*Accepted for presentation at the Fifth International Geostatistics Congress,
22–27 September 1996, University of Wollongong, Wollongong, NSW, Australia
(http://www.uow.edu.au/geostat96/)*

GEOSTATISTICS AND COST-EFFECTIVE ENVIRONMENTAL REMEDIATION

Christopher A. Rautman
Sandia National Laboratories
P. O. Box 5800; MS-1324
Albuquerque, New Mexico 87185
USA

1. Introduction

Numerous sites within the U. S. Department of Energy (DOE) complex have been contaminated with various radioactive and hazardous materials by defense-related activities during the post-World War II era. A common perception is that characterization and clean-up of these contaminated sites will be too costly using currently available technology. Consequently, the DOE Office of Technology Development has funded development of a number of alternative processes for characterizing and remediating these sites. The former Feed-Materials Processing Center near Fernald, Ohio (USA), was selected for demonstrating several innovative technologies. Contamination at the Fernald site consists principally of particulate uranium and derivative compounds in surficial soil.

A program was conducted during 1994 specifically to demonstrate the relative performance of seven advanced-characterization tools for measuring uranium activity of in-situ soils. These innovative measurement technologies are principally radiation detectors of different designs (table 1). Four industry-standard characterization technologies, including conventional, regulatory-agency-accepted soil sampling followed by laboratory geochemical analysis, were also demonstrated for comparison.

A risk-based economic-decision model has been used to evaluate the performance of these several characterization tools. The decision model computes the value of an objective function for each characterization approach. The objective function is defined as the total cost to remediate the site, including the costs of characterization and treatment plus the expected cost of failure to clean the entire site to regulatory standards. Geostatistical simulation is used to quantify the likelihood of such failure. The preferred characterization method is the one that minimizes the value of the objective function.

Although this application of the methodology at the Fernald site involves a number of site- and situation-specific considerations, the cost-risk-benefit decision framework is completely general. The methodology can assist site operators to choose among engineering alternatives for site characterization and remediation, and can also provide an objective and quantitative basis for decisions with respect to the completeness of site characterization. The quantitative uncertainty framework provided by geostatistics may provide a basis for more productive and focused discussions with regulatory agencies and other stakeholders in the environmental remediation arena.

2. Economic Cost-Risk-Benefit Decision Framework

The decision model employed at the Fernald demonstration program builds upon a comprehensive logical framework for economic-decision analysis outlined by Freeze *et al.* (1990). The framework attempts to quantify the various costs, risks, and benefits associated with the design and implementation of a given engineering decision. What distinguishes the approach of Freeze *et al.* from many other cost-benefit decision frameworks is the quantitative emphasis placed on *economic risk:* the likely costs that result from uncertainty regarding the ultimate performance of the system being evaluated. The entire analysis is denominated in currency units (dollars or otherwise), thus reducing all components of the analysis to a common reference familiar to business decision-makers. The goal of the economic-decision analysis is to maximize an objective function that represents the overall monetary return associated with a given project. The objective function is computed as the net present value of all revenues or benefits resulting from the project, less all relevant expenses or costs including the expected cost of uncertain performance, for a reasonable number of engineering alternatives.

TABLE 1: Classification, brief description, and sampling costs of alternative characterization technologies demonstrated at the Fernald site. [All costs in U.S. dollars]

ID	Technology Description	Detection Principle		Per-Sample Cost	Total Cost (this study)
Advanced Field-Measurement Technologies					
ATD	Alpha-track detector	passive	alpha particles	87	10,131
Beta	Beta scintillometer	passive	beta particles	167	15,391
EIC	Electret ionization chamber	passive	alpha particles	92	9,337
GMH	High-mount gamma spectrometry	passive	gamma rays	155	13,183
GML	Low-mount gamma spectrometry	passive	gamma rays	155	13,669
ICP	Laser-ablation inductively coupled plasma atomic emission spectrometry	active	visible and ultraviolet light	272	22,540
LRAD	Long-range alpha detection	passive	alpha particles	162	15,336
Industry-Standard Field-Screening Techniques					
FID	FIDLER scintillometer	passive	gamma rays	50	4,771
Lab	Mass spectroscopy	active	ionized elements	340	35,055
NAD	Sodium-iodide scintillometer	passive	gamma rays	42	3,709
XRF	X-ray fluorescence detector	active	photons	99	9,588

