

# Selective Remediation of Contaminated Sites Using a Two-Level Multiphase Strategy and Geostatistics

HIROTAKA SAITO\*

*Environmental and Water Resources Engineering,  
Department of Civil and Environmental Engineering, The  
University of Michigan, Ann Arbor, Michigan 48109-2125*

PIERRE GOOVAERTS

*Biomedware, Inc., 516 North State Street,  
Ann Arbor, Michigan 48104*

Selective soil remediation aims to reduce costs by cleaning only the fraction of an exposure unit (EU) necessary to lower the average concentration below the regulatory threshold. This approach requires a prior stratification of each EU into smaller remediation units (RU) which are then selected according to various criteria. This paper presents a geostatistical framework to account for uncertainties attached to both RU and EU average concentrations in selective remediation. The selection of RUs is based on their impact on the postremediation probability for the EU average concentration to exceed the regulatory threshold, which is assessed using geostatistical stochastic simulation. Application of the technique to a set of 600 dioxin concentrations collected at Piazza Road EPA Superfund site in Missouri shows a substantial decrease in the number of RU remediated compared with single phase remediation. The lower remediation costs achieved by the new strategy are obtained to the detriment of a higher risk of false negatives, yet for this data set this risk remains below the 5% rate set by EPA region 7.

## Introduction

Soil remediation goals set by the U.S. EPA are to ensure that averaged concentrations over remediated sites are not harmful to those being exposed over a certain period of time (1, 2). Contaminant concentrations corresponding to a permissible level of risk, referred to as “risk based concentrations,” are simply calculated from equations developed by U.S. EPA (1) depending upon the type of chemicals, exposure processes, duration of exposures, or acceptable risk levels. The target concentration is then compared with an average concentration computed from data obtained within an “exposure unit (EU),” defined as a spatial area to which humans or ecological subjects are exposed. The size of the exposure unit depends on the postremediation use of the site and is typically 5000 ft<sup>2</sup> for residential lots (3, 4).

Because of the uncertainty attached to concentration estimates at a site, the U.S. EPA recommends to use as the EU average concentration the 95% one-sided upper confidence limit (UCL) of the sample mean (5). One approach (3)

is to take three grids of 50 soil samples and to form three composite samples which are then analyzed separately to yield three individual measurements per unit. The three concentrations are averaged and a *t*-test is used to compute the 95% UCL (3). If the UCL exceeds the target concentration, the topsoil of the entire exposure unit is remediated. If the UCL is below the target, no remedial action is required. The question one might ask is then “do we really need to remediate the entire unit to meet the target level?”.

Selective remediation has been introduced as an alternative to cleaning the entire unit (3, 6–8). The basic idea is to treat only the hot spots within the unit so that the pollutant level of the entire EU becomes lower than the target concentration. In selective remediation, the exposure unit is subdivided into smaller units (remediation units, RU), the optimum size of which is case-specific. The average concentration of each RU is usually required prior to remediation to decide which RUs will be remediated. Unfortunately, the concept of selective remediation has been wrongly implemented in many remediation plans. The target or risk-based concentration was frequently applied to remediation units not to exposure units, potentially leading to unnecessary remediation and costs. In selective remediation, the final assessment is conducted only at the EU level, while the treatment decision is made at the RU level. Selecting remediation units is then a crucial step in the remediation procedure, and efficient selective remediation plans rely on a correct estimation of average hazardous level of remediation units.

Geostatistics provides a set of statistical tools for analyzing and processing sparsely sampled observations leading to a probabilistic model of the spatial distribution of pollutant concentrations. Geostatistical delineation of contaminated areas usually proceeds in two steps: (1) pollutant concentrations are mapped using interpolation techniques such as kriging, and (2) a decision rule is applied to the estimated concentrations; for example the sites where the estimated concentration exceeds a regulatory threshold are classified as hazardous. Most practitioners are aware that these estimates are uncertain and that such uncertainty must be accounted for in the decision making process. Geostatistical modeling of uncertainty has thus received an increasing attention the past few years (9), and techniques fall into two main classes: nonlinear kriging and stochastic simulation.

Among nonlinear kriging techniques, indicator kriging (10) is one of the most commonly used (11, 12). The uncertainty about unknown concentrations is modeled by a conditional cumulative distribution function (ccdf), which provides the probability that any regulatory threshold value is not exceeded at an unvisited location. The technique is referred to as nonparametric because neither a particular shape nor analytical expression is assumed for the ccdf. Instead, the ccdf value (i.e. cumulative probability) is determined for a series of threshold values discretizing the range of attribute values. For each threshold, the ccdf value is estimated by kriging of indicator transforms of data. The indicator coding is generally hard in that indicators are either zero or one, which means that there is no uncertainty about whether the threshold is exceeded or not at sampled locations. Measurement errors may however be nonnegligible, in particular when cheap recording devices are used. This source of uncertainty can be incorporated using soft indicators which are valued between zero and one (13).

The model of uncertainty provided by indicator kriging applies only to the measurement support (size of soil cores) which is usually much smaller than the size of remediation

\* Corresponding author phone: (734)615-5905; fax: (734)763-2275; e-mail: hirotaka@engin.umich.edu.

units and can be assimilated to a point. Point ccdfs thus need to be aggregated or “upscaled” to derive the probability that the RU average concentration does not exceed particular thresholds (block or RU cdf). Several authors (14–17) have used stochastic simulation to perform such change of support. The idea is to simulate the spatial distribution of pollutant concentrations across the site. The simulated RU concentration is then computed as the arithmetic average of simulated point values within the RU. Many realizations are generated, and the RU cdf is numerically approximated by the empirical distribution of simulated block values.

