Mathematical Geology, Vol. 26, No. 1, 1994

Practical Considerations in the Application of Simulated Annealing to Stochastic Simulation[1]

Clayton V. Deutsch[2] and Perry W. Cockerham[2]

Realizations generated by conditional simulation techniques must honor as much data as possible to be reliable numerical models of the attribute under study. The application of optimization methods such as simulated annealing to stochastic simulation has the potential to honor more data than conventional geostatistical simulation techniques. The essential feature of this approach is the formulation of stochastic imaging as an optimization problem with some specified objective function. The data to be honored by the stochastic images are coded as components in a global objective function. This paper describes the basic algorithm and then addresses a number of practical questions: (1) what are the criteria for adding a component to the global objective function? (2) what perturbation mechanism should be employed in the annealing simulation? (3) when should the temperature be lowered in the annealing procedure? (4) how are edge/border nodes handled? (5) how are local conditioning data handled? and (6) how are multiple components weighted in the global objective function?

KEY WORDS: conditional simulation, cosimulation, edge effects, objective functions, simulated annealing.

INTRODUCTION

In the "annealing" approach to stochastic simulation there is no explicit random function model (Haldorsen and Damsleth, 1990). Rather, the creation of a realization is formulated as an optimization problem to be solved with a numerical optimization technique (seminal references for the application of these techniques to spatial problems include Farmer (1992), Geman and Geman (1984), Kirkpatrick et al. (1983), and Metropolis et al. (1953)). The first requirement of this class of methods is an objective (or energy) function which is some measure of difference between the desired spatial characteristics and those of a candidate realization.

The global optimization technique most often used to obtain such realiza-

[1]Received 6 February, 1993; accepted 12 April, 1993.

[2]Exxon Production Research Co., P.O. Box 2189, Houston, Texas 77252-2189.

tions is based on an analogy with the physical process of annealing. Annealing is the process by which a material undergoes extended heating and is slowly cooled. Thermal vibrations permit a reordering of the atoms/molecules to a highly structured lattice, i.e., a low energy state. In the context of 3-D numerical modeling, the *annealing* process may be simulated through the following steps:

1. An initial 3-D numerical model (analogous to the initial melt in true annealing) is created by assigning a random value at each grid node by drawing from the population distributed. Note that this initial spatial distribution could be constructed so that it already shares additional features of the desired spatial distribution.
2. An energy function (analogous to the Gibbs free energy in true annealing) is defined as a measure of difference between desired spatial features and those of the realization. For example, the energy or objective function could be the squared difference between the semi-variogram of the realization and a model semi-variogram.
3. The image is perturbed by swapping pairs or sets of values taken at random locations in the 3-D numerical model (this mimics the thermal vibrations in true annealing).
4. The perturbation (thermal vibration) is accepted if the energy is decreased; it is accepted with a certain probability if the energy is increased (the Boltzmann probability distribution of true annealing). Technically the name "simulated annealing" only applies when the acceptance probability is based on the Boltzmann distribution (Aarts and Korst, 1989; Kirkpatrick et al., 1983). Through common usage, however, the name "annealing" is used to describe the entire family of methods that are based on the principle of stochastic relaxation.
5. The perturbation procedure is continued while reducing the probability with which unfavorable swaps are accepted (lower the temperature parameter of the Boltzman distribution) until a low energy state is achieved.

Low energy states correspond to plausible 3-D numerical models of the reservoir.

At first glance this approach appears terribly inefficient; millions of perturbations may be required to obtain an image having the desired spatial structure. These methods, however, are more efficient than they might seem as long as few arithmetic operations are required to update the objective function after a perturbation; virtually all conventional spatial statistics (e.g., covariances/correlations) may be updated locally (considering a few locations) after a perturbation rather than globally recalculated (considering all locations).

The objective function is defined as some measure of difference between a set of reference properties and the corresponding properties of a candidate re-

alization. The reference properties could consist of any *quantified* geological, statistical, or engineering property. Some examples include two-point transition probabilities (Farmer, 1992; Deutsch and Journel, 1991), seismic data (Doyen, 1988), multiple-point statistics (Deutsch, 1992a), and well test-derived effective properties (Deutsch, 1992b). Traditional two-point semi-variogram/covariance functions and correlation coefficients with secondary attributes will be considered in this paper. The real advantage of annealing, however, is the ability to integrate many disparate sources of data.

