OAK RIDGE
NATIONAL LABORATORY
MANAGED BY UT-BATTELLE
FOR THE DEPARTMENT OF ENERGY

ORNL/TM-2001/170

OPTIMIZED FIELD SAMPLING AND MONITORING OF AIRBORNE HAZARDOUS TRANSPORT PLUMES: A GEOSTATISTICAL SIMULATION APPROACH

November 2001

Prepared by
DI-WEN CHEN AND HAMILTON T. HUNTER

UT-BATTELLE
ORNL-27 (4-00)

ORNL/TM-2001/170

OPTIMIZED FIELD SAMPLING AND MONITORING OF AIRBORNE HAZARDOUS TRANSPORT PLUMES: A GEOSTATISTICAL SIMULATION APPROACH

Di-Wen Chen[1] and Hamilton T. Hunter[2]
[1]Clark-Atlanta University
[2]Oak Ridge National Laboratory

Date Published: November 2001

Prepared by
OAK RIDGE NATIONAL LABORATORY
P.O. Box 2008
Oak Ridge, Tennessee 37831-6285
managed by
UT-Battelle, LLC
for the
U.S. DEPARTMENT OF ENERGY
under contract DE-AC05-00OR22725

CONTENTS

Page

LIST OF FIGURES

Figure **Page**

ABSTRACT

Airborne hazardous plumes inadvertently released during nuclear/chemical/biological incidents are mostly of unknown composition and concentration until measurements are taken of post-accident ground concentrations from plume-ground deposition of constituents. Unfortunately, measurements often are days post-incident and rely on hazardous manned air-vehicle measurements. Before this happens, computational plume migration models are the only source of information on the plume characteristics, constituents, concentrations, directions of travel, ground deposition, etc. A mobile 'lighter than air' (LTA) system is being developed at Oak Ridge National Laboratory that will be part of the first response in emergency conditions. These interactive and remote unmanned air vehicles will carry light-weight detectors and weather instrumentation to measure the conditions during and after plume release. This requires a cooperative computationally organized, GPS-controlled set of LTA's that self-coordinate around the objectives in an emergency situation in restricted time frames.

A critical step before an optimum and cost-effective field sampling and monitoring program proceeds is the collection of data that provides statistically significant information, collected in a reliable and expeditious manner. Efficient aerial arrangements of the detectors taking the data (for active airborne release conditions) are necessary for plume identification, computational 3-dimensional reconstruction, and source distribution functions.

This report describes the application of stochastic or geostatistical simulations to delineate the plume for guiding subsequent sampling and monitoring designs. A case study is presented of building digital plume images, based on existing "hard" experimental data and "soft" preliminary transport modeling results of Prairie Grass Trials Site. Markov Bayes Simulation, a coupled Bayesian/geostatistical methodology, quantitatively combines soft information regarding contaminant location with hard experimental results. Soft information is used to build an initial conceptual image of where contamination is likely to be. As experimental data are collected and analyzed, indicator kriging is used to update the initial conceptual image. The sequential Gaussian simulation is then practiced to make a comparison between the two simulations. Simulated annealing is served as a postprocessor to improve the result of Markov Bayes simulation or sequential Gaussian simulation.

STOCHASTIC SIMULATION METHODOLOGY

Stochastic simulation is a Monte-Carlo procedure for generating outcome of digital images of a variable that are consistent with its values at sampled locations and with its *in situ* spatial variability, as characterized by histograms and variograms. There are many algorithms for simulating variables, for example, sequential Gaussian simulation, LU decomposition algorithm, indicator-based algorithms, object-based algorithms, simulated annealing etc. In this report the spatial distribution of contamination is simulated using Markov Bayes Simulation, sequential Gaussian simulation, and simulated annealing.

1. The Sequential Simulation Approach

The sequential simulation algorithm is described elsewhere;[1-3] however, to help readers better understand the principle, it is briefly repeated here.

The family of all "sequential" procedures makes use of the same basic algorithm. Consider the distribution over a field A of one or more attributes $z(u)$, $u \in A$. Stochastic simulation is the process of building alternative, equally probable, high-resolution models of the spatial distribution of $z(u)$; each realization is denoted with the superscript l: $\{z^{(l)}(u), u \in A\}$. The simulation is said to be "conditional" if the resulting realizations honor data values at their locations:

$$z^{(l)}(u_a) = z(u_a), \quad \forall l \ .$$

The variable *z(u)* can be continuous, such as concentrations over a contaminated site, or it can be categorical, e.g., indicating different levels of concentrations.

Approximation allows drawing the value of a variable *Z(u)* from its conditional distribution given the value of the most related covariate at the same location *u*. The sequential simulation principle is a generalization of that idea: the conditioning is extended to include all data available within a neighborhood of *u*, including the original data and all previously simulated values.

Consider the joint distribution of N random variables Z_i with N possibly very large. The N random variables Z_j may represent the same attribute at the N nodes of a dense grid discretizing the field A. Or, they can represent N different attributes measured at the same location, or they could represent a combination of K different attributes defined at the N' nodes of a grid with $N = K \cdot N'$.

Next consider the conditioning of these N random variables by a set of n data of any type symbolized by the notation $|(n)$. The corresponding N variate conditioning cumulative distribution function (ccdf) is denoted:

$$F_{(N)}(z_1, \ldots, z_N) = \text{Prob}\{Z_i \le z_i, i = 1, \ldots, N|(n)\}$$

The above expression is completely general with no limitation; some or all of the variables Z_i could be categorical. Successive application of the conditional probability relation shows that drawing an N variate sample from the ccdf can be done in N successive steps, each involving a univariate ccdf with increasing level of conditioning:

- Draw a value $z_1^{(l)}$ from the univariate ccdf of Z_1 given the original data (n). The value $z_1^{(l)}$ is now considered as a conditioning datum for all subsequent drawings; thus, the information set (n) is updated to $(n+1) = (n) \cup \{Z_1 = z_1^{(l)}\}$.
- Draw a value $z_2^{(l)}$ from the univariate ccdf of Z_2 given the updated data set $(n+1)$, then update the information set to $(n+2) = (n+1) \cup \{Z_2 = z^{(l)}\}$.
- Sequentially consider all N random variables Z_i.

The set $\{z_i^{(l)}, i = 1, \ldots, N\}$ represents a simulated joint realization of the N dependent random variables Z_i. If another realization is needed, $\{z_i^{(l')}, i = 1, \ldots, N\}$, the entire sequential drawing process is repeated.

Different simulation algorithms impart different global statistics and spatial features on each realization. For example, simulated categorical values can be made to honor specific geometrical patterns as in indicator-based simulation or the covariance of simulated continuous values can be made to honor a prior covariance model as in Gaussian-related simulation.

