Development of a GeoSpatial Exposure Model (GeoSEM) for Human Health and Ecological Risk Assessments.

William Thayer[1, 2], Philip Goodrum[1, 2], Gary Diamond[1] and James M. Hassett[2],

[1]Syracuse Research Corporation, 6225 Running Ridge Road, North Syracuse, NY 13212,
[2]Faculty of Environmental and Forest Engineering, Bray Hall, SUNY-College of Environmental Science and Forestry, Syracuse, NY 13210.

Introduction

One source of uncertainty in risk assessment is the exposure point concentration (EPC): the chemical concentration to which a human or ecological receptor may be exposed to for a toxicologically relevant time period within a geographic area called an exposure unit (EU). Biased sampling methods are often employed during site characterization. This, along with the assumption that contaminants are log-normally distributed, may contribute to overly conservative estimates of the EPC. Geostatistical methods allow the spatial information present in sample data to be incorporated in the EPC estimate, which should reduce the uncertainty in the EPC and risk estimates. However, the application of geostatistics introduces model uncertainty into risk estimates and management decisions. The available geostatistics software packages are not designed for exposure assessment and, therefore, require considerable experience in geostatistics to produce estimates for human health and ecological exposure assessment. We have developed the GeoSpatial Exposure Model (GeoSEM), a software tool that combines geostatistical algorithms and mapping capabilities in a format that does not require the user to be an expert in geographic information systems (GIS) or geostatistics.

Background

Typically, the 95th upper confidence level (95th UCL) on the mean chemical concentration measured on the stie is used as the EPC (EPA 1989, 1992). The use of geostatistics can reduce the uncertainty in the exposure point concentration, thereby producing more accurate estimates of risk.

We are developing a geospatial exposure model (GeoSEM) that merges data analysis capabilities (i.e., geostatistics algorithms) with a visual/mapping interface. The goal of the software development effort is to provide a user-friendly tool for risk assessors to incorporate spatial information in estimates of the EPC, thereby improving the accuracy of risk estimates and increasing the effectiveness of remedial actions at contaminated sites.

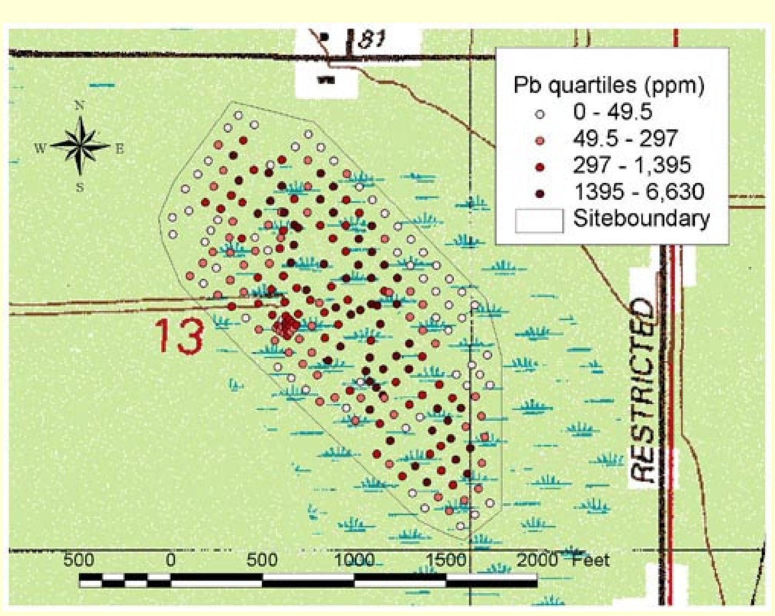

Figure 1. The spatial distribution of soil samples are shown. Samples were collected from the top 6-inches of soil. The figure indicates the sampling intensity is slightly greater near the center of the site and in one location near the south-central portion of the site, where high concentrations were anticipated. The spatial autocorrelation present in the data is indicated by the tendency for samples located next to each other to have similar lead concentrations.