Nowadays uncertainties attached to EU concentrations are modeled using mainly confidence intervals (5), implicitly assuming a normal distribution for concentrations and independence of observations. The main objective of this paper is to present a new geostatistically based methodology to account for uncertainty at both RU and EU levels in selective remediation. The technique is illustrated with a large soil data set that includes 600 dioxin concentrations measured in a 20 000 ft<sup>2</sup> (1858.60 m<sup>2</sup>) site.

## Theory

**Soft Indicator Kriging.** Consider the problem of modeling the uncertainty of the value of an attribute  $z$  (e.g., a dioxin concentration) at an unsampled location  $\mathbf{u}$ , where  $\mathbf{u}$  is a vector of spatial coordinates. The information available consists of  $z$ -values at  $n$  locations  $\mathbf{u}_\alpha$ ,  $z(\mathbf{u}_\alpha)$ ,  $\alpha = 1, 2, \dots, n$ . The uncertainty at  $\mathbf{u}$  is modeled by the conditional cumulative distribution function (ccdf) of the random variable (RV)  $Z(\mathbf{u})$ :

$$F(\mathbf{u}; z|n) = \text{Prob}\{Z(\mathbf{u}) \leq z|n\} \quad (1)$$

The function  $F(\cdot)$  gives the probability that the unknown is no greater than any threshold  $z$ . Ccdfs are modeled using a nonparametric (indicator) approach which estimates the probability (eq 1) for a series of  $K$  threshold values  $z_k$  discretizing the range of variation of  $z$ :

$$F(\mathbf{u}; z_k|n) = \text{Prob}\{Z(\mathbf{u}) \leq z_k|n\} \quad k = 1, \dots, K \quad (2)$$

The indicator approach requires a prior coding of each observation  $z(\mathbf{u}_\alpha)$  into a series of  $K$  values indicating whether the threshold  $z_k$  is exceeded. The error attached to a measurement at  $\mathbf{u}_\alpha$  is accounted for by replacing the single-valued datum  $z(\mathbf{u}_\alpha)$  by a normal distribution with a standard deviation  $s(\mathbf{u}_\alpha) = CV \cdot z(\mathbf{u}_\alpha)$  where  $CV$  is the coefficient of variation of the analytical error (18). To incorporate any bias in the laboratory measurement the mean of the distribution is shifted by an amount that is proportional to the observed value, that is the new mean  $z'(\mathbf{u}_\alpha)$  is  $z(\mathbf{u}_\alpha) \cdot (1 + \text{bias})$ . Thus, the  $K$  indicators are given by

$$j(\mathbf{u}_\alpha; z_k) = \text{Prob}\{Z(\mathbf{u}_\alpha) \leq z_k|n\} = G\{z_k - z'(\mathbf{u}_\alpha)/s(\mathbf{u}_\alpha)\} \quad (3)$$

where  $G(\cdot)$  is the standard normal cdf (19). The indicators  $j(\mathbf{u}_\alpha; z_k)$  have values between 0 and 1 (soft indicators), which expresses the uncertainty about whether thresholds are exceeded at sampled locations.

The ccdf value  $F(\mathbf{u}; z_k|n)$  at the unvisited  $\mathbf{u}$  is estimated by a linear combination of soft indicator transforms of neighboring  $z$ -data

$$[F(\mathbf{u}; z_k|n)]^* = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha(\mathbf{u}; z_k) j(\mathbf{u}_\alpha; z_k) \quad (4)$$

where the weights  $\lambda_\alpha(\mathbf{u}; z_k)$  are solutions of a system of  $(n(\mathbf{u})+1)$  linear equations for each threshold  $z_k$ . A technique, known as “linear interpolation between tabulated bounds”, is used to interpolate within each class ( $z_k, z_{k+1}$ ) and

extrapolate beyond extreme thresholds  $z_1$  and  $z_k$  (see ref 20, pp 284–328 for details).

**Probability Field Simulation.** The uncertainty about the average  $z$ -value over a unit  $V$  of much greater size than the soil core (i.e. RU or EU) can be quantified by the block ccdf  $F_V(\mathbf{u}; z|n)$  (21):

$$F_V(\mathbf{u}; z|n) = \text{Prob}\{Z_V(\mathbf{u}) \leq z|n\} \quad (5)$$

Despite its computational cost, stochastic simulation is increasingly preferred to multi-Gaussian block kriging and affine correction of point ccdfs for deriving block ccdf (9). The function  $F_V(\mathbf{u}; z|n)$  is numerically approximated by the empirical cumulative distribution of many simulated block values (22, p 511)

$$F_V(\mathbf{u}; z|n) \approx \frac{1}{L} \sum_{l=1}^L i_V^{(l)}(\mathbf{u}; z) \quad (6)$$

where  $i_V^{(l)}(\mathbf{u}; z) = 1$  if  $z_V^{(l)}(\mathbf{u}) \leq z$ , and zero otherwise. Each simulated block value  $z_V^{(l)}(\mathbf{u})$  is obtained by averaging a set of  $z$ -values simulated at  $J$  points  $\mathbf{u}_j'$  discretizing the block  $V(\mathbf{u})$ :

$$z_V^{(l)}(\mathbf{u}) = \frac{1}{J} \sum_{j=1}^J z^{(l)}(\mathbf{u}_j') \quad l = 1, \dots, L \quad (7)$$