2. Markov-Bayes Indicator Simulation

Markov Bayes Simulation is a coupled Bayesian/geostatistical methodology. It quantitatively combines soft information regarding contaminant location with hard experimental results. Soft information is used to build an initial conceptual image of where contamination is likely to be. As hard data are collected and analyzed, indicator kriging is used to update the initial conceptual image. Soft

information can include historical information, non-intrusive geophysical survey data, preliminary transport modeling results, etc.

A Bayesian approach assumes that parameters are unknown initially, but have some known probability distribution called the prior probability density function (pdf). For example, the presence of contamination at a point is unknown initially, but its prior pdf is known. As additional information, such as new experimental or sampling data becomes available, these prior pdfs can be updated quantitatively using Bayes' rule to produce posterior probability density functions:

$$P(X \mid Y) \infty P(X)P(Y \mid X) \ .$$

P(X|Y) is the posterior pdf for *X*, *P(X)* is the prior pdf for *X*, and *P(Y|X)* reflects the probability distribution associated with observing *Y* given the prior pdf of *X*.

From a Bayesian perspective, a two parameter beta distribution $B(\alpha,\beta)$ is a conjugate prior in the context of Bernoulli trials and the binomial distribution.[4] $B(\alpha,\beta)$ ranges between zero and one, and can assume a variety of shapes depending on the values of α and β. For a random variable ψ that follows a beta distribution, the expected value of ψ is given by:

$$E_{(\psi)} = \frac{\alpha}{\alpha+\beta} \ ,$$

where α,β = parameters associated with the beta pdf for ψ, $\alpha,\beta \geq 0$. The variance of ψ is given by:

$$\text{Var } (\psi) = \frac{\alpha\beta}{(\alpha+\beta)^2 \ (\alpha+\beta+1)} \ .$$

In the case of a binomial trial with an unknown underlying probability ψ of seeing a success in any given trial, if *X* successes are obtained in *N* trials, a prior for ψ of the form $B(\alpha,\beta)$ becomes the posterior $B(\alpha+x, \beta+N-X)$. *N* functions as the total amount of additional information supplied to the prior. As *N* grows large, $E(\psi)$ approaches the classical maximum likelihood estimator for ψ, *X/N*, and the $\text{Var}(\psi)$ decreases monotonically.

The issue now is how to update a prior beta distribution at a given point in space with results from hard samples or experimental points nearby in a manner consistent with the derivation of beta distributions as conjugate priors for binomial distributions and that recognizes their spatial autocorrelation. Two pieces of information are required from the set of samples: N^*_{xo}, the total amount of information represented by the set of samples appropriate for that point in space, x_o, and P^*_{xo}, the probability of encountering contamination at *xo* based on the experimental or samples' results. Indicator kriging provides a means for deriving these two pieces of information. An unbiased estimator of P^* at x_o is given by:

$$P^*_{xo} = \sum_{i=1}^{N} \lambda_i Z(x_i) \ ,$$

where

x_i = locations where samples or experimental data have been collected;

$Z(x_i)$ = 0 or 1, depending on whether the sample at x_i encountered contamination below or above a threshold

λ_I = kriging weights

The set of kriging weights, λ, can be derived by solving the following set of simultaneous linear equations:

$$\sum C_{ij}\lambda_i + \lambda_{N+1} = C_{jo} \qquad \text{for} \qquad j = 1,.....,N ,$$

where

C_{ij} = covariance between sample locations x_i and x_j;

C_{jo} = covariance between sample locations x_j and the point where the interpolation is taking place, x_o.

N^* at x_o can be tied to N, the number of samples taken, through the following relationship:

$$N^*_{xo} = \frac{2C_{oo}}{\text{var}_{\text{estim}}} - 1$$

$$\text{Var}_{\text{estim}} = C_{oo} - \left(\sum_{i=1}^{N} \lambda_i C_{io} + \mu \right) ,$$

where

$\text{Var}_{\text{estim}}$ = the estimation variance associated with the interpolation of p^* at location x_o;

C_{oo} = the variance of the indicator values;

μ = the average of the indicator values for the sample locations involved in the updating.

A general Bayesian/geostatistical approach to merging soft and hard data is the Markov Bayes model described by Deutsch and Journel (1998),[1] where the soft indicator data covariances and cross-covariances are calibrated from the hard indicator covariance models. The soft indicator data, i.e., the prior probability cdf's of type, are derived from calibration scattergrams using subroutine "bicalib" in the software package GSLIB developed by Deutsch and Journel (1998).[1] In all other aspects the Markov Bayes algorithm is similar to sequential indicator simulation.

3. Simulated Annealing

Simulated annealing is a generic name for a family of optimization algorithms based on the principle of stochastic relaxation.[5] Generating alternate conditional stochastic images of either continuous or categorical variables with the aid of simulated annealing is a relatively new approach.[1;3;6] The basic idea of simulated annealing is to continuously perturb an original image until it matches some prespecified characteristics written into an objective function. Each perturbation is accepted or not depending on whether it carries the image toward the objective. The following procedure is used here to generate realizations that reproduce both the histogram and semivariogram model:

A. Generate an initial realization $\{z^{(l)}_{(o)}, j = 1,\ldots, N\}$ by freezing data values at their locations and assigning to each unsampled grid node a z-value drawn at random from the target cdf $F(z)$.

B. Compute the initial value of the objective function corresponding to that initial realization:

$$O(0) = \sum_{s=1}^{S} \frac{[\gamma(h_s) - \hat{\gamma}_{(o)}(h_s)]^2}{[\gamma(h_s)]^2},$$

Where $\gamma(h_s)$ is the value of the target semivariogram model at lag h_s, S is lags $\hat{\gamma}_o(h_s)$ is the semivariogram value at lag h_s of the initial realization.

C. Perturb the realization by swapping z-values at any two unsampled locations u'_j and u'_k chosen at random: $z^{(l)}_{(o)}(\mathrm{u}'_\mathrm{j})$ becomes $\mathrm{z}^1_1(\mathrm{u})$ and vice versa. Assess the impact of the perturbation on the reproduction of target statistics by recomputing the objective function accounting for the modification of the initial image.

D. Accept all perturbations that diminish the objective function. Unfavorable perturbations are accepted according to a negative exponential probability distribution:

$$\text{Prob (Accept } i\text{th perturbation}\} = \begin{cases} 1 & \text{if } O(i) \le O(i-1) \\ \exp\dfrac{[O(i-1) - O(i)]}{t(i)} & \text{otherwise} \end{cases}$$

The $t(i)$ of the exponential distribution is analogous to the "temperature" in annealing metals. The success of the annealing depends on a slow cooling of the realization controlled by the contrl function, $t(i)$, that decreases with time. The idea is to start with an initially high temperature $t(0)$, which allows a large proportion of unfavorable perturbations to be accepted at the beginning of the simulation. As the simulation proceeds, the temperature is lowered so as to limit discontinuous modification of the stochastic realization.

