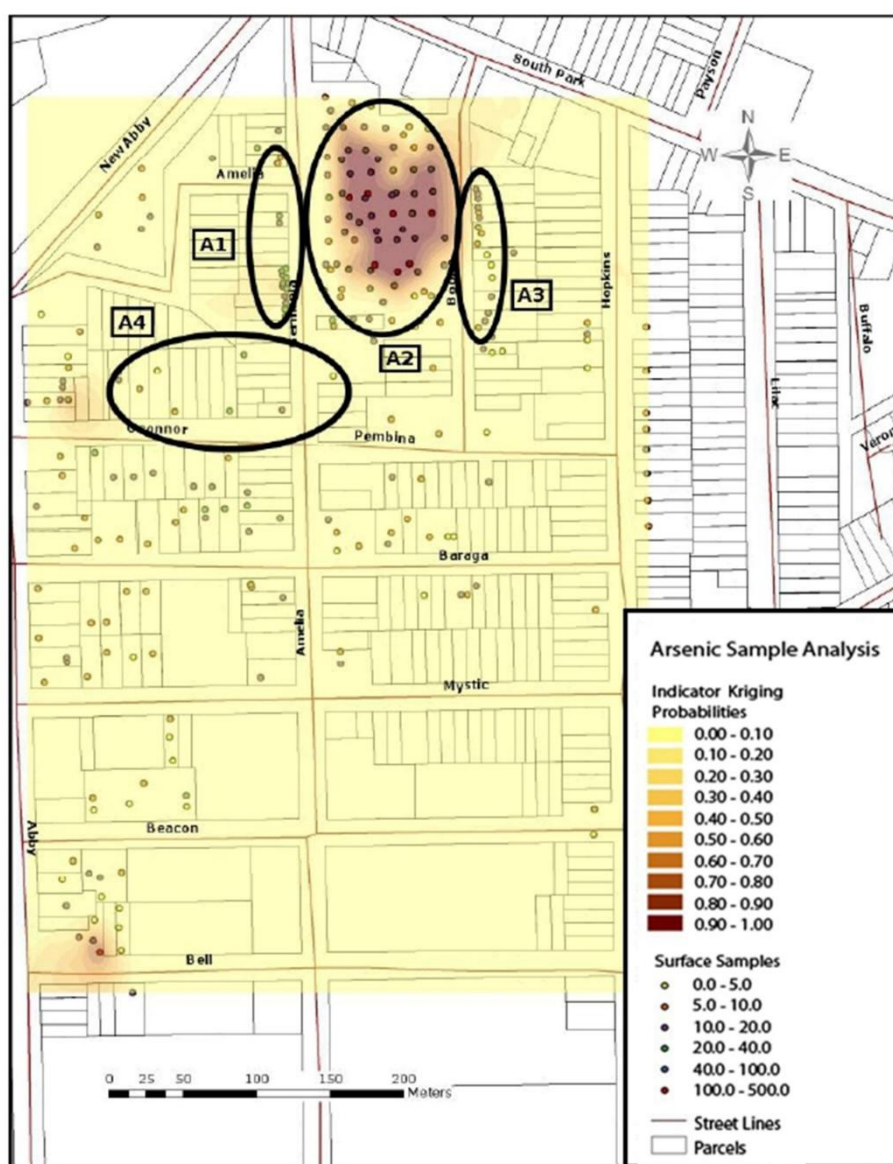


Figure 6. Indicator Kriging Map for Arsenic, constructed using a spherical variogram, threshold of 30 $\mu\text{g/g}$, with anisotropic correction. As the probability of exceeding the threshold value increases, the color darkens. Labeled outlines represent areas of concern.

4.2.3. Mapping using parcel boundaries to communicate with community members

An aspect that is often overlooked in environmental studies is the essential need to communicate results and open discussions with community members and neighborhood residents. Because of the way in which the geostatistical results are often displayed and communicated, residents and community members don't use the statistical analysis to its full potential. To address this limitation, complementary maps were deemed necessary to illustrate the benefits of establishing the site-specific background level and clarify the impact of demarcating ROIs. Parcels were chosen as mapping units to show how contextualizing sample concentrations with the respective site-specific background level helps in a more refined understanding of contamination levels and potential cleanup areas. Residents are concerned about direct impact to their property and neighborhood in which they live. By displaying the maps in this familiar way, the residents can become active partners and voices in the remediation planning process.

