[A version of this manuscript has been published in *Environrnental Modeling with GIS*, M.F. Goodchild et al., eds., Oxford University Press *(1993) 432-437]*

SPATIAL SIMULATION: ENVIRONMENTAL APPLICATIONS

Evan J. Englund
Environmental Monitoring Systems Laboratory
U. S. Environmental Protection Agency
Las Vegas, Nevada 89193-3478

ABSTRACT

GIS Systems excel at manipulating spatial information, but may not always adequately reflect the level of uncertainty associated with that information. Spatial simulations, conditioned to honor existing data as well as a variogram model, can provide both qualitative and quantitative evaluations of spatial uncertainty in interpolated data. Display of several alternate simulations can graphically illustrate the degree of uncertainty to the data interpreter. The use of alternate simulations as input to other models provides a mechanism for quantifying the sensitivity of complex systems to uncertainty or spatial variability.

INTRODUCTION

Spatial simulation is a geostatistical technique which has great potential as a tool for dealing with the various problems associated with spatial uncertainty. The method has seen only sporadic use since its practical implementation in the early 1970’s (Journel, 1974). It is noteworthy that the recent geostatistics text by Isaaks and Srivastava (1989) does not mention the technique. This lack of attention may have been due to the amount of computer time required - typically one or two orders of magnitude greater than that for kriging and contouring the same area. Fortunately, improved algorithms and fast desktop computers have combined to make the method feasible (each simulation shown below would require about 2 minutes on a 486-25 PC).

In this paper, two spatial data layers are created by contouring irregularly spaced sample data sets. The layers are then combined in a typical GIS logical operation, and spatial simulation is used to evaluate the resulting combined uncertainty. The two sample sets are drawn from much larger “exhaustive” sets which are examined to assess the validity of the simulation results.

KRIGING, SPATIAL SIMULATION, AND CONDITIONAL SIMULATION

Kriging is a spatial regression interpolation method which provides least squares estimates at unsampled locations. While kriged sulfates, like other regressions, may be good estimators, they are also unrealistically smooth and continuous.

Spatial simulations, conversely, fill in realistic-looking detail, but are poor estimators. Both kriging and simulation are controlled by variograrn models, which quantify the spatial variability of data. Spatial simulation refers to the generation of spatial data through the use of a random number generator, constrained to honor a specified variogram model. Conditional simulation forces the simulated data to also honor a pre-existing set of sample data.

A series of conditional simulations generated with different random seeds from the same initial data and variogram model, can be thought of as equally likely possibilities which might explain the available observed data. A kriged map, providing the “best” estimates at each location, is essentially equivalent to the average of a very large number of simulations. Conditional simulations can be used in a number of ways, from simple displays of the uncertainty of data, to much more complex sensitivity analyses where simulations are used to provide variable input to deterministic models.

AN EXAMPLE

The Intersection of Two Kriged Maps

Consider two variables, V 1 and V2, which have been measured in separate sampling campaigns over the same area. Each has an irregular network of sample locations, as shown in Figure 1. From each data set, variograms were computed and modeled (Figure 2). Kriged estimates of each variable were then made on a regular grid of nearly 20,000 points (Figure 3). The kriging process has taken the original sparse measured data and generated complete area] coverage of estimated values. Now imagine a hypothetical decision-maker who has determined that V 1 and V2 have a negative synergism; though neither variable is of interest by itself, when both have high values there is cause for concern. With both kriged maps as data layers in a GIS environment, it would be a simple matter to combine them and generate a new classification map showing where both V1 and V2 exceed specified limits (Figure 4).

How good is this classification map? Is it likely to be nearly perfect, mediocre, or totally wrong? How can the uncertainty be conveyed to the decision maker in an understandable and useful form?

One way to look at spatial uncertainty is with kriging errors. Although Isaaks and Srivastava (1989) rightly caution against the use of the kriging standard deviation for estimating confidence intervals, it is nevertheless often the only measure of spatial error available. A simple example