E. If the perturbation is accepted, update the initial realization into a new image $\{z^{(l)}_{(l)}(\mathrm{u}'_\mathrm{j}),\ j = 1, \ldots, N\}$, with object function value $O(1) = O_{\mathrm{new}}(0)$.

F. Repeat steps C to E until either the target low value $O_{\min} = 0.001$ is reached or the maximum number of attempted perturbations at the same temperature has been reached three times.

Other realizations $\{z^{(l')}(\mathrm{u}'_\mathrm{j}), j = 1,\ldots,N\}$, $\mathrm{l}' \neq l$, are generated by repeating the entire process starting from different initial realizations.

CASE STUDY

1. Objective of the study

The objective of our simulation study is to generate digital representations or maps of an airborne hazardous transport plume at the Prairie Grass Trials site. We will then use these maps for optimum and cost-effective field sampling design in the future. An instrument cluster carried by a blimp was proposed to successively sample the plume.

2. Available data

The available information at the Prairie Grass Trials site consisted of 138 hard experimental measurements of concentrations (in mg/m^3) and soft quick re-calculation data (we call them transport data) based on these measurements. The raw experimental data are filed in *A Field Program in Diffusion Reports* (Volumes I, II, and III), Morton L. Barad, editor, DTIC Numbers AD-152572, AD-152573, and AD-217076. The data file PR_GRASS_STATIONS contains the UTM coordinates using a central longitude of 99W and the Clarke 1866 ellipsoid. The tracer source location (42.4933N, 98.5717W = 535.200 east, 4704.427 west) is used as the anchor point for all other locations. Tracer sampler stations were calculated to form 50-, 100-, 200-, 400-, and 800-meter arcs to the north of the source location.

Figure 1 shows the locations and concentrations of these experiments. The file PR_GRASS_CALC holds a quick re-calculation of this experiment on a regular equi-spaced grid. The grid consists of 180 × 90 pixels spaced at 10 meter intervals. Considering the PC capacity and CPU time, rows and columns outside the plume were deleted, where the concentration value at each pixel of these rows and columns is zero. The grid is reformed by using short subroutines that we developed, Excel in Microsoft Office, and a window editor and pixels in the reformed grid are reduced to 90 × 80. Figure 2 shows the concentration distribution based on the quick re-calculated transport data. We can see that the plume travels to the north from the source location.

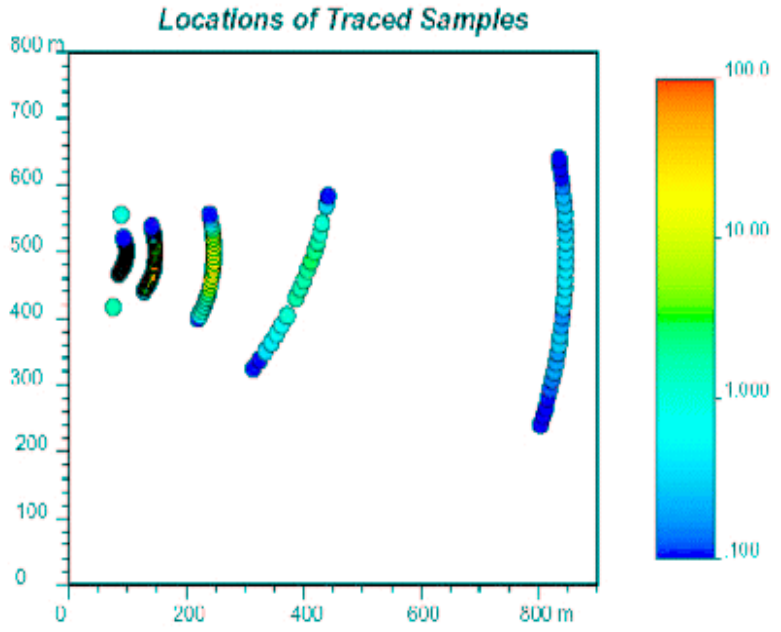

Figure 1. Locations of experiments.

3. Data Analysis

A. Histogram Models

The first step in our simulation study is to decide what probability distribution functions, or more prosaically, what histogram models are to be honored by our simulated concentrations in the plume. We

would like these histograms to be representative of the entire site. Because the experimental data are spatially clustered, the raw histogram (Figure 3) shows that it has sawtooth-like fluctuations. The raw histogram does not provide adequate models for our simulation. To obtain a representative distribution, we use the program declus in GSLIB that assigns declustering weights whereby values in areas/cells with more data receive less weight than those in sparsely sampled areas. The declustered histograms are also smoothed to remove such fluctuations and to increase the class resolution and extend the distribution beyond the sample minimum and maximum values by using the program histsmth in GSLIB. Figure 4 shows a declustered and smoothed histogram plot of the 138 experiments. Figure 5 is a lognormal probability plot of the 138 experiment data.

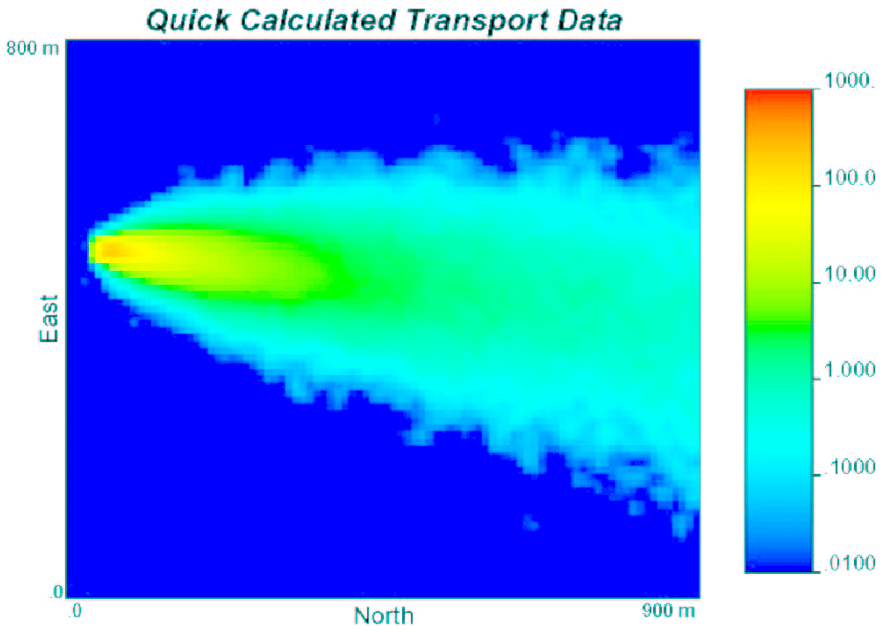

Figure 2. Plume generated by original transport data.