Figures 7a–b show the parcel based maps developed for communicating about lead and arsenic contamination, respectively. Neither the ICK predicted probabilities, nor original sample points, are shown on these maps. Each map is specific to the particular contaminant and does not factor in any information about contamination due to the other contaminant. The map symbology strategy is explained as follows. Surface and subsurface sample concentrations were classified by the following intervals: *background level, twice background level, and four times background level*. The parcel color corresponds to the class to which the surface sample point found on that parcel would be allocated to. Similarly, the single/double hash line symbols indicate the classification of the subsurface sample point found on the parcel. If a parcel contained multiple surface and subsurface sample values, the highest surface or subsurface value was used to determine the corresponding surface or subsurface class to be used for choosing the parcel symbology. ROIs from the corresponding maps for lead (Figure 4a) and arsenic (Figure 4b) were also overlaid in Figures 7a–b to compare how ROIs from ICK analysis correspond to the distribution of sample concentrations which far exceed the site-specific background level. It is apparent that the parcels within ROIs determined from geostatistical analysis are indeed characterized by high levels of contamination, well above the respective background levels.

5. Conclusions

In this paper, the goal was to present and show the successful implementation of a new geostatistics based decision-making framework for determining site-specific background levels that can guide soil remediation of contaminated sites. The existing practice of finding a nearby reference area to establish site-specific background levels is often difficult to implement due to lack of an acceptable reference site nearby and the substantial extra costs of collecting and processing samples from the reference site. It was shown in this paper that the new geostatistical decision-making framework based on indicator kriging algorithms, and supplemented with GIS mapping, local and regional background levels supporting, and other social, political and economic contextual information about the site is a viable and cheaper alternative method of determining a site-specific background level. A pre-requisite for the use of this method is that the study area for sampling contains at least 50% of uncontaminated soils. Thus, small site remediation is not possible with this method, since there will not be enough uncontaminated area available to determine site-specific background levels.