The practical implementation of this approach relies on the fast generation of many simulated fields. Probability field simulation (23, 24) is probably one of the most straightforward and less CPU demanding (i.e. fastest) geostatistical simulation techniques. The basic idea is to sample the set of ccdfs using a set of autocorrelated probability values  $\{p^{(l)}(\mathbf{u}_j'), j = 1, \dots, N\}$ , known as probability or  $p$ -field, so that the histogram and semivariogram of the resulting set of simulated values  $\{z^{(l)}(\mathbf{u}_j'), j = 1, \dots, N\}$  are close to target ones (i.e. sample histogram and model fitted to sample semivariogram):

$$z^{(l)}(\mathbf{u}_j') = F^{-1}(\mathbf{u}_j'; p^{(l)}(\mathbf{u}_j')|n) \quad j = 1, \dots, N \quad (8)$$

The probability field is generated as a nonconditional realization of a random function  $P(\mathbf{u})$  with a uniform marginal cdf. The  $p$ -field does not need to be conditional since at any datum location  $\mathbf{u}_\alpha$  the ccdf is a unit-step function centered on the datum value  $z(\mathbf{u}_\alpha)$ . Thus, whatever the simulated  $p$ -field value  $p^{(l)}(\mathbf{u}_\alpha)$  at that location

$$F^{-1}(\mathbf{u}_\alpha; p^{(l)}(\mathbf{u}_\alpha)|n) = z^{(l)}(\mathbf{u}_\alpha) = z(\mathbf{u}_\alpha) \quad \forall p^{(l)}(\mathbf{u}_\alpha) \in [0, 1] \quad (9)$$

Multiple realizations over large grids can be generated reasonably fast because ccdfs need to be modeled only once, and only multiple nonconditional  $p$ -fields need to be generated. The combination of soft indicator kriging and  $p$ -field simulation was used in a previous study to illustrate the impact of measurement errors in the modeling of the block ccdf (13).

**Selection of Remediation Units.** Once the RU ccdf,  $F_V(\mathbf{u}; z|n)$ , is modeled, the decision has to be made whether the unit  $V$  needs to be remediated. Among many available techniques (16, 17), the most straightforward approach consists of classifying as contaminated all RUs where the probability of exceeding the regulatory threshold  $z_i$  is greater than a critical probability threshold  $p_i$  (25):

$$\text{RU at } \mathbf{u} \text{ is hazardous if } \text{Prob}\{Z_V(\mathbf{u}) > z_i\} = 1 - F_V(\mathbf{u}; z_i|n) > p_i \quad (10)$$

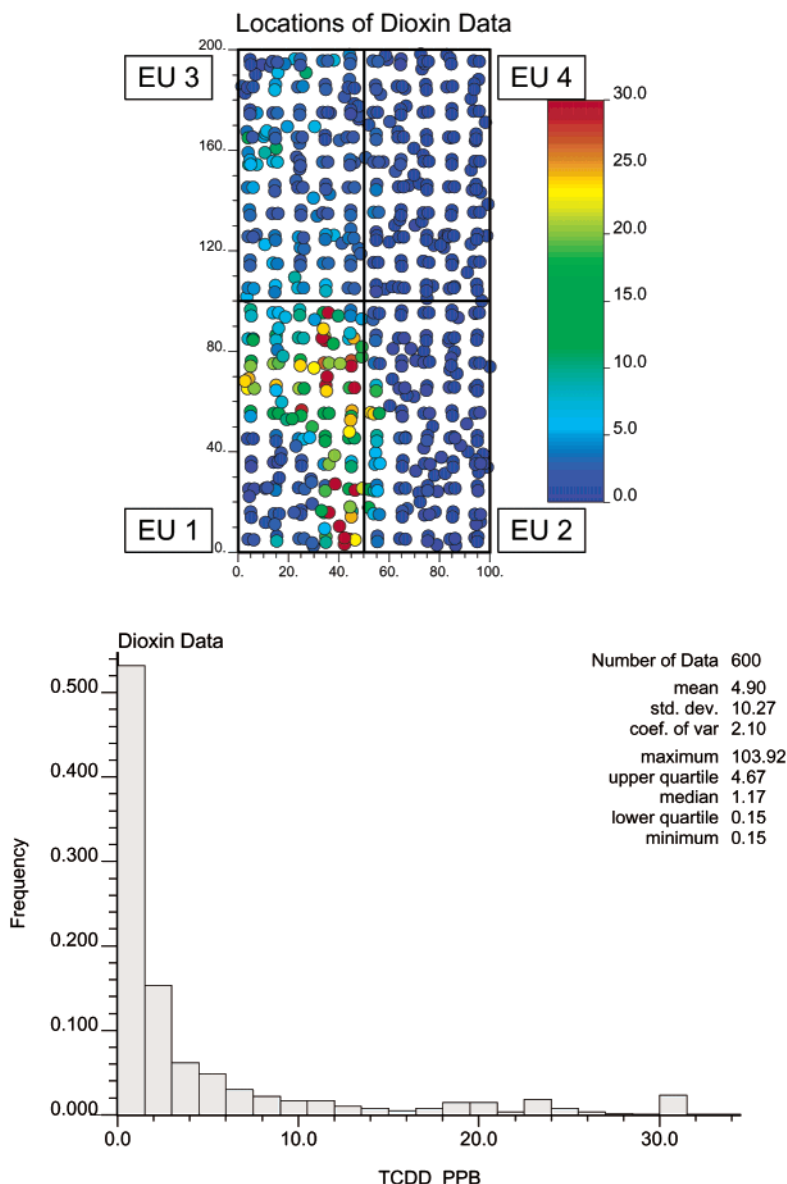


FIGURE 1. Dioxin data at Piazza Road, MO, and the corresponding sample histogram ( $\mu\text{g}/\text{kg}$ ).