B. Variogram model

The variogram is defined as a measure of spatial variability and is the key to any geostatistical study. The variogram analyses are performed on the 138 experiments (Figure 6) and the soft transport data (Figure 7) in different directions. The omnidirectional (0° ± 90°) and *NS*(0° ± 22.5°) directions represented by red and green are plotted in Figure 6 while the omnidirectional (0° ± 90°) and *EW*(90° ± 22.5°) represented by red and blue are plotted in Figure 7. The variogram of the transport data shows that the phenomenon is reasonably well structured, with a maximum correlation distance (range) of approximately 5 meters. Anisotropy is found and the shape of the variogram is exponential. The variogram of experiment data shows that the phenomenon is not normally structured with significant nugget effect of 1.6 meters and indeterminate maximum range.

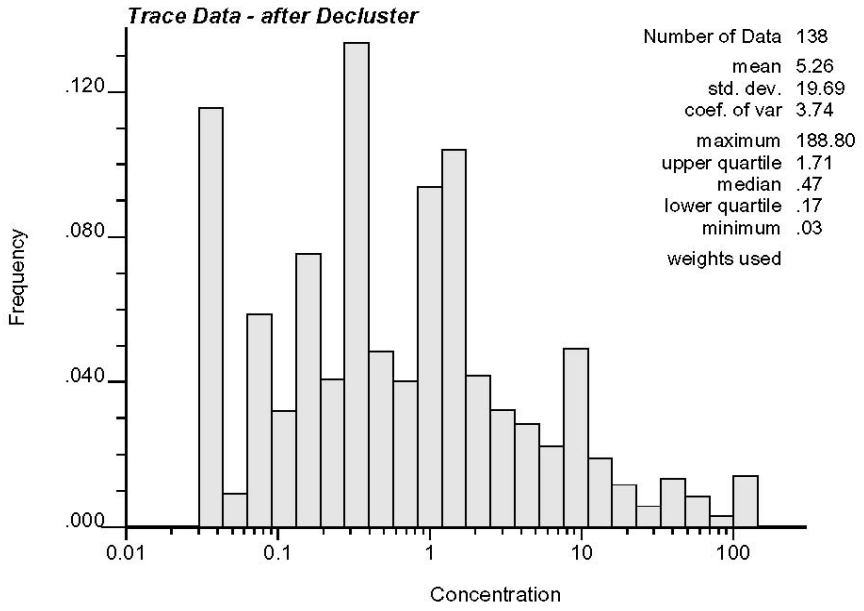

Figure 3. Histogram of the 138 tracer experiments.

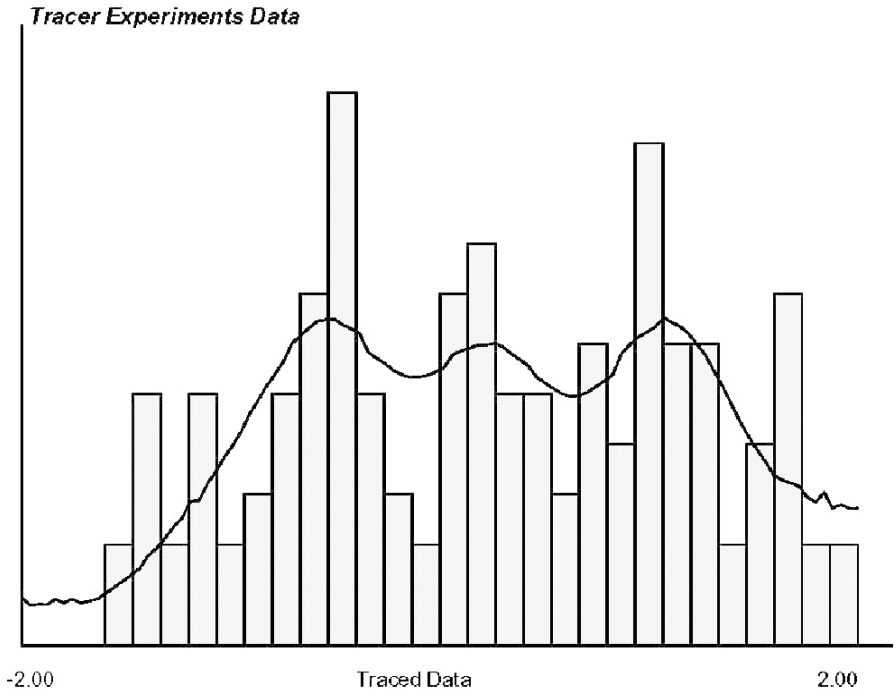

Figure 4. Smoothed histogram of 138 tracer experiments.

C. Bivariate Distribution

Before we can proceed to the stochastic simulation, one last step is required. Our simulation algorithm can be used to generate fields of only one variable at a time. However, we wish to simulate two variables, Z_1 (hard experiment data) and Z_2 (soft transport data), reproducing not only their respective spatial variation structures but also the relationship between them shown in Figure 8. We must therefore "decouple" the variables Z_1 and Z_2 so that we can simulate them independently. To do this, we use the following principal component transformation that yields the independent variables Y_1 and Y_2 from the correlated variables Z_1 and Z_2:

$$Y_1 \equiv Z_1 \text{ and } Y_2 = \frac{1}{\sqrt{1-\rho^2}} Z_2 - \frac{\rho}{\sqrt{1-\rho^2}} Z_1 ,$$

where ρ is the correlation coefficient between Z_1 and Z_2 which is found to be 0.887 (Figure 8).

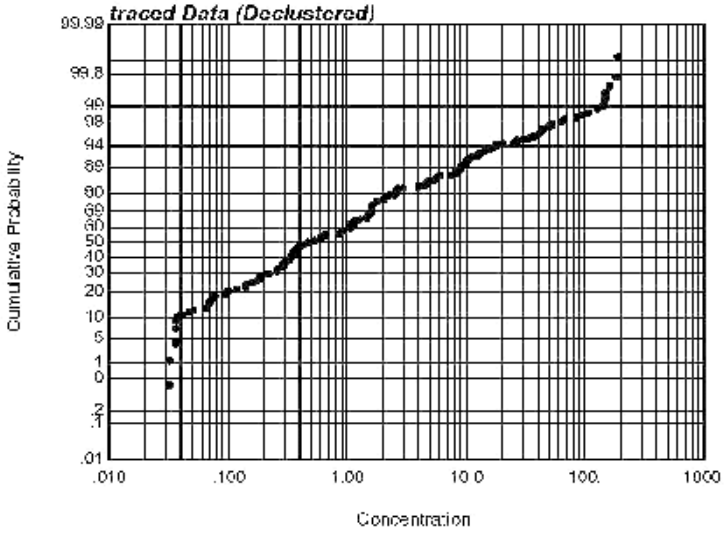

Figure 5. A cumulative probability plot of the 138 tracer data.