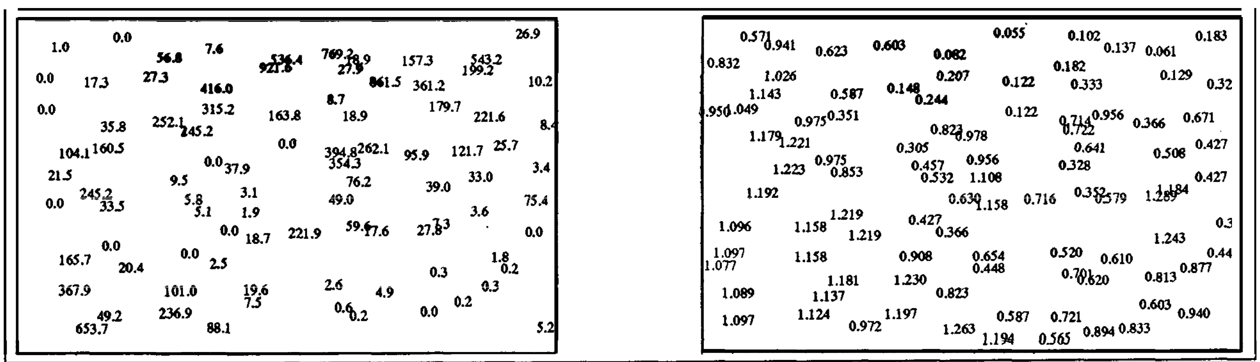


Figure 1 -- Sample locations for variables V1 (left) and V2 (right). Darker shades indicate higher values.

of an error zone for false negatives, based on kriging standard deviations is shown in Figure 5. Expected error rates within this zone range from 50% near the classification boundary to 16% or less at the outer boundary. A similar zone could be created for false positives.

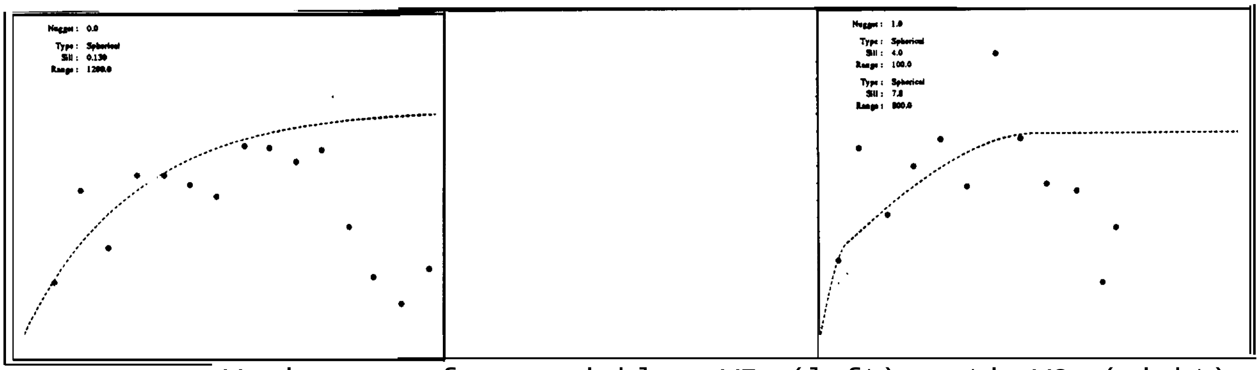


Figure 2 -- Variograms for variables VI (left) anti V2 (right). Observed values are shown as points; fitted models as lines.

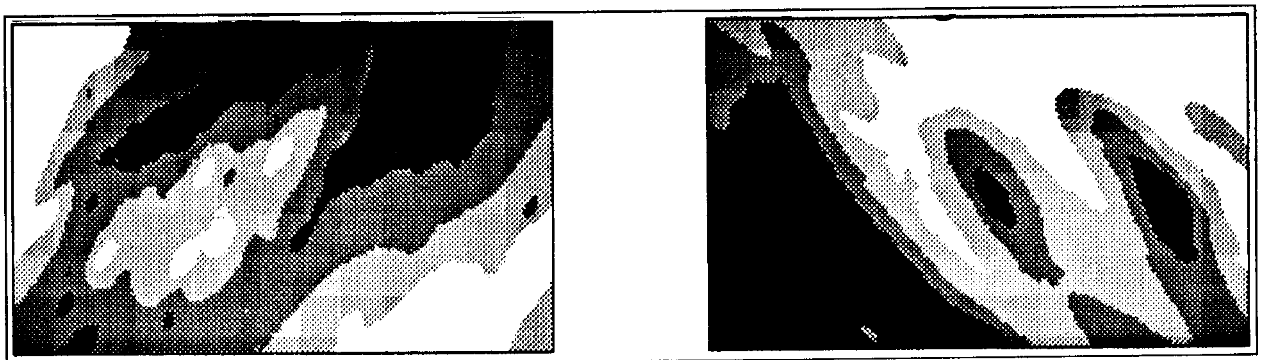


Figure 3 -- Maps of kriged estimates of variables VI (left) and V2 (right) . Darker shades indicate higher estimated values.