An alternative approach simply declares contaminated all units with an average concentration (i.e. estimated by the mean of RU cdf  $F_v(\mathbf{u};z|(n))$ ) greater than the threshold  $z_i$ . In both approaches the decision is made only at the RU level, hence the magnitude of EU contamination is not considered in this process.

In this paper we propose to conduct the decision-making process both at RU and EU levels. The basic idea is to treat only highly contaminated RU within the EU until the probability that the average EU concentration does not exceed the threshold  $z_i$  is above the probability threshold  $p_i$ . Although the remediation decision is made at the RU level, the final assessment is conducted only at the EU level. There are a few techniques available for the critical step of selecting RUs. For example, all locations with concentrations greater than an analytically derived “not to exceed” concentration are remediated to bring the average concentration of the site below the target concentration (7, 26). Another approach ranks RU according to their average concentrations and units are cleaned until the EU average concentration is below the cleanup target (8). A common shortcoming of these techniques is that the uncertainty attached to EU concentration is not accounted for in the selection of RU. Modeling simultaneously the ccdfs of EU and RU concentrations is

proposed as a way to reduce both the remediation cost and the risk of wrongly selecting RU. The new approach proceeds as follows:

1. Multiple realizations of the spatial distribution of pollutant concentrations over the site are generated using Monte Carlo simulation ( $p$ -field simulation). Simulated values falling within an EU are arithmetically averaged and the empirical distribution of simulated block values forms the EU cdf.
2. The probability of not exceeding the target concentration is computed from that cdf and used as the pre-remediation probability ( $p_{pre}$ ).
3. For each exposure unit, the postremediation cdf resulting from cleaning a single RU is derived by replacing the average simulated concentration of that RU by zero (mimicking the actual remediation process) for all realizations. The postremediation probability ( $p_{post}$ ) for cleaning that RU alone is then computed from the cdf. The process is repeated for each RU.
4. The difference between postremediation and pre-remediation probabilities is computed ( $p_{post} - p_{pre}$ ), and the remediation unit that increases the probability most is ranked first and is considered to have the greatest impact or influence on the EU cdf.

5. Starting with the most influential RU, units are remediated until the exposure unit meets a given cleanup goal, that is the probability of not exceeding the target threshold is greater than a specific probability  $p_c$ .

The selection of RUs is based only on the impact on the EU cdf, which implies that as long as the level of contamination of the EU after remediation is lower than the target the remediation is considered a success. If the postremediation exposure unit does not meet the goal, the remediation process is categorized as false negative (27). Conversely, the cleaning of RUs that are not necessary to make the EU safe (case when EU concentration is already small enough) is referred to as false positive.

## Materials and Methods

**Study Site.** The study area is a U.S. EPA Superfund site called Piazza Road located in Rosati, MO. In 1971, dioxin was introduced along with a waste oil serving as a dust suppressant; later the site was found to be contaminated (6, 15, 27, 28). Policy in EPA Region 7 has set the acceptable risk due to residual dioxin at less than one additional cancer in 1 000 000 population (e.g.  $10^{-6}$  carcinogenic risk), which corresponds to a concentration of  $1 \mu\text{g}/\text{kg}$  in a residential exposure scenario. The size of the study site is  $100 \text{ ft} \times 200 \text{ ft}$  (Figure 1), consisting of four  $50 \text{ ft} \times 100 \text{ ft}$  exposure units (average residential lot). In this paper, remediation units are  $10 \text{ ft} \times 10 \text{ ft}$ , yielding 50 RUs per EU. The RU size can be determined in many different ways depending upon, for example, the size of machinery in the remediation process: a well-trained operator of a front end loader can take 6 in. off a area of  $10 \text{ ft} \times 10 \text{ ft}$  in a single pass, resulting in a volume of slightly less than 2 cubic yards of material which is convenient for incineration or disposal activity.

Two sets of samples were collected: (1) 200 pairs of tablespoon soil samples and (2) 50 composite samples formed at 50 randomly selected locations in each of the four exposure units: at each location, nine tablespoon samples were taken from a 4-in. square grid (details of sampling procedure are summarized in ref 27). The sole isomer of dioxin at the site is 2,3,7,8-tetrachlorodibenzo-p-dioxin. The laboratory technical procedure (HRGC/MS/MS or HRGC/LRMS) has a high analytical precision (4.5%), a low bias (-2.6%), and a detection limit of  $0.3 \mu\text{g}/\text{kg}$ . Readings below the detection limit are reset to half the limit,  $0.15 \mu\text{g}/\text{kg}$ , in accordance with previous studies (29). For the purpose of this study, measurement errors include laboratory (analytical) errors and errors resulting from incomplete mixing of the field samples. The coefficient of variation (CV) for sampling errors ranges from 11% for composite samples to 15% for single spoon samples which were not homogenized, and these two rates were used for the soft indicator coding (eq 3) of the 200 and 400 concentration data, respectively.

Figure 1 shows the location and histogram of 600 dioxin data. The distribution is highly positively skewed (median =  $1.17 \mu\text{g}/\text{kg}$  < mean =  $4.94 \mu\text{g}/\text{kg}$ ). Of the observations 27% are below the detection limit, while the threshold of  $1 \mu\text{g}/\text{kg}$  is exceeded by 53% of the samples. High concentrations are located mainly in the left lower area of the site due to the accumulation of dioxin by water erosion (27). A previous study (30) showed larger spatial connectivity along the N-S direction which is consistent with the direction of the stream channel located in this plot.