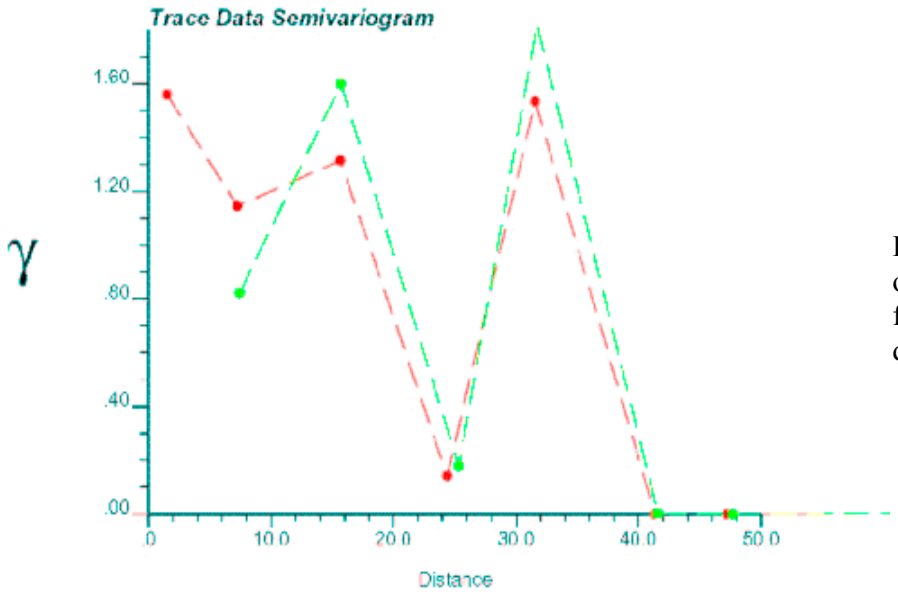

Figure 6. Measure of spatial variability from the 138 tracer data.

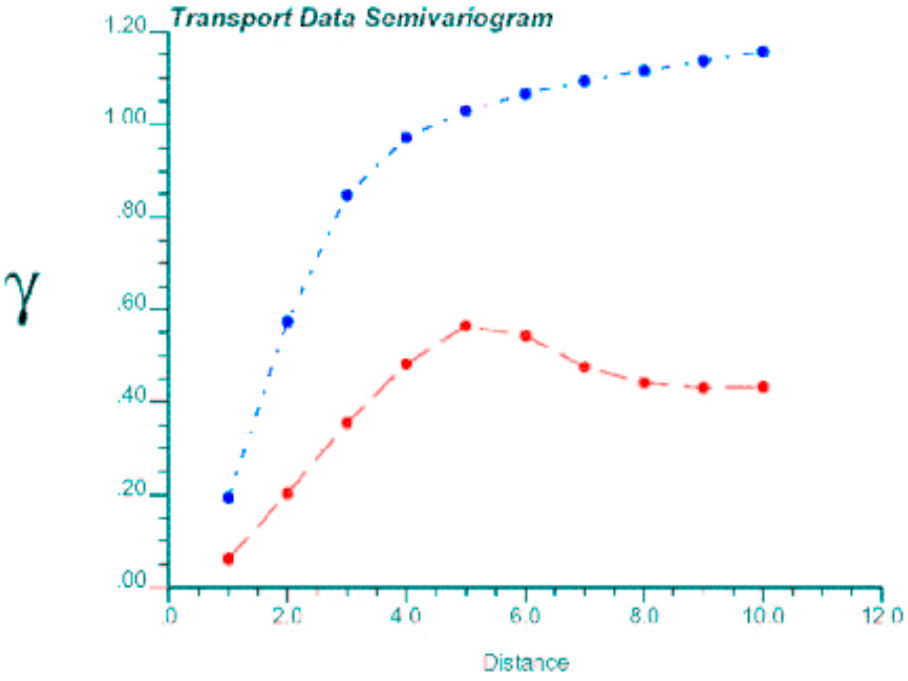

Figure 7. Measure of spatial variability calculated from the transport data.

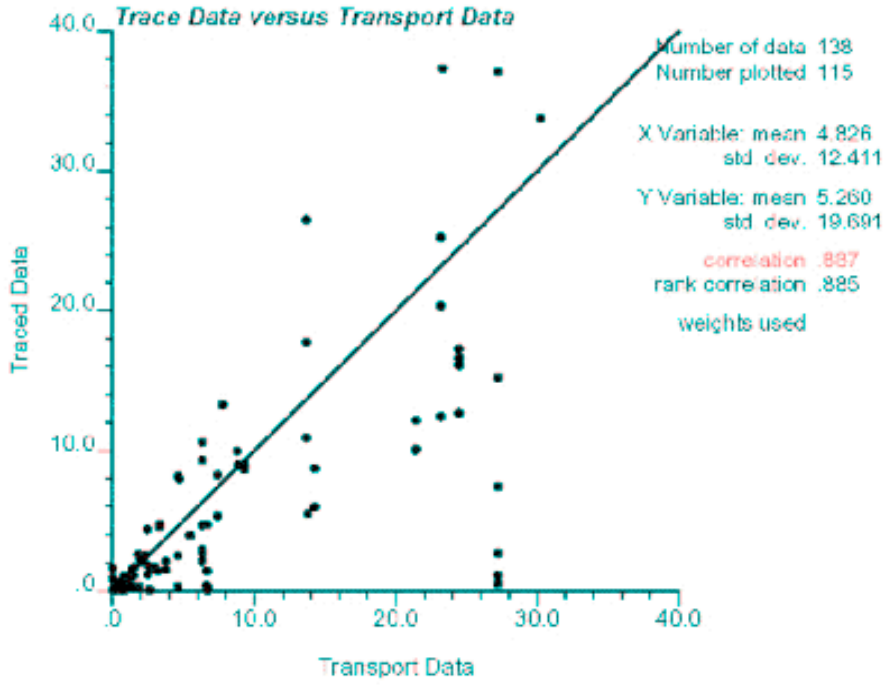

Figure 8. Scatterplot of tracer data versus transport data.

D. Simulations Performance

We are now ready to simulate fields of the two independent variables, Y_1 and Y_2. The simulations of Y_1 and Y_2 are to be conditioned on 138 experiment data (hard) and grided transport data (soft) with 90 × 80 pixels. Each pixel has 10 m × 10 m intervals.