In the case of environmental remediation activities, there generally are no particular economic benefits to the site operator in the sense originally implied by Freeze *et al.* For example, there are no dollar inflows to the DOE that result directly from cleaning up the Fernald site. This functionally transforms the profit-maximization framework into a total-cost minimization exercise, which can be stated mathematically as follows:

$$\text{Minimize:} \quad \Phi_i = Ctotal_i = Cchar_i + Ctreat_i + E\{Cfail_i\} \qquad (1)$$

across the suite of I different characterization methods. In equation (1), $Ctotal_i$ is the total cost associated with characterizing and remediating the site using characterization technology i. This cost comprises the cost of characterization ($Cchar_i$), the cost of treatment ($Ctreat_i$), and the expected cost of failure ($E\{Cfail_i\}$), which is defined as:

$$E\{Cfail_i\} = Pfail_i \cdot Cfail_i \ . \qquad (2)$$

Here, $Pfail_i$ is the probability of failure associated with use of technology i, and $Cfail_i$ is the cost incurred if that failure actually takes place. Uncertainty and economic risk enter the decision model through this probability-of-failure term. One minimizes the economic cost of a project by trading-off these different cost components against one another, and potentially, by trading costs (including risk costs) against variable benefits.

3. Application to the Fernald 1994 Demonstration Program

The 1994 Fernald demonstration program provided an almost ideal application of this form of economic-decision analysis. The problem is to choose a technique for measuring uranium contamination in soil from among a number of alternative characterization technologies. Some technologies may be relatively inexpensive to operate, but the readings may be somewhat inaccurate. Other technologies may be more costly, but provide highly accurate and precise measurements. A more conservative measurement technique, one which indicates more contamination than is actually present, may reduce the likelihood that the remediated site will fail to meet regulatory inspections, but at the increased cost of removing and treating soil that could have been left in place. Conversely, a less conservative measurement technique may suggest a lower initial remediation cost, but leave the operator exposed to a large potential liability if residual contamination is detected later.

In addition to the information from several innovative and field-screening technologies, measurements of uranium contamination were also obtained using field and laboratory analytical methods currently accepted by the cognizant regulatory agencies. Thus, based on the assumption that it is the regulator's perception of what is and what is not contaminated that determines the success or failure of a remediation effort, a reasonable model of ground truth was available as a benchmark against which to compare the other technologies that have been proposed as *alternatives* to current practices.

3.1. SITE DESCRIPTION AND THE DEMONSTRATION PROGRAM

The 1994 Fernald demonstration program was conducted at a site known as the incinerator area. A simple combustion unit operated at this site from 1954 until 1979, burning combustible wastes generated from both plant administrative and process areas. Some of the process wastes contained low levels of radioactive materials, principally uranium, and particulate contamination was dispersed across nearby areas as an aerosol plume. Uranium activities vary from background of 5–10 picocuries per gram (pCi/g) to more than several thousand pCi/g (DOE, 1992). Other site characterization activities at the Fernald site have indicated that the bulk of the uranium above likely action levels is present in the top 2 inches or so of the soil (Schilk *et al.*, 1993; Rautman *et al.*, 1995).

The 1994 demonstration program focused on a 2.9-acre portion of the incinerator area. A more-or-less regular grid consisting of a total of 85 sample locations on nested

60-, 30-, and 5-ft spacings was defined and marked in the field. All alternative technologies recorded measurements at these marked locations. Each technology was calibrated to a common set of native Fernald soils spiked with known levels of uranium; calibration beds also were sampled and analyzed by laboratory mass spectroscopy techniques.

3.2. GEOSTATISTICS AND IDENTIFICATION OF “CONTAMINATED” REGIONS

Excavation of contaminated soil probably will take place using mechanical equipment, thus limiting the selectivity of excavation units. A selective remediation unit is assumed to consist of a square panel 10-ft by 10-ft in size (approximately the width of a bulldozer blade) and 4 inches deep. This convention is consistent with previous geostatistical modeling of uranium contamination at Fernald (Rautman *et al.*, 1994). There are 2016 such potential remediation units at the demonstration site. The vast majority of these remediation units are not represented directly by measured sample values and the uranium activity of these parcels must be predicted using geostatistical (or other) methods.

