classroom, in discussions such as this, and through integrated projects with our colleagues.

ACKNOWLEDGMENT

Oak Ridge National Laboratory is managed by Martin Marietta Energy Systems, Inc., under contract DE-AC05-84OR21400 with the U.S. Department of Energy.

ADDITIONAL REFERENCES


1. INTRODUCTION

Gotway has described an approach to problem solving for an important class of spatial problems in which, for most cases, an analytical solution is well beyond present capabilities. The approach is illustrated through a spatial analysis using transmissivity data from a hydrogeologic unit above the potential repository site for the Waste Isolation Pilot Plant (WIPP), an interesting example of this type of problem that demonstrates many of the difficulties involved. Many spatial analyses of similar complexity would benefit from use of the expertise and methodology employed here. In this discussion, I will first elaborate on three points addressed by Gotway relating to variogram modeling and to the probabilities associated with specific realizations. The remainder of the discussion is then focused on three further points, two that present difficult issues in many spatial analyses but appear relatively straightforward for the WIPP problem and a third issue relevant to all problems involving stochastic simulation.

Several steps are performed in the "WIPP analysis" that might provide important insight to statis-
cians or geostatisticians facing this type of an analysis. Two points emphasized by Gotway that I feel are often overlooked in spatial data analyses are related to the modeling of variograms. These include the use of a robust estimator for the empirical variogram and the weighted least squares fitting of the parametric variogram model to the empirical variogram. A third point related directly to the WIPP analysis concerns probabilities associated with the generated realizations.

Many of the present geostatistical packages for computing and plotting the empirical variogram are flexible in allowing the user to specify the partition over the range of spatial differences between data points as well as various parameters associated with anisotropy, but the procedures often assume the desired computational formula to be as specified in Equation (3.5) of the WIPP analysis. Although this procedure might permit the user to note the effect of outliers through a variogram plot (depending on the user’s curiosity and available program options), it provides little guidance concerning alternative variogram models that might be applicable. Using the robust techniques discussed in the WIPP analysis can suggest alternative models.

Fitting a parametric model to the empirical variogram is another difficult step in many analyses of spatial data, with several techniques available. Many geostatistical packages feature a visual fitting procedure that permits the user to combine several variogram models into a single (linear combination of models) estimate. One problem with relying entirely on visual fitting or combining visual fitting with ordinary least squares estimation is that these procedures do not account for differences in the number of sample pairs within each “bin” or correlation between bin averages making up the empirical variogram. One possible consequence is that too much weight will be attached to highly variable empirical estimates. Using generalized least squares or weighted least squares in conjunction with these more standard techniques can help identify problems resulting from unequal bin sizes and can provide alternative models.

In describing the rationale behind stochastic simulation, Gotway (after Journel 1987 and Journel and Alabert 1989) refers to the generated realizations as “equiprobable,” a term that is accurate for the present analysis only in that any realization has probability 0 of occurrence (ignoring computer truncation of the random numbers). In the general case, the set of realizations represents a random sample based on a random field model that is determined by the statistical models input to the simulation algorithm (referred to here as input models—variograms or histograms, for example) and by the simulation algorithm used to generate the realizations. Once these parameters of the simulation method have been established, so have the probabilities of arbitrary sets of realizations.

The remainder of this discussion focuses on three practical issues relative to stochastic simulation that were not emphasized by Gotway and are not, in general, points likely to be made in a case study. The first issue deals with the question of stationarity. The second issue concerns the selection of sample statistics obtained from data at or near the site and statistical models based on these data. I will use the terms sample statistics and input models to refer to these quantities. The third issue focuses on the relationships between sample statistics, input models, and statistics computed for each generated realization (referred to here as local statistics). The question discussed for this issue is: What influences should and should not be contributing to differences between the models and sets of statistics?

2. STATIONARITY

The requirements for second-order stationarity were specified in Equations (3.2) and (3.3) of the WIPP analysis. If the assumption of Gaussianity of the random field is valid, then Equations (3.2) and (3.3) imply strong stationarity, for which all finite dimensional distribution functions $F$ satisfy

$$F_{z_1, \ldots, z_n}(z_1, \ldots, z_n) = F_{z_1+h, \ldots, z_n+h}(z_1, \ldots, z_n)$$

for all vectors $h$ and integers $n$.

The issue of stationarity is very important. For many applications an appropriate approach requires, first, trying to establish the thickness or shape of specific subsets of the region (different stratigraphic units, for example) and then simulating values within these (assumed stationary) subregions. The former step leads to categorical or discrete variable simulations, which in general involve a different set of assumptions and different stochastic simulation methods. These steps in a spatial data analysis are of critical importance because incorrect assumptions regarding the stationarity of the region could introduce large biases in the distribution of system responses.