A major difficulty with this type of display is that it does not account for spatial autocorrelation of errors, and thus does not provide the decision-maker with a realistic picture. If, say, 20% of the points in the error zone were actually misclassified, there is no way to tell whether they are scattered randomly through the zone, or whether they are strongly clustered to form occasional bulges in the boundary.

Figure 4 -- Classification based on two kriged maps. Dark area indicates where both VI and v2 are estimated to be high.

Intersections of Simulated Maps

Figure 6 shows three conditional simulations of V 1 which honor both the original data in Figure 1, and the variogram model in Figure 2. The only different parameter was the seed for the random number generator. Figure 7 shows three comparable simulations of V2.

While the kriged maps of the two variables appear to be generally similar in terms of the smoothness and complexity of the estimation sulfates, the simulations present a markedly different picture. The V2 simulations retain much of the appearance of the kriged map. The simulated surfaces are rougher, but still fairly continuous, and the shaded bands of the kriged map are generally present in the same places. By contrast, the VI simulations bear little resemblance to the kriged map. The simulations are very discontinuous, and many areas which are high on one simulation are low on another (e.g., compare the lower right-hand corners of the maps in Figure 6).

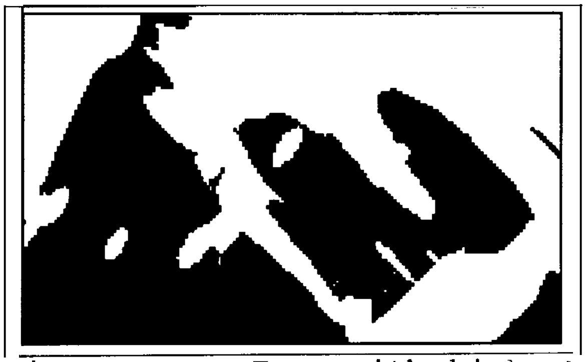


Figure 5 -- Zone with highest probability of false positives

These maps give us an intuitive grasp of the uncertainty of the estimates. If reality were to turn out to be like any of these simulations, it is obvious that the V2 kriged map would be a much better estimator than the V1 kriged map.

Note that the three V 1 simulations, while they differ considerably in detail, they share a common distinctive appearance which we could call “texture”. The same is true for the V2 simulations. This is basically a graphical illustration of the information contained in the variogram models, which describe the variability at all spatial scales, and control the simulation process.

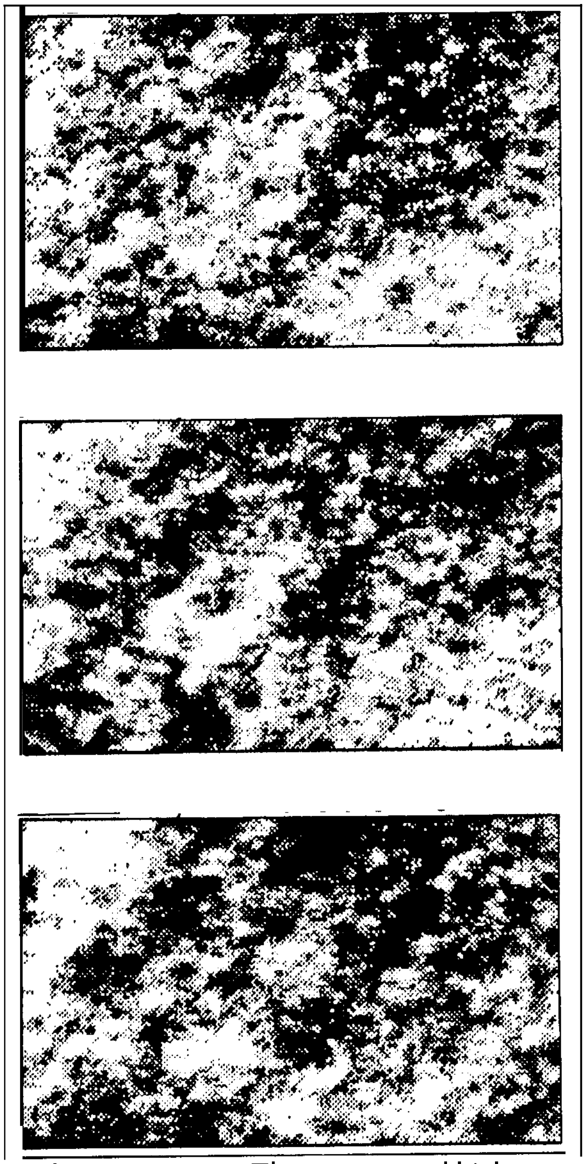


Figure 6 -- Three conditions simulations of VI .