We have used conventional variography and sequential gaussian simulation (Deutsch and Journel, 1992) conditioned on the relevant characterization data to model the exhaustive spatial distribution of uranium. Each different set of measurements exhibits its own statistical character and its own spatial continuity structure, thus requiring eleven different variograms. The approach results in a suite of equally likely maps of contaminant distribution for each characterization technique. Each map reflects the measured values at their spatial locations and exhibits essentially the same statistical and spatial character as the data. This stochastic approach explicitly acknowledges that both the models of contaminant distribution derived from the alternative characterization technologies and the model of true contamination derived from soil geochemistry are uncertain. There is no *a priori* means of determining which of the many possible simulated models is the “true” map of uranium contamination. It remains to evaluate the remediation consequences of the uncertainty.

A large number of simulations that differ only in their initial random-number seed are post-processed by determining for each grid location the number of simulations that exceed a specified uranium activity, Z^*. By converting this number exceeding Z^* to a proportion and by assuming that this proportion is a reasonable empirical approximation of the probability that a particular location actually exceeds Z^*, it is possible to create “probability maps” showing the likelihood of exceeding the action level at any location. The probability maps are transformed into remediation (excavation) maps by selecting all parcels that have a probability of exceeding Z^* greater than some acceptable value, p. The remediation map for each technology serves as input for determining the cost of treatment, $Ctreat_i$, in the objective function of equation (1).

3.3. ECONOMIC COSTS

3.3.1. *Characterization Costs*

Actual costs for the characterization part of the 1994 Fernald demonstration have been summarized in table 1 (Douthat *et al.*, 1995); costs vary by an order of magnitude among the several technologies. The industry-standard field screening methods generally were significantly less expensive than the proposed innovative characterization technologies. Cost of completing the field survey using the latter ranged from US$9,300 to more than $22,500. The currently accepted characterization method for regulatory purposes, consisting of soil sampling followed by conventional laboratory geochemical analysis, was, in fact, the most expensive method, costing in excess of $35,000.

3.3.2. *Treatment Costs*

The cost of treating Fernald soils is uncertain because a remediation technology had not yet been selected at the time of the field demonstration. However, several proposed treatment methods involve roughly comparable costs of approximately $200 per ton of soil processed. Adjusting this cost per ton for the volume and bulk density of a selective remediation unit results in a dollar cost of $437 per contaminated panel.

3.3.3. *Failure Costs*

In the case of environmental remediation, "failure" essentially results from a misclassification in predicting actual contaminant levels; figure 1 illustrates this problem. The true contaminant concentration (Z) of any given parcel of land is assumed to be unknown. The operator conducts site characterization and predicts the contamination level for all unsampled locations. Depending upon whether the predicted concentration, $\hat{Z}$, is above or below a regulatory threshold or action level, Z^*, the operator will decide to remediate that parcel or to leave the soil in place. Note that absolute accuracy of the predicted concentration does not matter particularly. The decision to remediate is based solely upon whether the predicted contaminant level for the parcel is above (contaminated) or below (not contaminated) the regulatory action level. It is the remediation *decision* that matters, not the accuracy of the prediction itself.

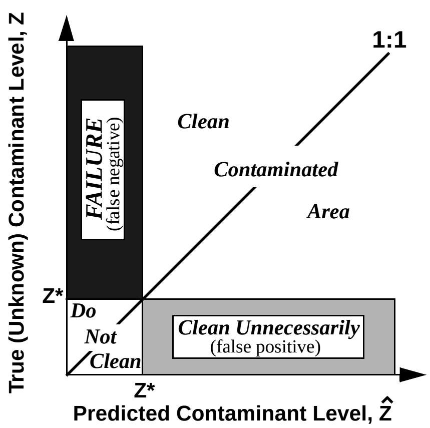

Figure 1. Classification problem for a contaminated site. Decision to remediate or leave as-is will be based upon a predicted contaminant level.