**Selective Remediation Performance.** Selective remediation amounts at subdividing each EU into smaller RUs and, starting from the most contaminated RU, RUs are cleaned until the remediation goal is met. The proposed two-level multiphase remediation technique is compared to two single phase remediation approaches based on either the probability for RUs not to exceed the threshold or the average RU dioxin concentrations (ccdf mean). For the multiphase

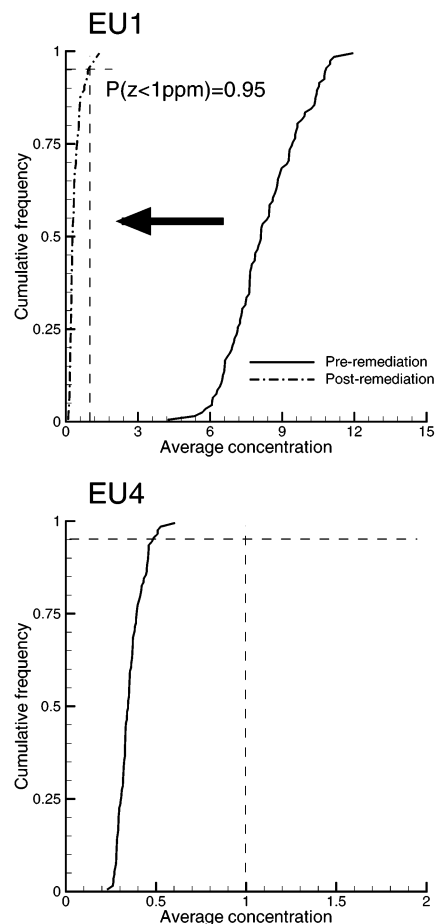


FIGURE 2. Preremediation ccdfs for EUs 1 and 4 which are numerically approximated by the empirical distributions of 100 simulated block values. The probability of not exceeding the threshold ( $1 \mu\text{g}/\text{kg}$ ) is above 95% (horizontal dashed line) for EU 4 which entails no remediation needed. For EU 1 selective remediation is conducted until the postremediation probability exceeds 95%.

strategy, the remediation goal is that the probability of exceeding a  $1 \mu\text{g}/\text{kg}$  dioxin concentration is less than 5% for the EU, which corresponds to the EPA remediation goal. For the single phase remediation techniques, RUs with a greater than 5% probability of exceeding  $1 \mu\text{g}/\text{kg}$  or an average dioxin concentration greater than  $1 \mu\text{g}/\text{kg}$  are selected. To investigate the impact of sampling intensity on remediation decision and to show the benefit of the new methodology, the following procedure is applied to the three different remediation strategies:

1. Select a sample size corresponding to a percentage  $S$  of observations with  $10\% \leq S \leq 90\%$ . For each size, 100 different random subsets were selected to account for sampling fluctuations.

2. For each random subset and sample size: (a) Model the ccdf at the nodes of a 5 ft spacing grid using soft indicator kriging and eight threshold values corresponding to the detection limit and the 7 deciles of the sample distribution that are larger than  $0.3 \mu\text{g}/\text{kg}$ . Generate 100 realizations of the spatial distribution of dioxin concentrations using  $p$ -field simulation. (b) Compute the arithmetical average of simulated values for each exposure and remediation units and construct the preremediation EU ccdf (see Figure 2). (c) Select RUs according to the three remediation strategies and record the number of RU selected for each strategy. (d) Compute the postremediation "true" EU concentration by replacing "true" dioxin concentrations of selected RUs by zero. The "true" concentration is taken as a weighted average of all