To perform our simulations, we use the Markov Bayes Simulation first. The sequential Gaussian simulation is then conducted to make a comparison between the two simulations. A simulated annealing simulation is served finally as a postprocessor to improve the result of the Markov Bayes simulation or sequential Gaussian simulation. Plume images in Figures 9, 10, and 11 are generated by Markov Bayes simulation, sequential Gaussian simulation, and simulated annealing, respectively. It can be intuitively perceived that the image produced by the Markov Bayes simulation is the best image.

More results will be obtained from the three different simulation algorithms reproducing correlation between soft and hard data sets. Uncertainty associated with those simulations will also be measured and reported at a future time.

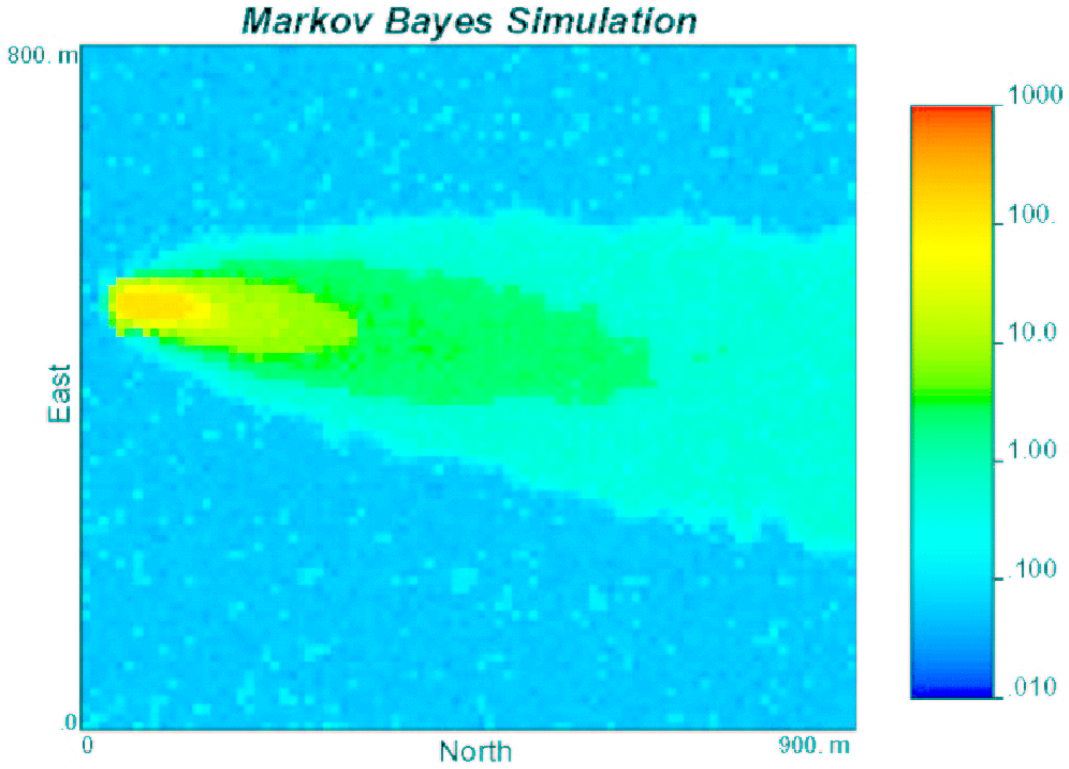

Figure 9. Plume generated by Markov Bayes simulation.

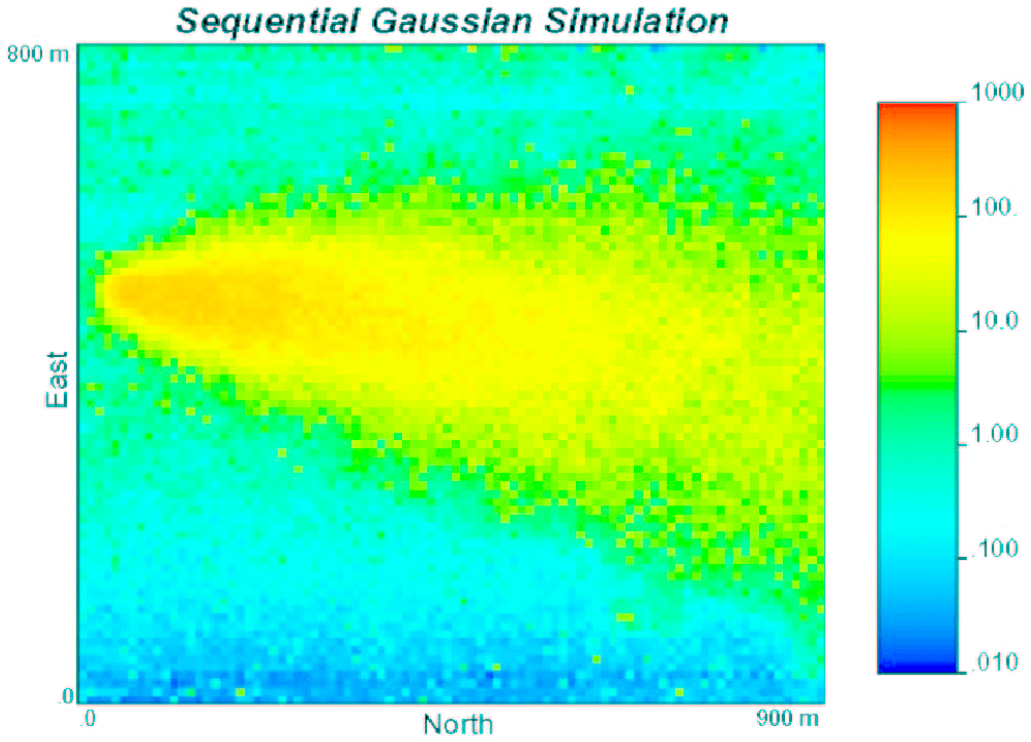

Figure 10. Plume generated by sequential Gaussian simulation.