Statistic	Traditional (unweighted)	Spatially-weighted
count	232	232
minimum (ppm)	0.5	0.5
maximum (ppm)	6,630	6,630
mean (ppm)	924	887
stdev (ppm)	1,322	1,340
CV (stdec/mean)	1.4	1.5
skew	1.9	1.9
kurtosis	3.0	3.0

Table 1. Descriptive statistics for the soil lead data are shown. Modest spatial clustering of soil samples near the center of the site, where the highest concentrations were observed, results in an unweighted mean that is slightly greater than the spatially-weighted mean. Spatial weights for the samples were determined using Thiessen polygons.

Objectives

1. Compare estimates of the EPC calculated using traditional methods with those that consider the geospatial information contained in sample data.
2. Show how spatial statistics can be used to assess uncertainty in the exposure point concentration.
3. Encourage the use of spatial statistics in exposure assessment.

Methods

Data collected from a skeet and trap range were used to compare traditional estimation methods with geospatial methods. Figure 1 and Table 1 show the spatial distribution of the samples and descriptive statistics for the data, respectively. The first comparison of traditional methods with geospatial methods is made using the data for the entire site. This is consistent with treating the entire site as one EU. Table 2 shows the 95[th] UCL on the mean (i.e. EPC) and the probability of the EPC exceeding selected concentration thresholds. The thresholds could represent risk-based concentration (RBC). The RBC represents a concentration above which an unacceptable risk of exposure occurs. The application of geostatistics in the design of future sampling plans is illustrated in Figures 2-4. The site has been divided into 10 EUs. Kriging and traditional methods were then used to estimate the EPC (i.e., mean soil concentration) for each EU. GeoSEM was used to estimate the probability that the EPC exceeded an risk-based threshold of 1000 ppm. Figure 2 shows the results using the ordinary kriging method; maps prepared using traditional methods are also shown for comparison (Figures 3-4).

Estimate	Method				
	Traditional (unweighted)		Spatially-weighted		Kriging
	normal	log-normal	normal	log-normal	
mean	924		887		867
95[th]UCL	1,067	8,887	1,032	10,629	917
Pr EPC>900	0.66	>0.99	0.44	>0.99	0.14
Pr EPC>1000	0.19	>0.99	0.10	>0.99	<0.01
Pr EPC>1100	0.02	>0.99	0.01	>0.99	<0.01

Table 2. Comparison of traditional and geospatial methods for modeling uncertainty in the EPC. Traditional methods indicate an unacceptable risk if the RBC is less than 1,067 ppm or 8,887 ppm, depending upon whether a normal or lognormal distribution is assumed. The kriging method indicates an unacceptable risk if the RBC is less than 917 ppm, which is less than both of the traditional methods. The lower estimates produced by the geospatial methods reflect additional information extracted from the data: i.e., the spatial arrangement of the data (in this case, the clustering of samples in high concentration areas) and, in the case of kriging, the spatial autocorrelation ('continuity') structure.

Exposure unit	Count (n)	Traditional (unweighted)		Spatially-weighted		Kriging
		normal	log-normal	normal	log-normal	
1	17	<0.01	0.67	<0.01	0.47	<0.01
2	11	0.02	0.87	0.01	0.80	<0.01
3	38	<0.01	0.06	0.01	0.16	<0.01
4	27	0.96	0.99	0.95	0.99	>0.99
5	27	0.01	0.40	<0.01	0.64	<0.01
6	25	0.94	>0.99	0.95	>0.99	>0.99
7	25	0.82	>0.99	0.72	>0.99	>0.99
8	19	0.90	0.97	0.74	0.84	>0.99
9	26	0.55	0.98	0.25	0.95	<0.01
10	17	<0.01	0.60	<0.01	0.42	<0.01

Table 3. Geospatial methods are compared to traditional methods for estimating the probability of exceeding an RBC of 1,000 ppm for each of the 10 exposure units (EUs) on the site (see Figures 2-4). This information can be used to help decide on future courses of action: future sampling to reduce uncertainty, remediation or no action. Comparison of the methods for EU 9: reveal that kriging indicates a low likelihood that the EPC exceeds the RBC while traditional methods indicate the opposite.