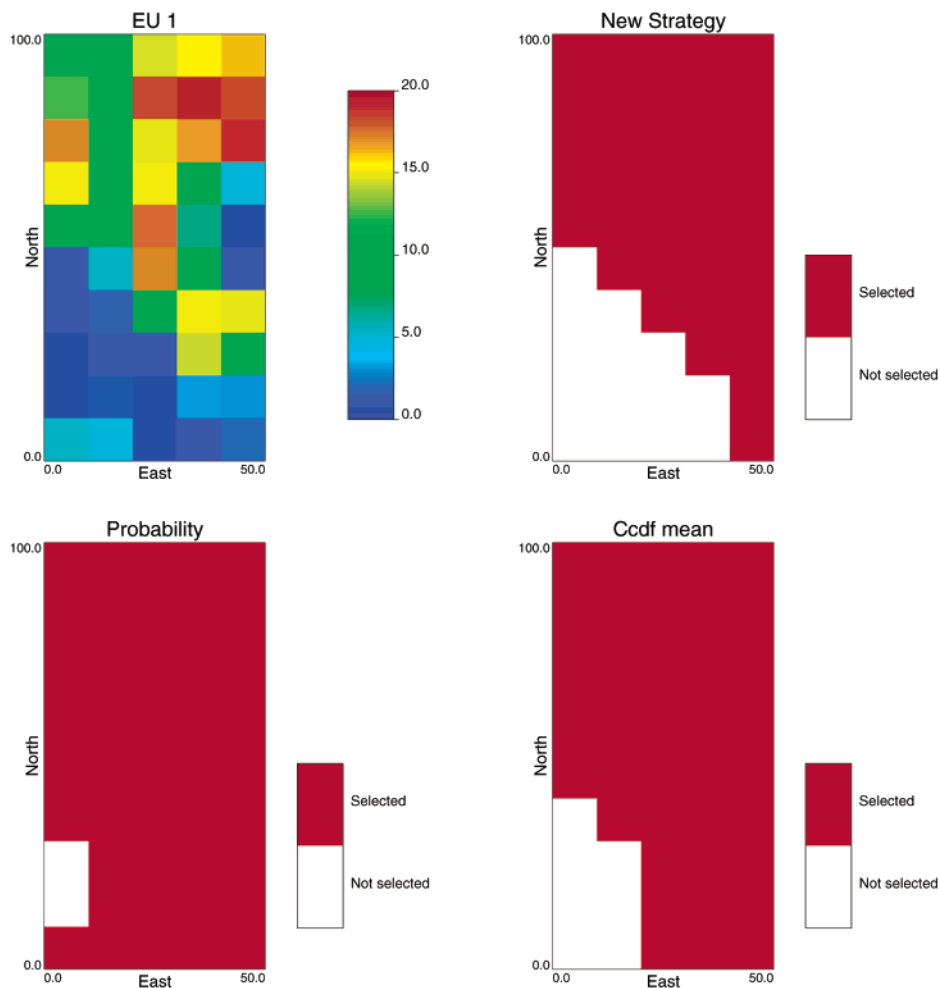


FIGURE 3. First realization of the spatial distribution of dioxin concentration within EU 1 generated using  $p$ -field simulation conditionally to 50 randomly selected dioxin data. 36 RUs are selected as a result of the new selective remediation procedure, while 48 and 43 RUs are selected under the two single phase strategies.

observations within that unit: paired observations received half the weight of randomly located composite samples. If the postremediation concentration is greater than the threshold, the remediation is categorized as failure (false negative). (e) Compute the proportion of selected RUs whose true concentration actually exceeds the threshold to investigate the efficiency of the three remediation strategies.

## Results and Discussion

Figure 2 shows EU 1 and 4 cdfs which were derived by  $p$ -field simulation using 50 randomly selected observations. Unlike EU 4, the probability of exceeding the threshold ( $1 \mu\text{g}/\text{kg}$ ) for EU 1 is greater than the target probability of 5%, which leads to remediation. Using the multiphase strategy, the 5% threshold is reached after 36 RUs have been cleaned, see the postremediation cdf in Figure 2 (top graph, dashed line). Figure 3 (top left graph) shows the first realization of RU values with the ones selected according to the new selective remediation strategy (top right graph) and two single phase remediation strategies (bottom graphs). This example already indicates the tendency of the proposed strategy to reduce the number of RUs selected for remediation. This result is confirmed by Figure 4 that shows the average number of RU selected over 100 random subsets as a function of the relative sampling intensity. The 95% probability intervals (both upper and lower bounds) of the distribution of 100 values are also plotted to illustrate the impact of sampling fluctuations. The thick horizontal line depicts the optimum number of RU that should be selected as obtained from the

original data set: RUs are ranked according to the weighted average concentration of all observation within these units and are removed until the EU dioxin concentration is below  $1 \mu\text{g}/\text{kg}$ . The number of selected RUs generally decreases as the sampling intensity increases. This is not the case for EU 3 because of the narrow range of concentrations in that unit (see Figure 1), which facilitates its modeling using fewer observations. Overall, the proposed method outperforms (i.e. lower remediation costs) the two single phase remediation strategies except when few data are sampled in EU 2. Unlike EU 3, there are a few extreme values along the left edge of this unit (Figure 1), which would have a large impact on the remediation plan especially when only a few data are available for site characterization. When either of single phase strategies is adopted, the number of RU selected is far more than necessary leading to false positive in almost all cases (i.e. the postremediation concentration is below the regulatory threshold). This overselection of RUs results in additional cleaning costs which usually increase as a linear function of the number of units selected. Despite the lack of regulation regarding the proportion of false positives, the budget for the remediation project is always limited and the lower the cost, the better the design. This unnecessarily large remediation effort can be easily lowered by the new remediation strategy.

The false negatives generated by each remediation strategy are investigated by computing for each EU the proportion of random subsets where the post remediation “true” concentration exceeds the regulatory threshold. In all cases

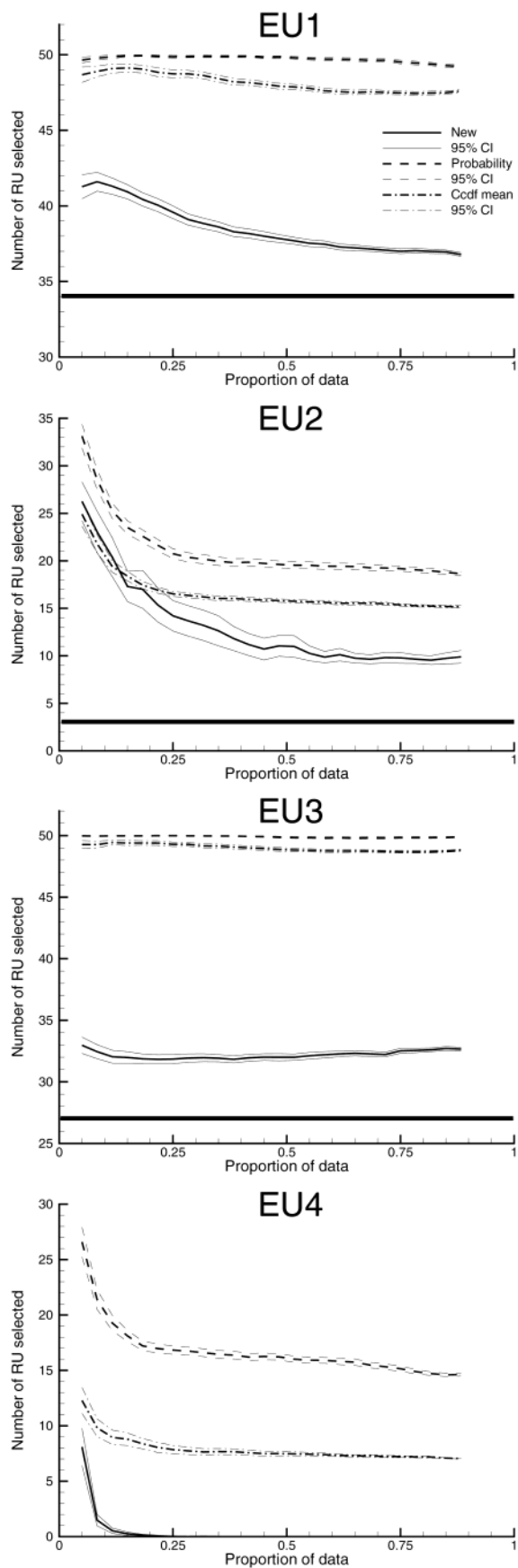


FIGURE 4. Impact of the sample size on the average number of RU selected for EUs 1 to 4 by the three remediation strategies. Both lower and upper bounds of the 95% probability intervals are also plotted.