This problem was relatively straightforward for the WIPP analysis. Here, the analysis was performed over a single member of the Rustler formation in which the variable being analyzed (transmissivity) is a measurement representative of the entire member (vertically) at a given location. Hence the problem is two-dimensional (many of the nonstationarities in a hydrogeologic region are vertical reflecting different physical processes occurring over time). In addition, assumptions of stationarity are more likely to be applicable to this analysis than many others because the entire sampled unit (the Culebra member
of the Rustler formation) was created through the same depositional and post-depositional processes. This is not to imply that assumptions of stationarity at the WIPP site should be (or were) made without question. In many cases, very little information concerning the stationarity of a region can be obtained from site information or from available data.

3. WHICH SAMPLE STATISTICS AND INPUT MODELS SHOULD BE RETAINED?

The determination of which statistics should be calculated from sample data, modeled, and input to the simulation algorithm is an important concern in most spatial applications, complicated in many cases by the processing of realizations through a nonlinear transfer function. Different transfer functions will respond differently to different spatial features in the generated realizations. For any particular application, this relationship must be evaluated.

There are several ways to model a real-valued process over a finite lattice of points. Alternatives include Gaussian random fields and Markov random fields in which the latter may assume one of several conditional distributions describing the dependence of the value at one point on values at other points within a fixed neighborhood. Although progress has been made on characterization of Markov random fields (c.g., Besag 1974; Cressie 1991, secs. 6.4, 6.6), most of the attention for geostatistical applications has focused on methods for generating Gaussian random fields. These methods include LU decomposition and turning bands as mentioned in the WIPP analysis, as well as the sequential Gaussian approach (Deutsch and Journel 1992) and spectral approaches such as those developed by Mejia and Rodriguez-Iturbe (1974) and Gutjahr (1989). If the process to be generated is Gaussian and the second-order stationarity conditions are applicable, then any one of these methods might be appropriate.

One problem with using a single set of statistics to specify the spatial relationship for an entire region is that this set of statistics is then applicable to all values over the range of the variable to be simulated. This may prove inappropriate if (a) spatial relationships between neighboring values differ at different levels of the variable or (b) there are distinct subsets of the region in which the general pattern of the spatial relationship no longer applies. Situations in which (a) is applicable provided one of the primary motivations for partitioning the range of possible values of the simulated variable into an ordered set of discrete categories and allowing different (indicator) covariance relationships for each category (Journel 1987; Journel and Alabert 1989). These references cite other advantages to this "sequential indicator approach," including avoiding maximum entropy characteristics of realizations generated as Gaussian random fields (a possible problem if the region is not Gaussian) and the ease of incorporating “soft” information (information about the region other than fixed data values—a restricted range of values at a point, for example) in the simulation.

In a comparative study of commonly used geostatistical simulation algorithms, Gotway and Rutherford (1994) obtained results that indicated that for sequential indicator simulation realizations generated based on a continuous variable region, the potential advantages of this approach were more than compensated for by the loss of information suffered in the discretization transformation and back transformation. This appeared to be largely the result of the transformation method (as opposed to other aspects of the sequential indicator simulation method), which introduces a “between category” nugget effect in the generated realizations. This is an area where further research should prove profitable because more appropriate transformation methods might make available several discrete-variable simulation techniques for continuous-variable simulation problems.

For case (b), in which there are distinct subregions with differing spatial characteristics, the general approach for each realization is first to resolve (through simulation) uncertainties concerning the boundaries of the subregions and then generate values within each subregion according to the appropriate input models. Techniques for simulating the discrete subregions include the truncated Gaussian approach (Dowd 1992; Galli, Guerillot, Ravenne, and HERESIM Group 1990) or a discrete version of the sequential indicator approach (Deutsch and Journel 1992). The former method is restricted to a single covariance function for expressing the continuity within subregions and the spatial relationship between subregions, whereas the latter method permits an indicator variogram to be specified for each type of subregion but provides no guidance for between subregion relationships beyond that given in conditioning data (conditioning data will help order the subregions). For cases in which the specific shape of the subregions is important, methods used for Boolean or more general germ-grain models may be more appropriate or statistics used for the characterization of random sets may be adapted for simulation. These techniques were described and reviewed by Stoyan, Kendall, and Mecke (1987, chaps. 4, 5 and 6).

The question of which random-field model is appropriate for a specific application is a question that must be investigated by the practitioner. In the WIPP case study, the observations made by Freeze (1975) concerning the distribution of transmissivity measurements could not be disputed on a basis of site
data, and hence the log transmissivity field was assumed to be Gaussian.

4. **HOW CLOSELY SHOULD THE LOCAL STATISTICS MATCH EMPIRICAL INFORMATION?**

To address the question “How large should the deviations be between the sample statistics and values of these same statistics computed over each of the generated realizations?” one must determine which sources of uncertainty or variability should contribute to these differences and which should not. At least three sources come into play here. The first two relate to uncertainties in the sample statistics and input model, and the last source applies strictly to variability in models calculated from the local statistics for each realization (these terms were defined in the last paragraph of Section 1 of this discussion). These sources of uncertainty or variability are as follows:

1. Uncertainty in the sample statistics that results from their computation based on limited data. This source of uncertainty was grouped with uncertainties about stationarity and isotropy by Mantoglou and Wilson (1982) and called “model-to-reality fitting error.”