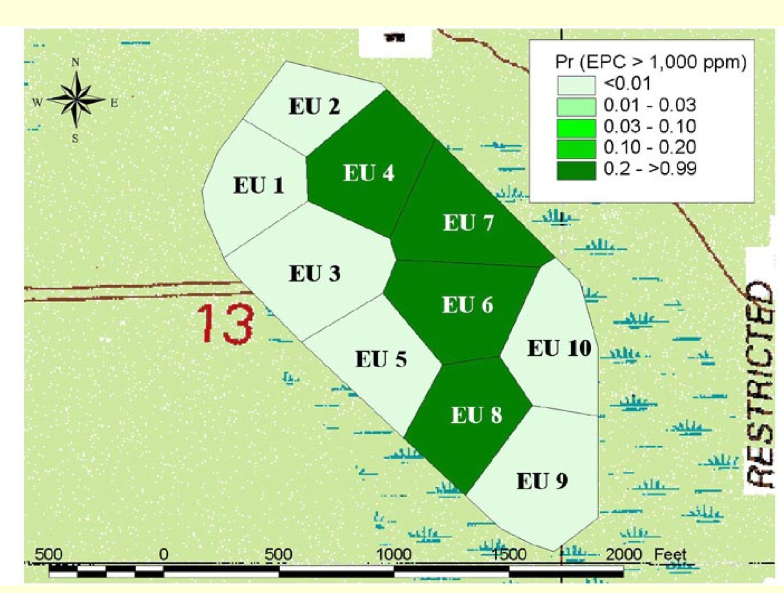

Figure 2. This map shows the probability of exceeding the RBC of 1,000 ppm, using kriging to estimate the probabilities. Note that approximately half of the EUs do not require remediation. The form of kriging used in this analysis ('ordinary kriging') assumes a normal distribution for the sampling distribution of the EPC. Kriging and the normal-based approach use different methods to estimate the variance of the EPC. Kriging considers the spatial autocorrelation of the soil concentration data and the geographical distribution of the sample locations in calculating the estimate; the traditional methods do not.