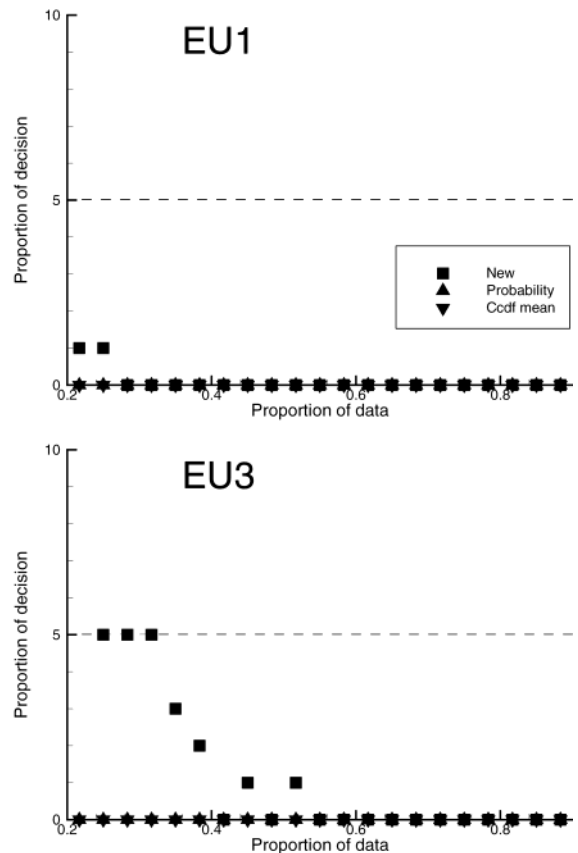


FIGURE 5. Impact of sampling intensity and remediation strategy on the proportion of false negatives for EUs 1 and 3.

there is no false negatives for EU 2 and 4 which contain the lowest pollutant concentrations, recall Figure 1. Figure 5 displays the proportion of false negatives for the two other EUs. The horizontal line corresponds to an acceptable false negative rate of 5%, which was set by EPA Region 7 (13). As expected, the lower remediation costs achieved by the new strategy are obtained to the detriment of a higher risk of false negative. Yet, for EU 1, the proportion of false negatives is below 5% for most sampling intensities. For EU 3, as long as more than 20% of the original data are used (i.e. 120 observations), the proportion of false negatives generated by the new method is acceptable.

A risk associated with selective remediation is that, because of the smaller number of RU selected, one may fail to meet the remediation goal (false negative). Results however demonstrate the overall superiority of the new technique, especially its ability to decrease the remediation cost while keeping the proportion of false negatives acceptable. The efficiency of the technique can also be assessed by calculating the proportion of selected RU whose true concentrations actually exceed the threshold, which indicates how well the technique pinpoints the units to be cleaned. The proportion for EU 1 is plotted as a function of sample size in Figure 6. Clearly, the new technique selects RUs more effectively than the other two methods, and the same trend is observed for other exposure units.

The technique presented in this paper indeed minimizes the remediation cost, while keeping the risk of leaving the site hazardous (i.e. false negative) below the regulatory target (e.g. 5%). Regardless of the technique used, the risk is never drastically reduced as long as only a limited number of samples is available, since estimation results are conditioned to samples. The detection of all the hot spots is ensured only by remediating or sampling the entire site. However, in practice, time and cost always limit the number of samples

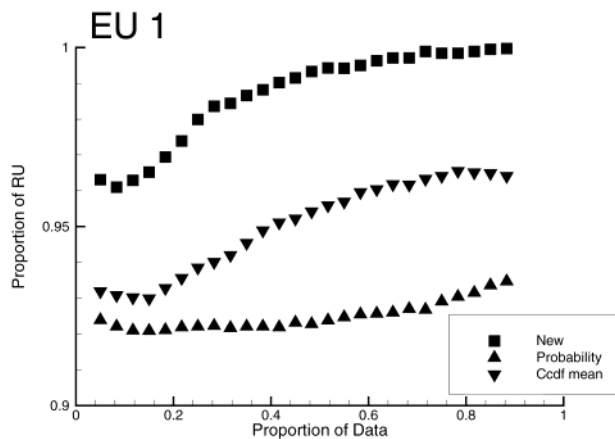


FIGURE 6. Proportions of RUs that actually exceed the target concentration ( $1 \mu\text{g}/\text{kg}$ ) among units selected within EU 1 as a function of the sample size and remediation strategy.

or the area being remediated. Minimizing the risk of missing any hot spot could be achieved by a multiphase sampling scheme (i.e. Bayesian updating approach), in which additional samples are collected based upon the results from initially sampled observations. The combination of the multiphase selective remediation strategy and the multiphase sampling design would be the best approach to minimize the false negative in soil remediation designs.