The public domain *sasim* source code documented in (Deutsch and Journel, 1992) was taken and modified for the examples presented below.

FAST UPDATING

Annealing techniques rely on many perturbations to achieve a final acceptable realization. This implies that each component of the objective function must be reasonably simple for fast updating after each perturbation. Also, there should not be too many components nor conflicting components.

For example, consider the semi-variogram at a specific lag **h**:

$$\gamma(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [z(\mathbf{u}) - z(\mathbf{u} + \mathbf{h})]^2 \qquad (1)$$

The number of pairs $N(\mathbf{h})$ contributing to that semi-variogram value depends on the size of the grid and the separation vector **h**. In the case of a unit vector **h** in the x-direction of a square $nx = 200$ by $ny = 200$ grid, a global calculation of $\gamma(\mathbf{h})$ would require an evaluation of $N(\mathbf{h}) = (nx - 1) \cdot ny = 39800$ pairs of values when only two pairs in $\gamma(\mathbf{h})$ must be updated. This is illustrated on Fig. 1 with a small 5 by 5 grid: regardless of how large the grid is, the updating of $\gamma(\mathbf{h})$ for a specific lag requires an updating of only two of the pairs that contribute to that lag.

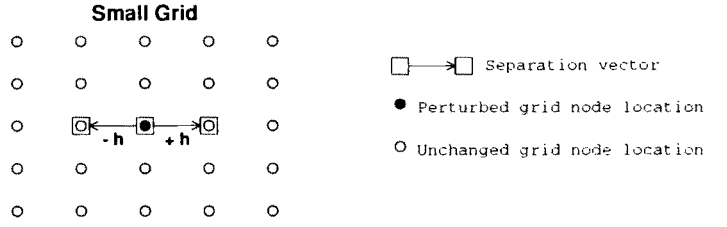

Fig. 1. An example of the two pairs that require updating after perturbing a grid node location. Only the two lags ($+\mathbf{h}$ and $-\mathbf{h}$) require updating regardless of how large the grid network.

When the value of one node location $z(\mathbf{u})$ is perturbed to $z'(\mathbf{u})$ the semivariogram is updated to $\gamma_{new}(\mathbf{h})$ from $\gamma_{old}(\mathbf{h})$ by:

$$\gamma_{new}(\mathbf{h}) = \gamma_{old}(\mathbf{h}) - [z(\mathbf{u}) - z(\mathbf{u} + \mathbf{h})]^2 - [z(\mathbf{u} - \mathbf{h}) - z(\mathbf{u})]^2 + [z'(\mathbf{u}) - z(\mathbf{u} + \mathbf{h})]^2 + [z(\mathbf{u} - \mathbf{h}) - z'(\mathbf{u})]^2 \quad (2)$$

Even when the objective function is updated efficiently the large number of perturbations required by annealing techniques *can* consume a vast amount of CPU time. Therefore, it is desirable to restrict the number and complexity of components entering the objective function. This may, however, be contrary to the goal of integrating a maximum amount of prior data. A slow technique is preferred if it accounts for more prior information. Given the exponential increase in computer speed over the past few years and projected future increases there is no reason to limit the development of annealing programs to overly simplistic objective functions.

PERTURBATION MECHANISM

The annealing algorithm calls for the image to be perturbed in a way that simulates thermal agitation. Swapping the values of two randomly chosen nodal locations is the perturbation mechanism presently most commonly used in geostatistical applications (Deutsch, 1992a; Farmer, 1992). A number of other perturbation mechanisms have been considered such as swapping sets of values or changing any single grid node value according to some local probability distribution function.

The following perturbation mechanisms did *not* lead to an improvement over the standard approach:

1. Randomly select a node and consider replacing its value with a random selection from the global histogram. There is no direct control on the histogram of the final realization and no speed or convergence advantages were obtained.
2. During real annealing, the entire alloy is at the same temperature and thermal agitations occur simultaneously throughout the alloy. In simulated annealing, however, we can process only one agitation/perturbation at a time. By choosing both nodal locations at random, some areas could remain unchanged longer than other areas. To ensure that all nodes are visited more or less regularly we could select the first node in the pair to be swapped according to a random path that sequentially visits each node once. A new random path would then be established after visiting all the nodes at least once. The authors have not found any speed or convergence advantages.