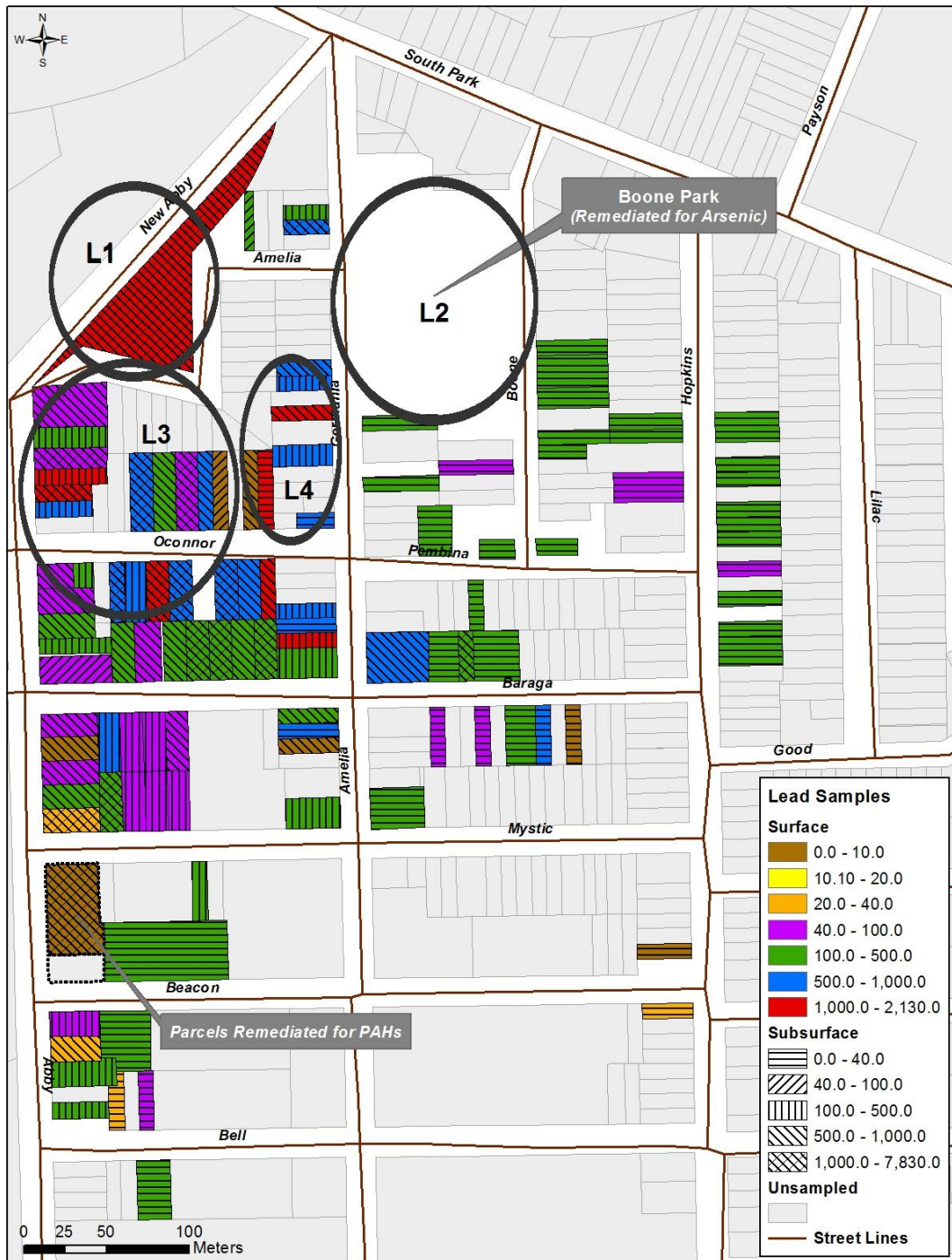


Figure 7a. Map using parcel boundaries to show lead contamination levels in relation to the site-specific background level of $50 \mu\text{g/g}$. Both the highest surface and highest subsurface sample values applicable to a parcel are shown simultaneously. ROIs developed from the map of the chosen lead indicator co-kriging interpolation are also overlaid for reference.