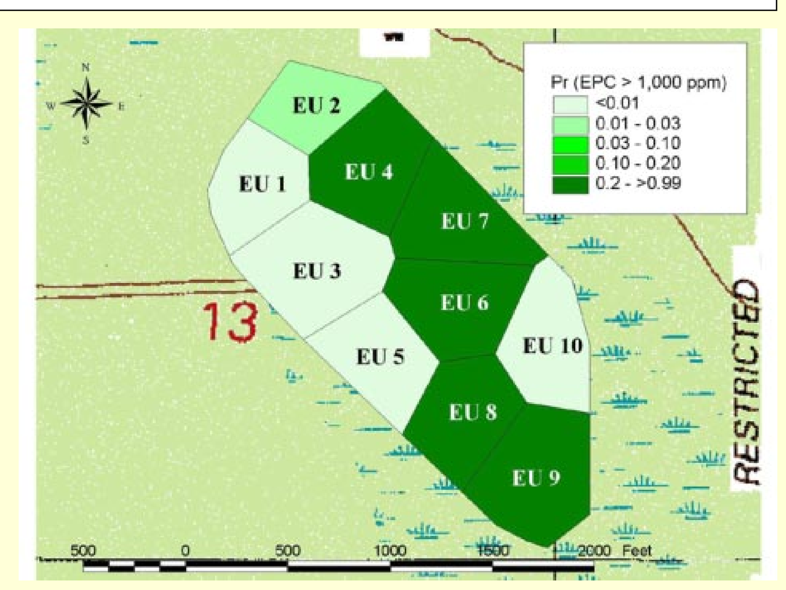

Figure 3. The probability of exceeding a RBC of 1,000 ppm, assuming a normal distribution for the sampling distribution of the EPC is shown. Typically, probabilities greater than 0.05 indicate some type of remedial action is required (i.e., 95th UCL > RBC). This map indicates there is a very low probability that remedial action is required in exposure units (EUs) 1, 2, 3 and 5, while it appears likely that some response would be required for EUs 4 and 6-9.

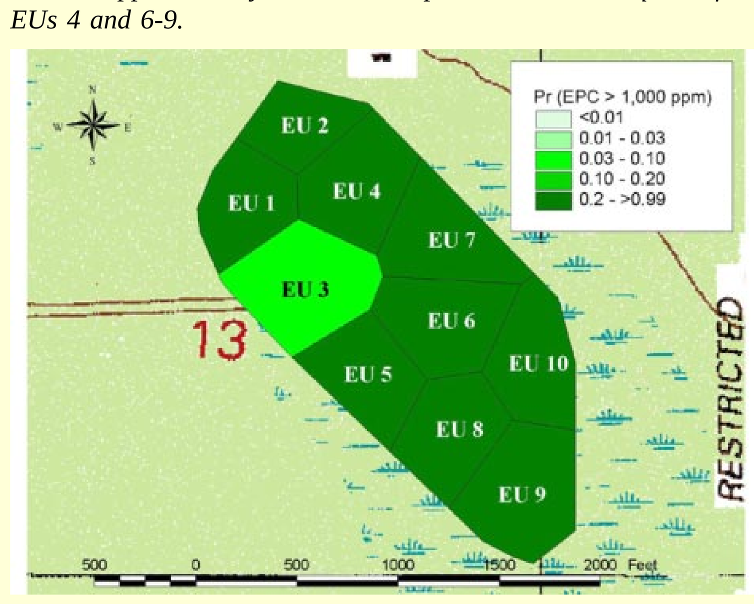

Figure 4. This probability of exceeding an RBC of 1,000 ppm, assuming a log-normal distribution of soil concentration. The EPC for all but one EU is likely to exceed the RBC. The probability that the EPC for EU3 exceeds the RBC is close to the 'decision point' of 0.05, indicating additional sampling to reduce uncertainty in the estimate of the EPC would be worthwhile. The assumption of a log-normal distribution for the soil concentration increases the area where remedial action is indicated by approximately 50%.