The recommended perturbation mechanism is to swap values at two randomly selected grid node locations (conditioning data are not selected, see later section).

ANNEALING SCHEDULE

Another practical concern when applying simulated annealing is "when do we change the temperature?" The timing and magnitude of the temperature reductions are jointly known as the *annealing schedule*. The following empirical annealing schedule has been used by others (Farmer, 1992; Press et al., 1986) and the authors with success.

The initial objective function is rescaled to 1.0 so that a dimensionless annealing schedule can be used. The total number of nodes (n_{nodes}) is used in the following definitions to keep them independent of the grid size. The concept is to start with an initial high temperature t_0 and lower it by some multiplicative factor λ whenever enough perturbations have been accepted ($K_{\text{accept}} \cdot n_{\text{nodes}}$) or too many have been tried ($K_{\max} \cdot n_{\text{nodes}}$). The algorithm is stopped when efforts to lower the objective function become sufficiently discouraging, i.e., when $K_{\max}$ has reached a certain number (S) or the objective function has reached a low value $O_{\min}$. Table 1 gives parameters for typical annealing schedules.

CORRECTING EDGE EFFECTS

Figure 2 illustrates typical artifacts that can be generated by annealing-based simulation approaches. These *edge effects* are particularly noticeable when using a fast annealing schedule or when the univariate distribution is skewed; the univariate distribution for the realization shown on Fig. 2 is lognormal with mean 10 and a high coefficient of variation of 1.5.

Extreme values are pushed to the edges because, in the standard way of computing semi-variograms, they contribute only once to the semi-variogram calculation, as opposed to a value in the center of the grid which enters the semi-variogram calculation twice,

$$[z(\mathbf{u}) - z(\mathbf{u} + \mathbf{h})]^2 + [z(\mathbf{u}) - z(\mathbf{u} - \mathbf{h})]^2 \tag{3}$$

Table 1. Recommended Dimensionless Annealing Schedules

Schedule	t_0	λ	$K_{\max}$	K_{accept}	S	$O_{\min}$
Quench	0.0	0.0	—	—	—	0.001
Very fast	0.5	0.01	10	2	3	0.001
Fast	1.0	0.05	50	5	3	0.001
Default	1.0	0.10	100	10	3	0.001

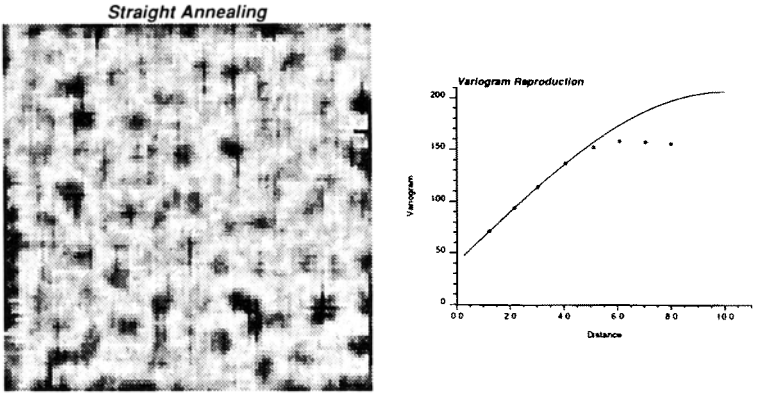

Fig. 2. A realization of a lognormal field (mean = 10, standard deviation = 15) generated with no allowance for weighting the semi-variogram pairs near the border. The 2-D realization is 100 grid units on each side. The isotropic semi-variogram reproduction is shown on the right.

Thus, the objective function is lowered by swapping any extreme value from the center of the grid to the borders of the grid. Edge effects in one form or another are also encountered in the physical process of annealing.

An obvious fix is to "wrap" the grid so that a vector that leaves the grid on one side enters on the other. This is done routinely in geophysics and image processing. The problem, however, is that the spatial distribution becomes correlated across the edges making it more difficult to model border trend effects. Another way of correcting this problem is to increase the weight on the pairs where the corresponding $-\mathbf{h}$ or $+\mathbf{h}$ vector is not in the grid.