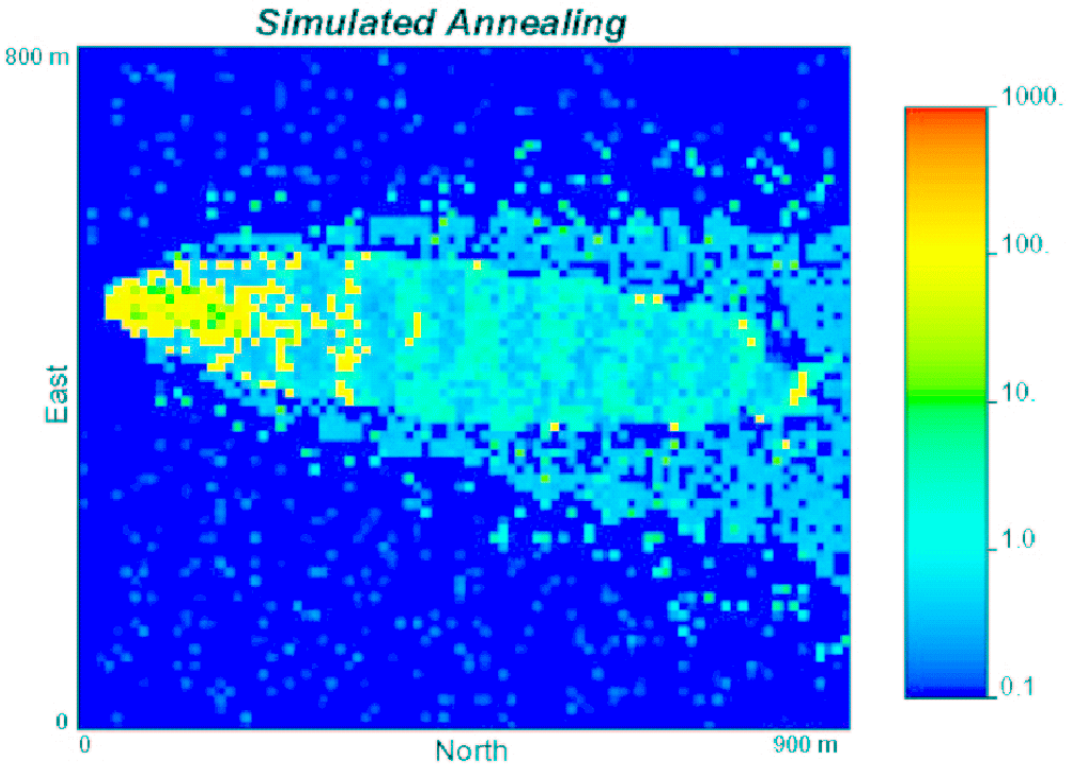

Figure 11. Plume generated by simulated annealing.

CONCLUSION

Stochastic simulation is emerging on environmental and geotechnical fronts as an invaluable tool for characterizing spatial or temporal phenomena. The challenge for sampling programs at an airborne hazardous site is providing real-time sampling program support. Supporting field sampling and monitoring programs requires the ability to estimate the nature and extent of contamination based on available information, to measure the uncertainty associated with those estimates, and to direct sampling and monitoring designs so that sample locations maximize information gained. The case study at the Prairie Grass Trials site shows that geostatistical analysis and stochastic simulation are well suited to quantitative sampling program support.

By using Markov-Bayes indicator simulation, prediction of contaminant concentration can be **improved** by integrating effectively the direct hard experimental data into the soft preliminary transport modeling results. It means that a plume can be updated when additional samples are available. The updated high-resolution, real-time plume images will have a realistic texture that mimics an exhaustive characterization, while maintaining the overall statistical character of the experimental data. In Markov-Bayes indicator simulation, it is crucial that the bivariate relationship existing between the hard and soft information be correctly rendered. Very often this relationship cannot be captured by classical parametric description based on the means and variances of the marginal distributions and the correlation coefficient.

Simulated annealing is a relatively new approach that constructs stochastic simulation by iterative trial and error based on optimization algorithms. This approach is very general and can be used to address many different situations. It should be pointed out, however, that the relevance of its results depends heavily on how meaningful the physical relationship is between the main and co-attribute as described by the bivariate probabilistic model.

ACKNOWLEDGMENTS

The work presented in this report was supported by the Radiation Safety Information Computational Center (RSICC), Atomic Computation Experts Institute (ACES), Oak Ridge National Laboratory, and the U.S. Nuclear Regulatory Commission's Historically Black Colleges and Universities (HBCU) Faculty Research Participation Program.

APPENDIX A
Histogram Smoothing

The process of histogram smoothing not only removes sawtooth-like data fluctuations; it also allows increasing the class resolution and extending the distribution(s) beyond the sample minimum and maximum values. Geostatisticians have developed their own smoothing routines rather than relying on traditional kernel-type smoothing since traditional methods do not honor sample statistics, such as the mean and variance, and simultaneously important quantile-type information such as the median and zero probability of getting negative values. The more flexible annealing approach has been retained for smoothing histogram[1] (Deutsch and Journel, 1998).

Consider the problem of assigning N probability values p_i, $i = 1, \ldots, N$, to evenly spaced z values between given minimum $z_{\min}$ and maximum $z_{\max}$. The equal spacing of the z_i values is

$$\Delta z = \frac{1}{N}(z_{\max} - z_{\min}) \quad ,$$

with $p_i \geq 0, \forall_i = 1, \ldots, N, p_i = 0, z_1 = z_{\min}$ and $p_N = 0, z_N = z_{\max}$. The idea is to choose a large N (100-500) so that the resulting distribution can be reliably used for geostatistical simulation.

The final set of smoothed probabilities is established from an initial set of probabilities by successive perturbations. The probabilities mechanism consists of selecting at random two indices i and j such that $i \neq j$ and $i, j \in]1, N[$. The probability values at i and j are perturbed as follows:

$$p_i^{new} = p_i + \Delta p$$

$$p_j^{new} = p_j + \Delta p$$

j and i are chosen such that $p_j \leq p_i$ and the incremental change Δp is calculated as

$$\Delta p = 0.1 U p_j \quad ,$$

where 0.1 is a constant chosen to dampen the magnitude of the perturbation and U is a random number in [0,1}. There is no need to check that $p_i \leq 1.0$; it must be ≤ 1.0 since p_j^{new} is greater than zero.

If the initial p_i, $i = 1, \ldots, N$, values are legitimate (i.e., $p_i \in [0,1], \forall_i$ and $\sum_{i=1}^{N} pi = 1$), then any set of probabilities derived from multiple applications of this perturbation mechanism is also legitimate.

The following component objective functions have been considered:

- For the mean:

$$O_m = (\bar{z} - m_z)^2 \quad ,$$

where mz is the target mean (from the data or specified by the user) and $\bar{z}$ is the average from the smoothed distribution:

$$\bar{z} = \sum_{i=1}^{N} p_i z_j \quad .$$

- For the variance:

$$O_v = (s^2 - \sigma^2)^2 ,$$

where σ^2 is the target variance (from the data or specified by the user) and s^2 is the variance from the smoothed distribution:

$$S^2 = \sum_{i=1}^{N} p_i z_i^2 - \bar{z}^2 .$$

- For a number of quantiles:

$$O_q = \sum_{i=1}^{n_q} [cp_i - F(z_i)]^2 ,$$

where n_q is the number of imposed quantiles, cp_i is the smoothed cdf value for threshold z_i, and $F(z_i)$ is the target cumulative probability (from the data).