2. The range over which the data are gathered not being “very large” compared to the range of the model for the stochastic process assumed to apply to this area so that large discrepancies between the sample statistics and the stochastic process model can be expected. Similarly, if the region of interest is not “very large” compared to the range of the model for the stochastic process, then there should be large discrepancies between the local statistics and the stochastic process model.

3. The final source of variability being the result of a random component introduced by the procedures used in the simulation algorithm. It has nothing to do with uncertainties in the site-related variables. Alternative Gaussian-based simulation methods can have deviations of different magnitude, on average, between models based on local statistics and the input models. Differences in the magnitude of deviations from the input models are even possible for the same simulation method depending on the options selected. Mantoglou and Wilson (1982) showed that different numbers of lines selected for the turning-bands method can effect deviations from the input variogram model. Guardiano (1993) acknowledged that alternative methods of path selection for sequential algorithms may effect properties of the resulting realizations.

In a typical geostatistical analysis, it is not unusual for source 3 to be included in the analysis (often without knowledge) and for sources 1 and 2 to be omitted. In these cases, the input models represent the closest fit to the sample statistics, and these models are provided as input to the simulation algorithm for all realizations. Unfortunately, this means that deviations from the input models (which play the role of the stochastic process model to the simulation algorithm) are a function of the simulation method used and not of the real uncertainties of the problem as specified in sources 1 and 2. If so, very misleading conclusions can be reached.

An alternative approach that appears to merit investigation would be to attempt to quantify the uncertainties associated with sources 1 and 2 and incorporate this uncertainty into the input models for the simulation algorithm. The simulation algorithm could then be altered to force an exact match between models based on local statistics and input models. Matheron (1965, chap. XIII) investigated the magnitude of expected differences in variograms resulting from sources 1 and 2 for some simple variogram models. Simulation offers another possible means of approximating variability from these sources. Uncertainties of these types can be input in similar ways to the input models) to the simulation algorithm. Output from the simulation (the realizations) could be altered so that models based on them conform exactly to the input models through one of several post processing procedures (techniques for deriving alternative realizations from existing realizations). Two promising techniques include annealing, a technique that has already been applied to spatial analyses (e.g., see Deutsch and Journel 1992) and genetic algorithms that have achieved success in optimization (for this case, minimizing differences between models based on local statistics and input models based on empirical information) for other types of problems.

**ADDITIONAL REFERENCES**


**TECHNOMETRICS, MAY 1994, VOL. 36, NO. 2**
I appreciate the time the discussants took to read the article and formulate their replies. The in-depth nature of their comments and the quality of their questions and suggestions adds greatly to the methodology being illustrated. In trying to unify their points into a few themes for a reply, I found four broad topics to address in greater detail—the nature of modeling, relaxing the multivariate Gaussian assumption, limitations of the turning-bands algorithm and comparison to other simulation algorithms, and sensitivity/uncertainty analyses.

1. THE NATURE OF MODELING

Clearly, trying to make predictions about the physical characteristics of the Earth 1,000 to 10,000 years into the future is difficult. Wendelberger and Beckman use the word “ludicrous,” but I find this label harsh. Perhaps “unrealistic” is a better adjective. Nevertheless, such demonstrations are required by Federal regulations, and there seems to be little tolerance for an “Emperor has no clothes” revelation. Rather, the attitude has been to develop the best models possible and to concentrate on defending the methodology used in the models. Although Toran suggests that modeling is “best used for understanding what we do not know about our system rather than for prediction” (p. 150), I believe that prediction is a vital function of any modeling effort. Often, it provides a technique for model validation, and ideally, as the models become more refined, it can provide valuable information on which to base decisions. One example that comes to mind is the collection of models used by meteorologists in weather forecasting. We all realize that these models are not perfect, but we often rely on their predictions when making decisions about outdoor events, automobile trips, vacations, and so forth. Although predictions about next year’s weather are rather tentative, it may be possible, given a long cyclic history of Ice Ages and warming trends, to make some predictions about the next major climatic fluctuation thousands of years in the future. So, what makes predictions about the Earth 1,000 years into the future unrealistic is not the basic idea that we cannot extrapolate far into the future with predictive models, but, as Toran points out, the fact that we are lacking 1,000 years of hydraulic record on which to base such predictions. As scientists become more involved in the formulation of regulations, I am hopeful that the regulations will become more realistic and more meaningful.

2. RELAXING THE MULTIVARIATE GAUSSIAN ASSUMPTION

One of the questions posed by several of the discussants is how to simulate processes that are not based on a multivariate Gaussian assumption. This is a tricky problem because investigating the assumption of a complete multivariate normal distribution is very difficult, particularly in geostatistical formulations in which only one partial realization of a random function is observed. A stationarity assumption (discussed in detail by Rutherford) is required to create the “replication” necessary to estimate parametric functions such as the semivariogram. This situation is somewhat different from classical multivariate statistics in which there are \( n \) observations on \( p \) variables. The best one can do with geostatistical data is to try to assess the nature of the underlying marginal distribution, and, if data are adequate, the nature of the bivariate distribution.