interestingly, geostatistical simulation is a method which contains fractal simulation as a subset, but is more flexible. According to Burroughs (1983), the slope of a variogram model plotted on log-log axes can be translated directly into fractal (Hausdoff-Besicovitch) dimension. A linear variogram model on log-log axes would therefore indicate constant fractal dimension, and lead to simulations with similar textures at all scales. Neither variogram model used in this example is linear on the log-log plot, resulting in different textures at different scales.

The simulations by themselves give us an idea of the spatial uncertainty of each variable. By combining pairs of simulations in the same way we combined the kriged maps, we can look at the uncertainty in our original classification. Figure 8 shows the resulting three simulated

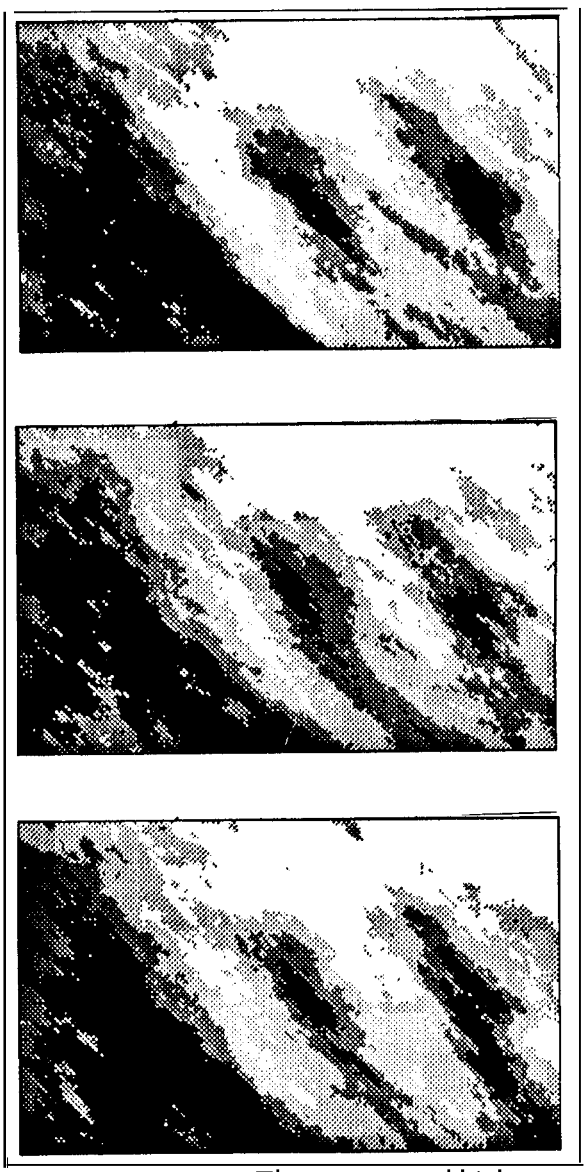
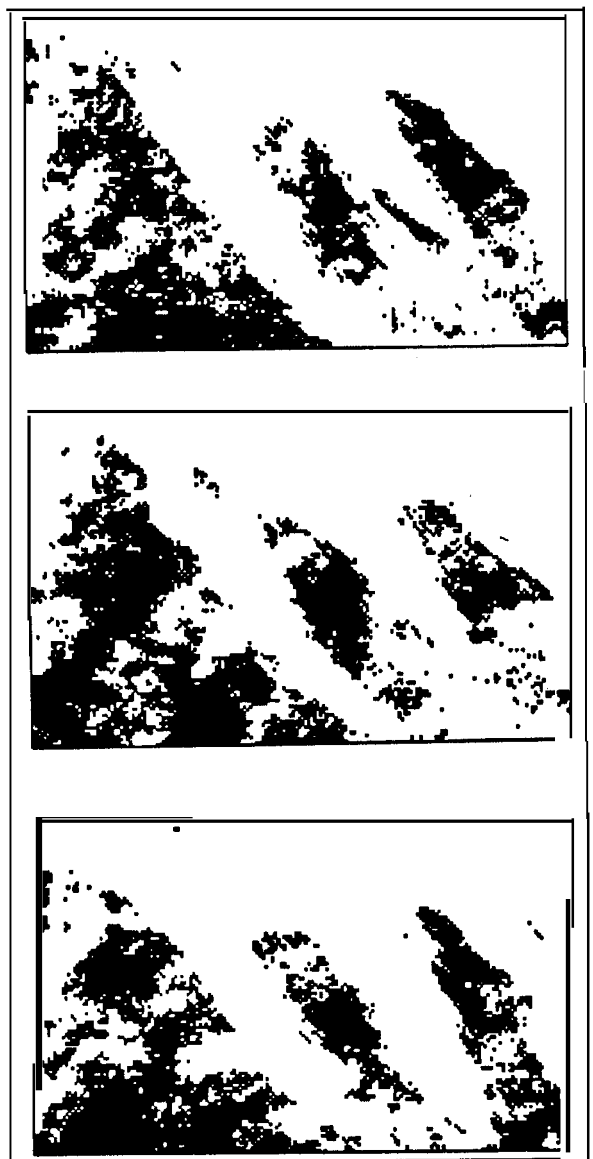