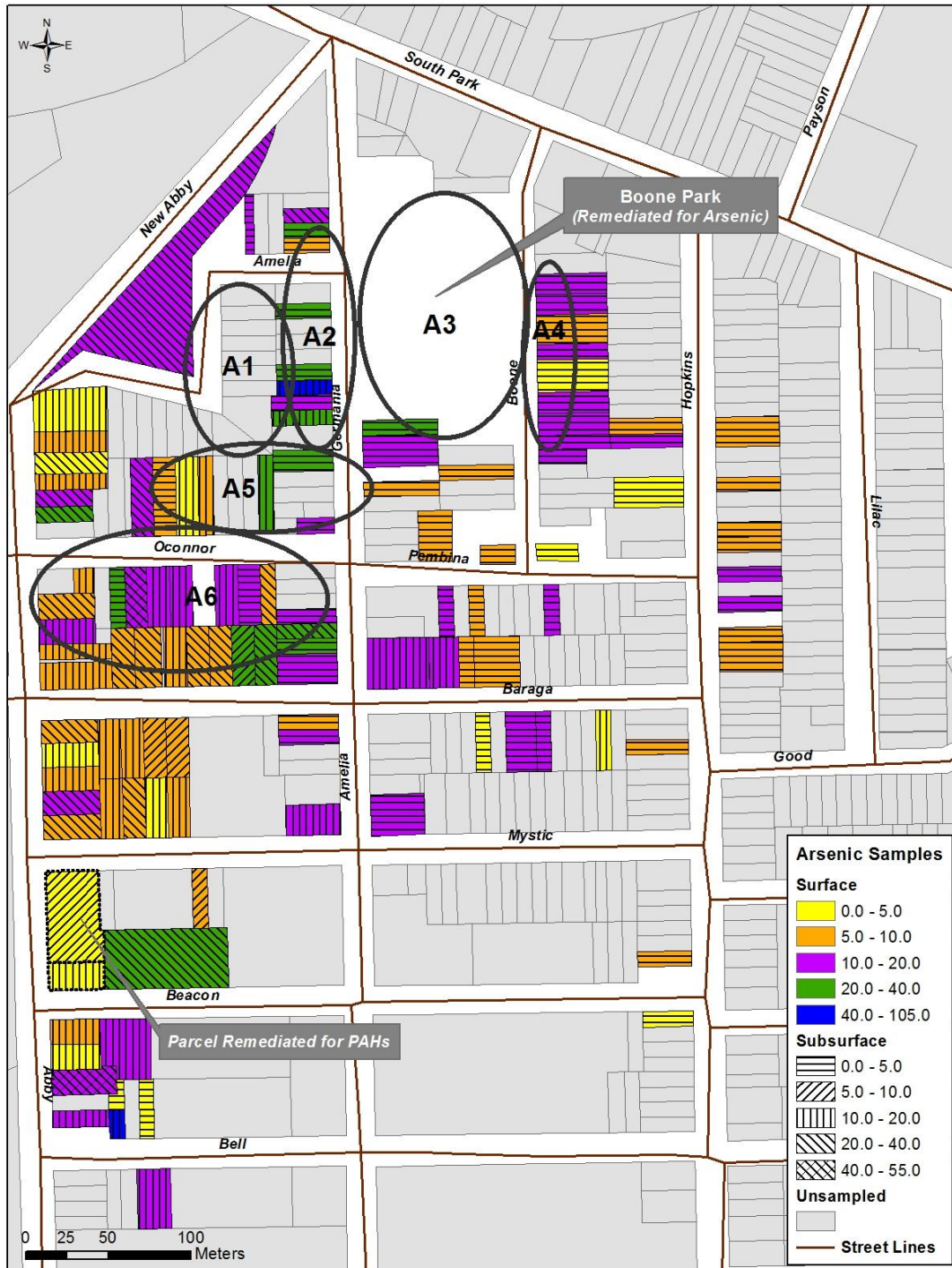


Figure 7b. Map using parcel boundaries to show arsenic contamination levels in relation to the site-specific background level of $10 \mu\text{g/g}$. Both the highest surface and highest subsurface sample values applicable to a parcel are shown simultaneously. ROIs developed from the map of the chosen arsenic indicator co-kriging interpolation are also overlaid for reference.

Based on the developed framework, the interpolation techniques of IK and ICK were applied to datasets collected for the Hickory Woods neighborhood to identify site-specific concentrations based on samples from the contaminated site itself. ICK with subsurface samples was determined to be marginally better than IK with only surface samples. The ICK models indicate the site-specific background levels to be 50 µg/g for lead and 10 µg/g for arsenic. These concentrations coincide with guidelines for lead and arsenic (see Table 1). The lead concentration of 50 µg/g is significantly lower than the EPA action level of 400 µg/g, but consistent with rural background levels reported in New York State (20–50 µg/g). ROIs identified with high probabilities of exceeding the background level also encompass known areas of high concentration values well above the allowable limit. The results of this analysis also confirmed that Boone Park is characterized by localized contamination behavior, uncharacteristic of the rest of the study area. Remediation has already occurred in Boone Park; however, areas adjacent to the park have not been remediated and need to be.

Acknowledgements

This work was supported by grants from the National Science Foundation Chemistry Division (CHE-0316735 and CHE-0616916). The authors acknowledge Christopher Case for his support on some aspects of the project that led to this paper.

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. US Environmental Protection Agency. Terminology Reference System. Available from: https://iaspub.epa.gov/sor_internet/registry/termreg/searchandretrieve/termsandacronyms/search.do.
2. Wyoming Department of Environmental Quality, Establishing Site-Specific Background Metal Concentrations in Soil, 2000.
3. US Environmental Protection Agency, *Establishing Background Levels*, 1995.
4. US Environmental Protection Agency, *Guidance for Comparing Background and Chemical Concentration Levels in Soil for CERCLA Sites*. Office of Emergency and Remedial Response, Washington, DC 20460, EPA 540-R-01-003, 2002.
5. Stensvold KA, Scientific Investigations Report 2011-5202, US Department of the Interior, US Geological Survey, 2012, 1-27.
6. Vosnokis KAS, Perry E (2009) Background versus Risk-Based Screening Levels: An Examination of Arsenic Background Soil Concentrations in Seven States. *Int J Soil Sediment Water* 2: 2.
7. Mortefolio MJ, Derivation of Site-Specific Arsenic Background in Soil: A Case Study. *Proceedings of the Annual International Conference on Soils, Sediments, Water and Energy*. 2010, 12: 6.
8. Love D, Loock D, Kuzyk ZZ, et al. (2005) Development of Site-Specific Environmental Criteria from Background Data. In: *Assessment and Remediation of Contaminated Sites in Arctic and Cold Climates (ARCSACC)*, Edmonton, Alberta.