The semi-variogram for a specific lag is weighted as follows:

$$\gamma(\mathbf{h}_l) = \frac{\sum_{i=1}^{nl} w_i \cdot [z(\mathbf{u}_i) - z(\mathbf{u}_i + \mathbf{h}_l)]^2}{2 \cdot \sum_{i=1}^{nl} w_i} \tag{4}$$

where both $z(\mathbf{u}_i)$ and $z(\mathbf{u}_i + \mathbf{h}_l)$ are in the grid, w_i is 1.0 if $z(\mathbf{u}_i - \mathbf{h}_l)$ is in the grid, and w_i is some user specified weight greater than 1.0 if $z(\mathbf{u}_i - \mathbf{h}_l)$ is outside the grid. Intuitively, this weight should be 2.0, however, we found a value of $w_i = 4.0$ to be more efficient. The higher value of w_i accounts for the effect of corner regions where values enter the objective function less than half of the time; for a vector at 45° there are *no* pairs to weight.

A realization with this border weighting is shown on Fig. 3. Note that the semi-variogram is no longer reproduced exactly because pairs are weighted differently in the simulation and in the semi-variogram calculation.

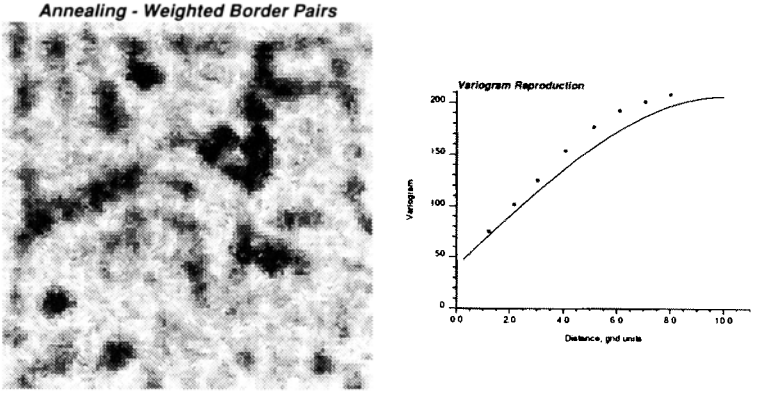

Fig. 3. A realization generated under the same conditions as those shown on Fig. 2 except the border pairs have been weighted. The semi-variogram reproduction is shown on the right. The reproduction is not exact since the border pairs were weighted during the simulation but not weighted when the experimental semi-variogram was calculated.

CONDITIONING TO LOCAL DATA

One simple way to honor local data in the simulated realizations is to never perturb them. This scheme often produces discontinuities next to the conditioning data. This occurs because the conditioning data do not contribute equally to the global objective function since they have no opportunity to be perturbed.

One proposed correction (Deutsch and Journel, 1991, 1992) is to split the objective function into two parts:

$$O = \sum_{l=1}^{n_l} \{[\gamma_s(\mathbf{h}_l) - \gamma_{\text{model}}(\mathbf{h}_l)]^2 + [\gamma_c(\mathbf{h}_l) - \gamma_{\text{model}}(\mathbf{h}_l)]^2\} \quad (5)$$

where n_l is the number of lags in the objective function, $\gamma_{\text{model}}(\mathbf{h}_l)$ is the model semi-variogram, and $\gamma_s(\mathbf{h}_l)$ and $\gamma_c(\mathbf{h}_l)$ are the semi-variograms of the realization. The subscript s denotes pairs of data where each endpoint is a simulated node and the c subscript indicates pairs where at least one endpoint is a conditioning data.

This approach unduly constrains the outcomes near the conditioning data. A more satisfactory approach is to maintain a one-part objective function:

$$O = \sum_{l=1}^{n_l} [\gamma(\mathbf{h}_l) - \gamma_{\text{model}}(\mathbf{h}_l)]^2 \quad (6)$$

where the pairs contributing to the semi-variogram value $\gamma(\mathbf{h}_l)$ are weighted according to whether or not conditioning data are included in the pair, similar to Eq. (4). In this case, $w_i = 1.0$ if both $z(\mathbf{u}_i)$ and $z(\mathbf{u}_i + \mathbf{h}_l)$ are simulated

nodes and $w_i > 1$ if either $z(\mathbf{u}_i)$ or $z(\mathbf{u}_i + \mathbf{h}_l)$ are conditioning data. Testing to date has shown that no apparent discontinuities are generated with $w_i \approx 15$ for pairs that include conditioning data.