TABLE 2: Contingency table showing alternative costs of failure. [all values in US dollars]

	Small-Value† Failure	Large-Value Failure
Small No.‡ of failures	$a-437/parcel	$b=2,000/ parcel
Large No. of failures	$c=35,000	$d=500,000

†distinction set at twice the standard deviation of the soil geochemistry/laboratory analyses

‡distinction set at 30 percent of the 2016 parcels covering the demonstration site

After the site operator deems remediation complete, we assume that the regulatory agency examines some or all of the site area and makes its own assessment of whether or not the various remaining parcels are below the regulatory threshold. Although the "true" concentration for a given panel may still be unknown (even unknowable) to the regulatory agency, we assume that the agency's determination is what matters in assessing success or failure. If the assessment is that the parcel is still above threshold—i.e., that the parcel has been misclassified (a false negative; figure 1)—we count the site characterization-remediation effort for that parcel a failure.

Note that not all classification errors count as failures under this scenario. If the site operator treats a parcel that is not above threshold, there has been a classification error. However, this false positive (figure 1) does not constitute *regulatory* failure because the parcel is presumed still "clean" following treatment. There are economic costs to this second type of classification error; however, the cost of failure is not one of them.

Actual costs of failure for the Fernald site are more problematical than costs involved in either site characterization or remediation. We have defined a cost of failure, *Cfail*, that varies according to the degree of that failure, as given in table 2. A "small-value" failure is one in which the contaminant level of a given "uncontaminated" parcel exceeds the regulatory threshold by only a limited amount. Conversely, a "large-value" failure is one in which the prediction from site characterization was below threshold when, in fact, the actual contaminant level exceeded that threshold by a large amount.

The values of ***$a***, ***$b***, ***$c***, and ***$d*** are arbitrary, but they follow the logic of a small penalty (***$a***) for small errors, and a large penalty (***$b***) for large errors that could be construed as "flagrant." This approach is consistent with the power-curve concept related to the data-quality-objective process published by the U.S. Environmental Protection Agency (EPA, 1993). If there are so many small errors as to constitute "negligence," an additional one-time cost of failure (***$c***) is assessed as a penalty. A complete breakdown of a characterization program that results in an unacceptably large number of large-value failures is penalized by a very large cost of failure (***$d***).

We have elected to set ***$a*** = *Ctreat*, in keeping with a philosophy that a few small errors do not constitute systematic failure of the characterization effort. Because this failure cost, ***$a,*** is intended merely to include the cost of follow-up treatment, it is necessary to accrue this cost for every parcel that fails. Similar logic applies to failure cost ***$b***, only in this case the cost is greater to cover expenses associated with additional sampling and paperwork intended to assure the regulatory body that such large-value errors are relatively isolated occurrences. Failure costs ***$c*** or ***$d*** are assumed to be incurred only once.

3.4. STOCHASTIC EVALUATION OF THE OBJECTIVE FUNCTION

The objective function [equation (1)] was evaluated using the logic presented in figure 2. Uncertainty in site characterization was incorporated probabilistically into the analysis through design of the remediation plans based on probability mapping as described in section 3.2. Uncertainty in the true distribution of uranium contamination is incorporated here through evaluation of those remediation plans against each of the multiple stochastic simulations based on the EPA-accepted laboratory measurements. There are several important features about this formulation of the stochastic approach that follow. (1) The loop structure focuses on the uncertainty and the expected cost associated with each remediation panel, $\mathbf{x}$. Thus, although the comparison of Φ_i is technology by technology, the results allow direct spatial mapping of the cost in dollars for any technology, i. This technique may be effective for communicating comparisons, particularly to lay audiences. (2) It is possible to iterate over several probability levels and to plot $Ctotal_i$ as a function of p to identify any prominent break in slope (change in behavior). (3) It is also possible to iterate over several remediation threshold values, Z*, to demonstrate changes in $Ctotal_i$ or differences in technology rankings as a function of clean-up threshold.