9. New York Department of Environmental Conservation, Determination of Soil Cleanup Objectives and Cleanup Levels, 1994.
10. New Jersey Department of Environmental Protection, Soil Cleanup Criteria (Site Remediation Program), 2015.
11. New Hampshire Department of Environmental Services, NHDES Contaminated Sites Risk Characterization and Management Policy, 1998.
12. California Regional Water Quality Control Board, Application of Risk-Based Screening Levels and Decision Making to Sites with Impacted Soil and Groundwater, 2001.
13. State of Colorado Department of Public Health and Environment, Hazardous Materials and Waste Management Division (2014) Arsenic Concentrations in Soil Risk Management Guidance for Evaluating.
14. U.S. Environmental Protection Agency (2014) Regional Screening Level (RSL) Summary Table.
15. State of Connecticut, Department of Energy and Environmental Protection (2013) Regulation of Department of Energy and Environmental Protection Concerning Remediation Standard.
16. Michigan Department of Environmental Quality (2013) Cleanup Criteria Requirements for Response Activity (Formerly the Part 201 Generic Cleanup Criteria and Screening Levels).
17. Minnesota Pollution Control Agency (1999) Site Remediation Section: Risk-Based Guidance for Soil-Human Health Pathway.
18. Milillo TM, Sinha G, Gardella JA (2012) Use of Geostatistics for Remediation Planning to Transcend Urban Political Boundaries. *Environ Pollut* 170: 52-62.
19. Goovaerts P, Kriging versus Stochastic Simulation for Risk Analysis in Soil Contamination. In: geoENV I—Geostatistics for Environmental Applications. Springer Netherlands, 1997: 247-258.
20. Webster R, Oliver MA, Geostatistics for Environmental Scientists. John Wiley & Sons, 2007, 330.
21. Fairbanks P, 'Legitimate questions' on toxics and illness. *Buffalo News*, p. B1, 25 March 2001.
22. Fairbanks P, Taking Action. *Buffalo News*, p. B1, 8 September 2001.
23. Fairbanks P, Herbeck D (2002) LTV Steel to Help Pay for \$16 Million Cleanup. *Buffalo News*, 16 July 2002.
24. Groll M (2002) Toxic Neglect: City Had Early Warnings of Contamination. *Buffalo News*, p. A1, 11 February 2002.
25. Gardella JA, Milillo TM, Sinha G, et al. (2007) Linking Community-Service Learning and Environmental Analytical Chemistry. *Analytical Chemistry* 79: 811-818.
26. Gardella JA, Milillo TM, Sinha G, et al. (2009) Linking Advanced Public Service Learning And Community Participation With Environmental Analytical Chemistry: Lessons From Case Studies In Western New York. In: Redlawsk D, Rice T (Eds.), Service Learning with Government Partners. San Francisco, CA: John Wiley & Sons, 98-112.
27. Verstraete S, Van Meirvenne M (2008) A Multi-Stage Sampling Strategy for the Delineation of Soil Pollution in a Contaminated Brownfield. *Environ Pollut* 154: 184-191.
28. Environmental Science Research Institute (ESRI). Available from: <http://www.esri.com>.
29. United States Census Bureau (2000) 108th Congressional District Census TIGER Line File. Available from: <ftp://ftp2.census.gov/geo/tiger/tgred10>.
30. Brewer C, Harrower M. ColorBrewer 2.0. Available from: <http://colorbrewer2.org>.

31. Goovaerts P (2009) AUTO-IK: A 2D Indicator Kriging Program for the Automated Non-Parametric Modeling of Local Uncertainty in Earth Sciences. *Comput Geosci-UK* 35: 1255-1270.
32. Lin YP, Chu HJ, Wu CF, et al. (2011) Hotspot Analysis of Spatial Environmental Pollutants Using Kernel Density Estimation and Geostatistical Techniques. *J Environ Res Public Health* 8: 75-88.
33. Antunes I, Albuquerque MTD (2013) Using Indicator Kriging for the Evaluation of Arsenic Potential Contamination in an Abandoned Mining Area (Portugal). *Sci Total Environ* 442: 545-552.
34. Goovaerts P (1994) Comparative Performance of Indicator Algorithms for Modeling Conditional Probability Distribution Functions. *Mathamatical Geology* 26: 389-411.
35. Johnston K, Ver Hoef JM, Krivoruchko K, et al. Using ArcGIS Geostatistical Analyst. Redlands: ESRI, 2001.
36. Isaaks EH, Srivastava RM, An Introduction to Applied Geostatistics, New York: Oxford University Press, 1989, 592 p.



AIMS Press

© 2017 Gaurav Sinha et al., licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)