- The smoothness of the set of probabilities p_i, $i = 1, \ldots, N$, can be measured by summing the squared difference between each p_i and a smooth $\hat{p}_i$ defined as an average of the values surrounding i, that is:

$$O_s = \sum_{I=1}^{N} (p - \hat{p}_i)^2 ,$$

where the p_i, $i = 1, \ldots, N$, are the smoothed probability values and $\hat{p}_i$ $i = 1, \ldots, N$, are local average of the p_i values:

$$\hat{p}_i = \frac{1}{2n_0} \sum_{k=-n_o, k\neq i}^{n_o} p_{i+k} \qquad i = 1,\ldots., N ,$$

where n_0 is the number of values in the smoothing window (say, 5-10), $p_i = 0 \; \forall_i \leq 1$ and $i \geq N$. The global objective function is defined as the sum of these four components:

$$O = \nu_m O_m + \nu_v O_v + \nu_q O_q + \nu_s O_s$$

The weights ν_m, ν_v, ν_q, and ν_s are computed such that each component has an equal contribution to the global objective function.

REFERENCES

1. Clayton V. Deutsch and Andre G. Journel, *GSLIB: Geostatistical Software Library and User's Guide*, New York : Oxford University Press, 1997.

2. R. Mohan Srivastava, "An overview of stochastic methods for reservoir characterization," in *Stochastic Modeling and Geostatistics: Principles, Methods, and Case Studies*, ed. J. M. Yarus and R. L. Chambers., *AAPG Computer Applications in Geology No. 3*, AAPG (1994), 1-16.

3. Pierre Goovaerts, *Geostatistics for natural resources evaluation*, New York: Oxford University Press, 1997.

4. P. M. Lee, *Bayesian Statistics: An introduction*, New York: Oxford University Press, 1989.

5. C. Farmer, "The generation of stochastic fields of reservoir parameters with specified geostatistical distribution," pg 235-252 in *Mathematics in Oil Production*, ed. S. Edward and P. King, Oxford: Clarendon Press, 1988.

6. Jean Paul Chiles and Pierre Delfiner, *Geostatistics modeling spatial uncertainty*, New York: Wiley, 1999.

INTERNAL DISTRIBUTION

1. Central Research Library
2. ORNL Laboratory Records (OSTI)
3. ORNL Laboratory Records (RC)
4. Jim A. Bucholz
5. S. N. Cramer
6. Mark D. DeHart
7. Keith F. Eckerman
8. Ronald James Ellis
9. Margaret B. Emmett
10. Jess C. Gehin
11. Norman Maurice Greene
12. J. N. Herndon
13. David M. Hetrick

14-18. H. T. Hunter

19. Daniel T. Ingersoll
20. Jeffrey O. Johnson
21. Ronald W. Lee
22. Richard A. Lillie
23. Cecil V. Parks
24. Thomas E. Potok
25. Robert W. Roussin
26. Robert T. Santoro
27. Royce O. Sayer
28. Charles O. Slater
29. Timothy E. Valentine
30. John C. Wagner
31. Jy-An John Wang

32-36. RSICC

EXTERNAL DISTRIBUTION

37. Shahrouz Aliabadi, Clark Atlanta University, Dept. of Engineering, Atlanta, GA 30314
38. David Beck, U.S. Department of Energy, 4A-019, 1000 Independence Ave. SW, Washington, DC 20585
39. Timothy E. Beville, U.S. Department of Energy, 19901 Germantown Rd., Germantown, MD 20874
40. Thomas M. Black, U.S. Department of Energy, 19901 Germantown Road, MS A-359, Germantown, MD 20874-1290
41. William M. Blumberg, U.S. Nuclear Regulatory Commission, One Flint North, Mail Stop 10 H4, 11555 Rockville Pike, Rockville, MD 20852-2738
42. John R. Boisseau, Ph.D., Director, Texas Advanced Computing Center, University of Texas at Austin, Campus Mail Code: R8700 Austin, TX 78712
43. D. E. Carlson, U.S. Nuclear Regulatory Commission, Reactor and Plant System Branch, Division of System Research, Office of Nuclear Regulatory Research, MS T-10 G6, RM T-10, 17, Washington, DC 20555-0001

44-48. Di-wen Chen, Clark Atlanta University, Engineering, 3037 Science Research Center, Atlanta, GA 30314

49. Jay Cook, U.S. Department of Energy, Bldg, FORS MS SO42, Washington, DC 20585
50. David H. Crandall, U.S. Department of Energy, DP-18/4A-045, Dir., Inertial Fusion/Natl. Ignition Facility Project, 1000 Independence Ave. SW, Washington, DC 20585
51. Steven P. Curtis, U.S. Department of Energy, Nevada Operations Office, 232 Energy Way North, Las Vegas, NV 89030
52. James S. Ellis, Lawrence Livermore National Lab., P.O. Box 808 (L-103), Livermore, CA 94551-0808
53. Abebe Fissehe, Clark Atlanta University, Mathematical Sciences, 142 McPheeters-Dennis, Atlanta, GA 30314
54. Lance Kaplan, Clark Atlanta University, Dept. of Engineering, Atlanta, GA 30314
55. Libby Kittrell, Oak Ridge Institute for Science and Education, P.O. Box 117, Oak Ridge, TN 37831-0117

56. Jane Long, Lawrence Berkeley Laboratory, Building 50E, Earth Science Division, Berkeley, CA 94720
57. Randal L.N. Mandock, Clark Atlanta University, Department of Physics, Atlanta, GA 30314
58. Doug Minnema, U.S. Department of Energy, MS DP-45, 19901 Germantown Road, Germantown, MD 20874
59. Bruce E. Rittmann, Northwestern University, 2145 Sheridan Rd, #A228, Evanston, IL 60208-3109
60. E. Sartori, OECDNEA, Le Seine St-Germain 12, Boulevard Iles, 92130, Issy-les-Moulineaux, France
61. Williams Seffens, Clark Atlanta University, Biological Sciences, 4029 Science Research Center, Atlanta, GA 30314
62. Richard E. Toohey, ORISE, RIDIC, Mail Stop 51, P.O. Box 117, Oak Ridge, TN 37831-0117
63. Robert S. Van Pelt, Environmental Restoration Engineering, Westinghouse Savannah River Company, Building 730-2B, Aiken, SC 29803
64. Paul Vogel, U.S. Department of Energy, 1000 Independence Avenue, S.W., Washington, DC 20585