Create j simulated uranium-activity models, $\hat{Z}_{gc,j}$, using soil geochemistry data to represent collectively the uncertain ground truth, Z.
For each alternative characterization technology, i:
- Map the boundary of contaminated vs. uncontaminated soil for a threshold concentration, Z*, using geostatistical simulation of $\hat{Z}_j$ to produce a probability map at some desired level of confidence, p. Set remediation flag $R_{i,\mathbf{x}}$ = .true. if panel **x** is to be treated; set $R_{i,\mathbf{x}}$ = .false. otherwise.
- For each selective remediation grid block, **x**:
  - Incur the appropriate characterization cost, $Cchar_i$,**x**.
  - If $R_{i,\mathbf{x}}$ = .true. then clean the panel; accumulate $Ctreat_i$,**x**.
  - If $R_{i,\mathbf{x}}$ = .false. then do not clean the panel; evaluate for potential error.
  - For each geochemistry realization, j:
    - If $\hat{Z}_{gc,\mathbf{x},j}$ > Z* then this is a "failure;" incur failure cost ***$a***; increment number of failures, $smNfail_i$.
    - If $\hat{Z}_{gc,\mathbf{x},j}$ >> Z* then this is a "flagrant" failure: incur failure cost ***$b***; increment number of failures, $lgNfail_i$.
  - Next geochemistry realization, j.
  - Compute expected (average) characterization, treatment, and failure costs for panel **x**.
- Next location, **x**.
- If $smNfail_i$ or $lgNfail_i$ exceeds relevant threshold, there were "too many errors;" incur failure cost ***$c*** or ***$d***, as appropriate.

Next technology, i.

Figure 2. Pseudo-computer-code algorithm for stochastic evaluation of the objective function using geostatistical simulation to capture uncertainty in the actual level of uranium contamination.

4. Results

The values of the objective function, Φ_i, from the stochastic comparison of the alternative characterization methods at three different action levels and at one specific probability level are shown in figure 3; the different components of the objective function are indicated. Note that the value of Φ_i varies according to the question being asked (the specified action level). The total expected costs to remediate the site for an action level of 60pCi/g, including the likely cost of regulatory failure, vary by a factor of 2 from $554,000 (Beta) to more than $1.24 million (ATD and EIC). The cost of treatment alone varies from a low of $57,000 to the *a priori* full-site treatment cost of $880,992 (2016 panels@$437). The lowest *treatment-cost* option was the EIC characterization method, one of the relatively low-cost innovative demonstration techniques. However, despite the low characterization and indicated treatment costs, this method is actually the most expensive overall because of the high probability of failure. Note that four characterization technologies (FID, LRAD, NAD, XRF) are associated with a zero expected cost of failure. Each of these methods indicated that the entire site is contaminated at an action level of 60 pCi/g. Conversely, the ATD technology indicated virtually no contamination and thus its costs are overwhelmingly related to regulatory failure.

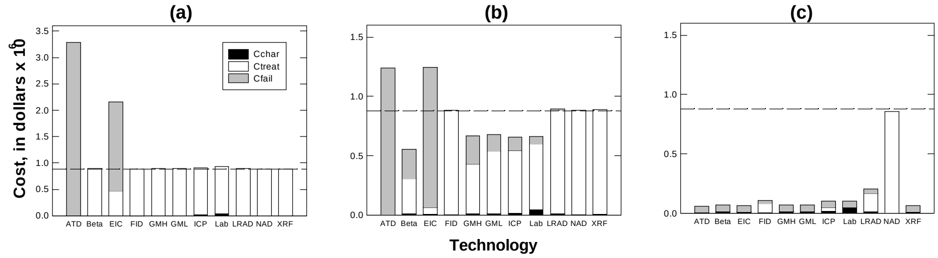

Figure 3. Components of the objective function value (Φ_i) for the 11 different characterization technologies at a probability level of 0.30 and a uranium action threshold of (a) 35 pCi/g, (b) 60 pCi/g, and (c) 100 pCi/g. Horizontal line is the *a priori* cost to treat the expanded site. Note expanded cost scale in (a).

Comparison of the different characterization techniques is expanded in figure 4, in which the value of Φ_i is shown as a function of the probability level (p) used to define the remediation plan. If the site operator is unwilling to accept any risk of an incorrect remediation decision, p is equal to zero and there is no alternative but to treat the entire site area. Characterization in this case is worthless, and the cost of site measurements is in addition to the zero-risk cost of treatment. As the probability level acceptable to the site operator increases (i.e., as the operator becomes less risk-averse), figure 4 indicates that Φ = *Ctotal* typically decreases. There are two characterization technologies for which this reduction does not occur: the ATD and EIC methodologies. They are adaptations of indoor radon-monitoring devices, and their performance has been documented as notably inaccurate and sensitive to changes in environmental conditions in the field (Rautman, 1996).