The last concern one might have is the potential existence of small areas of high dioxin concentration (i.e. hot spots), while the average hazardous level of the EU is lower than the regulatory threshold. Then one needs to reexamine how the remediation goal has been set. In general, long-term exposure is of concern (31), especially when the site is used as a residential lot, and the proposed method is designed to deal with such conditions. When the focus is on short-term exposures or in the presence of other site uses or exposure pathways (e.g. a child's sand box set on the ground), alternatives in terms of EU size and remediation strategy should be considered. In such a case, simply drawing contours (i.e. delineating areas, eq 10) to identify locations where a contaminant concentration or a risk level is above a given threshold (32, 33) may be an appropriate approach. The size of EU/RU in this direct hot-spot detection method can be anything but not less than the size of soil sampling cores.

#### Literature Cited

- (1) U.S. EPA Office of Research and Development. *Risk Assessment Guidance for Superfund*; Volume 1: Human Health Evaluation Manual, Part B, Development of Risk-based Preliminary Remediation Goals, EPA/540/R-92/003; U.S. EPA: Washington, DC, 1991.
- (2) Blacker, S.; Goodman, D. *Environ. Sci. Technol.* **1994**, *28*, 466A–470A.

- (3) Blacker, S.; Goodman, D. *Environ. Sci. Technol.* **1994**, *28*, 471A–477A.
- (4) Myers, J. C. *Geostatistical Error Management*; Van Nostrand Reinhold: New York, 1997.
- (5) U.S. EPA Office of Emergency and Remedial Response. *Supplemental Guidance to RAGS: Calculating the Concentration Term*; Intermittent Bulletin, Vol. 1, No. 1, May 1992.
- (6) Rytty, R. T.; Neptune, D.; Groskinsky, B. *Environ. Test. Anal.* **1992**, *1*, 26–31.
- (7) Bowers, T. S.; Shifrin, N. S.; Murphy, B. L. *Environ. Sci. Technol.* **1996**, *30*, 1437–1444.
- (8) Brakewood, L. H.; Grasso, D. *Environ. Sci. Technol.* **2000**, *34*, 3837–3842.
- (9) Goovaerts, P. *Geoderma* **2001**, *103*, 3–26.
- (10) Journel, A. G. *Math. Geol.* **1983**, *15*, 445–468.
- (11) Ginevan, M. E.; Splitstone, D. E. *Environ. Sci. Technol.* **1997**, *31*, 92A–96A.
- (12) Juang, K. W.; Lee, D. Y. *Environ. Sci. Technol.* **1998**, *32*, 2487–2493.
- (13) Saito, H.; Goovaerts, P. *Environmetrics* **2002**, *13*, 555–567.
- (14) Journel, A. G. In *Geostatistics Tróia '92*; Soares, A., Ed.; Kluwer Academic Publishers: Dordrecht, 1992; pp 213–224.
- (15) Englund, E. J.; Heravi, N. *Environ. Ecol. Stat.* **1994**, *1*, 247–263.
- (16) Goovaerts, P.; Webster, R.; Dubois, J. P. *Environ. Ecol. Stat.* **1997**, *4*, 31–48.
- (17) Kyriakidis, P. C. In *Geostatistics Wollongong '96*; Baafi, E. Y., Schofield, N. A., Eds.; Kluwer Academic Publishers: Dordrecht, 1997; pp 973–983.
- (18) Journel, A. G. *Math. Geol.* **1986**, *18*, 269–286.
- (19) Van Meirvenne, M.; Goovaerts, P. *Geoderma* **2000**, *102*, 75–100.
- (20) Goovaerts, P. *Geostatistics for Natural Resources Evaluation*; Oxford University Press: New York, 1997.
- (21) Lahiri, S. N.; Kaiser, M. S.; Cressie, N.; Hsu, N. J. *J. Am. Stat. Ass.* **1999**, *94*(445), 86–97.
- (22) Journel, A. G.; Huijbregts, C. J. *Mining Geostatistics*; Academic Press: New York, 1978.
- (23) Srivastava, R. M. In *SPE Annual Conference and Exhibition, Washington D.C.*; No. 24753; Soc. Petroleum Eng.: Washington, DC, 1992; pp 927–938.
- (24) Froidevaux, R. In *Geostatistics Tróia '92*; Soares, A., Ed.; Kluwer Academic Publisher: Dordrecht, 1993; pp 73–84.
- (25) De Oliveira, V.; Ecker, M. D. *Environmetrics* **2002**, *13*, 85–101.
- (26) Schultz, T. W.; Griffin, S. *Risk Anal.* **2001**, *21*, 43–52.
- (27) Rytty, R. T. *J. Air Waste Manage. Assoc.* **1993**, *43*, 197–202.
- (28) Abt, M.; Welch, W. J.; Sacks, J. *Math. Geol.* **1999**, *31*, 1–22.
- (29) Zirschky, J.; Keary, G. P.; Gilbert, R. O.; Middlebrooks, E. J. *J. Environ. Eng.* **1985**, *111*, 777–789.
- (30) Saito, H.; Goovaerts, P. *Environ. Sci. Technol.* **2000**, *34*, 4228–4235.
- (31) U.S. EPA Office of Research and Development. *Risk Assessment Guidance for Superfund*; Volume 1: Human Health Evaluation Manual, Part A, EPA/540/1-89/002; U.S. EPA: Washington, DC, 1989.
- (32) Qian, S. S. *Environ. Ecol. Stat.* **1997**, *4*, 1–29.
- (33) Diggle, P. J.; Tawn, J. A.; Moyeed, R. A. *Appl. Statist.* **1998**, *47*, 299–350.

Received for review May 15, 2002. Revised manuscript received February 6, 2003. Accepted February 11, 2003.

ES020737J