Figure 7 -- Three conditions simulations of V2 .



Fifure 8 -- Three simulated classification maps . Each is based on a pair of conditional simulations . Dark areas indicate where both VI and V2 were simulated to be high.

classification maps. The decision-maker looking at a series of such maps and realizing that reality might be like any one of them, will have a much better idea whether or not the current information is adequate for the decision.

Figure 9 shows the true distributions of V 1 and V2, and the true classification map based on their intersection. Note that the simulated maps do not perfectly represent all of the features observed in the actual maps; for example, the simulations of V2 appear to be more variable than the real thing. In Figure 2, the type of model which was fitted to the experimental V2 variogram is one which implies that the data represent a continuous but relatively rough surface. The experimental variogram, however, is so noisy that many alternate models could be considered equally valid. An investigator with prior knowledge that the measured phenomenon has a continuous but relatively smooth surface could have used a more appropriate model and produced more realistic simulations.

Figure 9 -- Actual maps of V1 (top), V2 (center), and Classification (bottom).

QUANTIFYING THE CONSEQUENCES OF UNCERTAINTY

When visualization of uncertainty alone is not sufficient for the decision-maker, the simulations may be used to make quantitative estimates of errors and their consequences. In the example above, a sensitivity analysis could be conducted by assuming that decisions were to be made from the existing kriged classification map. Then, for any one of the simulations, the number of false positives and false negatives which would occur if that simulation were reality, can be computed. Repeating this process on say, 100 simulations will give both the expected numbers of errors, as well as best case and worst case numbers.

If it is possible to estimate the costs associated with incorrect decisions, simulations can be used in a more elaborate procedure to determine the optimal amount of additional data required. A number of additional samples can be drawn from a simulated model and added to the conditioning data; a new kriged estimate made; and the benefits due to better decisions can be compared to the costs of obtaining the additional data. The process can be repeated with different sampling schemes until the most cost-effective one is found. While such an approach is computer intensive, it is possible on any desk-top computer system capable of serious GIS applications.

DISCUSSION

Both kriging and conditional simulation require the same input - sample measurements with spatial coordinates, and a variogram model. Kriging is becoming more frequently used in GIS applications, and some standard GIS packages already have, or are adding kriging options. Thus, there is no inherent reason why conditional simulation should not be used as routinely for uncertainty analysis as kriging is used for interpolation. It is unlikely, however, that this will occur in the GIS environment until a substantial demand has been established. As happened in the case of kriging, this is likely to require the gradual accumulation of case studies in the literature, as well as wider availability of textbooks, training, and user-friendly software.

REFERENCES

Isaaks, E. H., and Srivastava, R. M., 1989, Applied Geostatistics, Oxford Univ. Press, 561 p.

Journel, A.G, 1974, Geostatistics for conditional simulation of ore bodies. Econ. Geol. 69:5:673-687.

Palmer, M.W., 1988, Fractal geometry: a tool for describing spatial patterns of plant communities, Vegetatio 75:91-102.