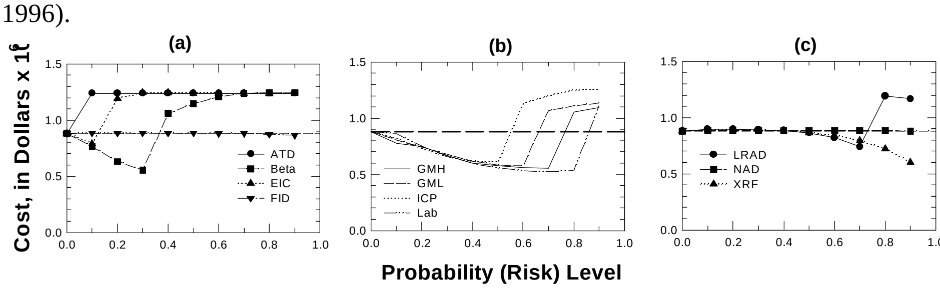

Figure 4. Objective function values (Φ_i) for the different field-characterization technologies as a function of probability level for a uranium action threshold of 60 pCi/g. Horizontal line is *a priori* cost to treat entire site.

Figure 4 also indicates that at some increased probability level, Φ begins to increase abruptly. This increase reflects the changing interplay of the likely costs of treatment and failure with changes in probability level that determine the initial remediation plan. The individual cost components that result from this trade-off between *Ctreat* and E{*Cfail*} are shown in figure 5 for three of the characterization technologies. These graphs emphasize that one can only achieve a net reduction in total remediation cost by trading part of the cost of treatment for a finite probability of failure corresponding to a non-zero

E{*Cfail*} [equation (2)]. At some probability level, the likelihood of failure becomes too great, and it is cheaper simply to clean up the parcel. Note however, that very low probability levels, perhaps corresponding to the "95-percent confidence level" of classical statistics, may not be the most cost-effective level of site characterization and remediation.

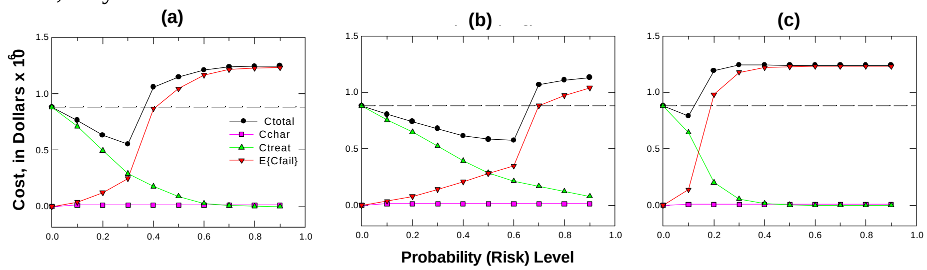

Figure 5. Individual cost curves for three technologies at an action level of 60 Pci/g: (a) beta scintillometer, (b) low-mount gamma-ray spectrometer, and (c) electret-ionization chambers. Horizontal line is *a priori* cost to treat entire site.

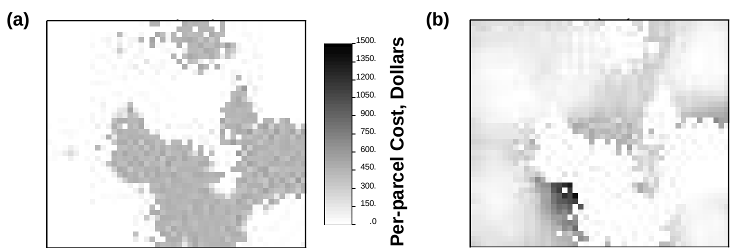

Figure 6. Maps showing the spatial distribution of (a) *Ctreat* and (b) *E*{*Cfail*) for the beta scintillometer characterization technique at a probability level of 0.30; action level 60 pCi/g.