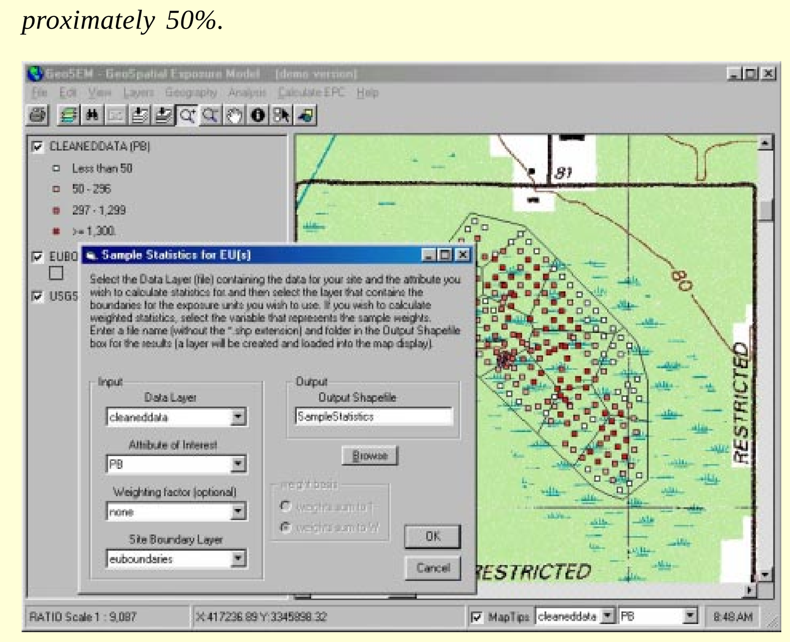

Figure 5. The GeoSEM interface for calculating statistics for each exposure unit is shown. The interface allows the user to select the chemical for which statistics are desired and provides these either for the entire site or by EU. The spatial querying tools allow the risk assessor to include only those samples within a geographic region of interest in the analysis. Standard GIS tools allow the user to zoom, pan and display information on geographic features. GeoSEM can display many standard image formats, including USGS topographic maps shown here.

Results

The geospatial estimates of the mean and 95[th] UCL shown in Table 2 account for the cluster of samples in the region of the site with the highest concentrations. In effect, the geospatial methods account for the redundant information produced by the clustered samples; in addition, kriging (and other geostatistical methods) account for the spatial continuity (i.e., spatial autocorrelation) present in the data. The lognormal-based estimate of the 95[th] UCL (8,887 ppm) is indicative of a commonly encountered problem with assuming a log-normal distribution for concentration data. Considering the number of samples collected on the site (232) and the preferential sampling of the areas with high concentration, it is hard to accept that the estimate of the mean could possibly be underestimated by over 900%. This work indicates the value of geospatial analysis in the assessment of contaminated sites. Incorporating geospatial information in the estimate of the EPC is intuitively appealing. If the potential receptor is assumed to have random access to the contaminated medium (e.g., soil) (EPA 1989), the EPC should be estimated by a method that considers the spatial distribution of soil contamination. Figures 2-4 illustrate the value of using a geospatial approach to sampling design and remedial action decision-making. This type of geospatial analysis can be used to decide if additional samples should be collected to reduce uncertainty in the EPC (e.g., if the probability of exceeding the EPC is greater than a pre-determined level, such as 0.05) or if additional effort should be focused on remedial design of the EU. Geospatial analysis is also valuable in remedial design to determine which portions of the EU require remediation.

A screen capture of the GeoSEM interface is shown in Figure 5. GeoSEM will connect risk assessors with easily implemented, robust geospatial statistical routines, such as kriging (ordinary, indicator, log-normal and normal score kriging) and simulation (Gaussian and indicator); and area weighting approaches such as Thiessen polygons, within a single software platform. GeoSEM is a GIS-based application that is designed to run in the Microsoft Windows® environment. GeoSEM is capable of using a wide variety of GIS vector and raster file formats including DOD vector product format (VPF), ESRI ArcView® Shapefiles, ARC/INFO® coverages and spatial database engine (SDE®) layers, computer-aided design (CAD) drawings, binary and ascii grids and many types of standard georeferenced image formats such as geoTIFFs, bitmaps, GIFs, JPEGs and ERDAS.

Future Research

We will compare the performance of simulation and kriging algorithms to identify user-friendly tools that provide accurate estimates of contaminant concentrations that are robust to departures from normal and log-normal based assumptions. Linkage of GeoSem to the Integrated Stochastic Exposure (ISE) model (developed by SRC) will provide risk assessors with a complete geospatial/probabilistic exposure model that can be applied to risk assessments at Superfund sites. The ultimate goal is to add decision analysis tools such as Value-of- Information analysis to GeoSEM to provide geospatial software that can be used on-site or in the office to make cost effective decisions for contaminated sites.

References

U.S. Environmental Protection Agency. 1989. Risk Assessment Guidance for Superfund. Vol I. Human Health Evaluation Manual (Part A). EPA/540/1-89/002. December.

U.S. Environmental Protection Agency. 1992. Supplemental Guidance to RAGS: Estimating the Exposure Point Concentration