Maps of the per-parcel expected costs are shown in figure 6 for the lowest-cost characterization alternative (Beta). Figure 6(a) simply shows the initial remediation plan developed using probability mapping of the Beta scintillometer data at the minimum-cost probability level of 0.30. The cost of treatment is presumed constant at $437 per parcel. Figure 6(b) shows the per-parcel expected cost of failure. The white areas correspond exactly to the grey parcels in figure 6(a). This map of figure 6(b) clearly contains information that is relevant to continued site characterization. The darker parcels are those associated with the highest expected cost of failure, up to nearly $1,500 per parcel. Because the combined cost of taking an additional sample and cleaning up that parcel (if necessary) is only $167 + $437 = $604, one should be able to reduce the total expected cost through additional sampling in regions where the E{*Cfail*} exceeds this amount. Note that because of spatial correlation, a single additional sample probably will provide information about more than one remediation unit.

5. Conclusions

A quantitative cost-risk decision framework using geostatistical simulation to quantify uncertainty related to less-than-exhaustive site characterization can be used to identify cost-effective plans for environmental remediation. This framework has been applied to the problem of selecting among alternative technologies for measuring the uranium activity of in-situ soil adjacent to an aerosol contamination source at the U. S. Department of Energy Fernald (Ohio USA) site. The decision framework involves minimizing the total expected cost of site characterization, site treatment, and potential regulatory failure to clean the site to established standards across a reasonable number of engineering alternatives. The preferred alternative technology has been shown to be a function of the regulatory action level and the degree of risk-aversion of the decision maker. The lowest total-cost alternative may not be associated with the lowest probability of making incorrect treat vs. leave-in-place decisions for individual remediation units. Characterization, treatment, and potential failure costs vary spatially and maps of these likely costs can provide useful information with respect to staged sampling efforts and stopping criteria for site characterization work.

6. Acknowledgment

This work was supported by the U.S. Department of Energy, Office of Science and Technology under contract DE-AC04-94AL85000, TTP AL-26PL-41, Subsurface Contamination Focus Area (formerly the Uranium-in-Soils Integrated Demonstration Program).

7. References

Deutsch C.V., and Journel, A.G., 1992, GSLIB Geostatistical software library and user's guide, New York: Oxford University Press, 340 p.

DOE (U.S. Department of Energy), 1992, Contaminated Soils adjacent to the sewage treatment plant incinerator, removal action #14 work plan, U. S. Department of Energy, Cincinnati, Ohio.

Douthat, M.M., Steward, R.N., and A.Q. Armstrong, 1995, Cost results from the 1994 Fernald characterization field demonstration for uranium-contaminated soils: Report ORNL/TM-12985, Oak Ridge National Laboratory, Oak Ridge, Tenn., 20 p.

EPA (U.S. Environmental Protection Agency), 1993, Data quality objectives process for Superfund, Interim final guidance 9355.9-01, EPA540-R-93-071, PB94-963203, U.S. Environmental Protection Agency, Office of Emergency and Remedial Response.

Freeze, R.A., Massmann, J., Smith, L., Sperling, T, and James, B., 1990, Hydrogeological decision analysis: 1. A framework: *Ground Water*, v. 28, p. 738–766.

Rautman, C.A., McGraw, M.A., Istok, J.D., Sigda, J.M., and Kaplan, P.G., 1994, Probabilistic comparison of alternative characterization technologies at the Fernald Uranium-in-Soils Integrated Demonstration Project, *in* Proceedings of Waste Management-94, Tucson, Ariz., February 27–March 3, 1994, p. 2117–2124.

Rautman, C.A., Cromer, M.V., Newman, G.C., and Beiso, D.A., 1995, 1994 Fernald field characterization demonstration program data report: Sandia Report SAND95-3024, Sandia National Laboratories, Albuquerque, N. Mex., 170 p.

Rautman, C.A., 1996, Evaluation of measurement reproducibility using the standard-sites data, 1994 Fernald field characterization demonstration program: Sandia Report SAND96-0262, Sandia National Laboratories, Albuquerque, N. Mex., 38 p.

Schilk, A.J., Perkins, R.W., Abel, K.H., and Brodzinski, R.L., 1993, Surface and subsurface characterization of uranium contamination at the Fernald environmental management site, Report PNL-8617, Pacific Northwest Laboratory, Richland, Wash., 43 p.