An example is shown on Fig. 4. The top illustration shows two strings of conditioning data; the gray is sandstone and the black is shale. The middle illustration shows a stochastic realization generated by straight annealing (without a two-part objective function). Note the discontinuities next to the data locations. The bottom illustration shows a stochastic realization generated where the pairs involving a conditioning datum have been weighted by $w_i = 15.0$.

WEIGHTING COMPONENT OBJECTIVE FUNCTIONS

In general, the objective function O is made up of the weighted sum of C components:

$$O = \sum_{c=1}^{C} w_c O_c \tag{7}$$

where w_c and O_c are the weights and component objective functions, respectively. The component objective functions measure how certain features of the simulated image differ from the desired control or reference properties. For example, one component could be a measure of difference between a semi-variogram model and the semi-variogram of the realization, a second component could measure reproduction of a specific short scale spatial pattern quantified by a particular four-point (quadrivariate) statistic, while a third component could measure the fidelity to a secondary attribute through reproduction of a correlation coefficient.

Each component objective function O_c could be expressed in widely different units of measurement. For example, a component measuring semi-variogram departure may be in units of variance squared (>1000), while a component measuring the reproduction of a correlation coefficient may be quite small (<0.5). Weighting components equally causes the component with the largest magnitude to dominate the global objective function.

The weights w_c allow equalizing the contributions of each component in the global objective function. All decisions of whether to accept or reject a perturbation are based on the change to the objective function,

$$\Delta O = O_{\text{new}} - O_{\text{old}}, \text{ with}$$

$$\Delta O = \sum_{c=1}^{C} w_c [O_{c_{\text{new}}} - O_{c_{\text{old}}}] = \sum_{c=1}^{C} w_c \Delta O_c \tag{8}$$

The weights w_c, $c = 1, \ldots, C$ are established so that, on average, each component contributes equally to the change in the objective function ΔO. That

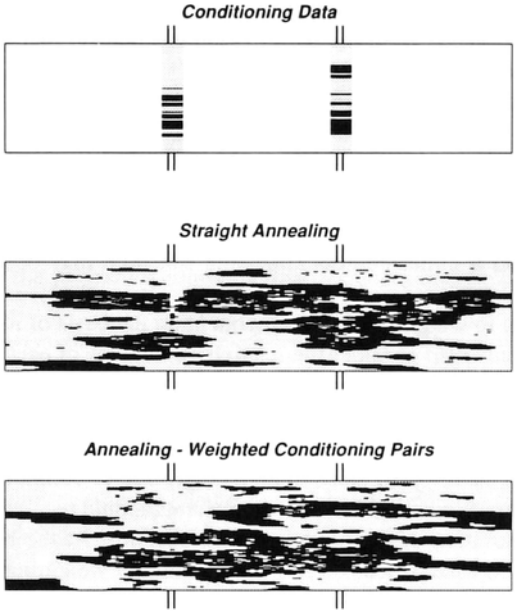

Fig. 4. The top illustration shows two strings of conditioning data; the gray represents sandstone and the black represents shale. The middle illustration shows a stochastic realization generated by straight annealing (the conditioning data are simply not perturbed). Note the discontinuities next to the well locations. The bottom illustration shows a stochastic realization generated where the pairs involving a conditioning datum have been weighted by 15.

is, each weight w_c is made inversely proportional to the average change in absolute value of its component objective function:

$$w_c = \frac{1}{|\Delta O_c|}, \qquad c = 1, \ldots, C \tag{9}$$

In practice, the average change of each component $\overline{|\Delta O_c|}$ may not be computed analytically; it can, however, be numerically approximated by evaluating the average change due to a certain number M (say 1000) of independent perturbations:

$$\overline{|\Delta O_c|} = \frac{1}{M} \sum_{m=1}^{M} |O_c^{(m)} - O_c|, \qquad c = 1, \ldots, C \tag{10}$$

where $\overline{|\Delta O_c|}$ is the average change for component c, $O_c^{(m)}$ is the perturbed

objective value, and O_c is the initial objective value. Each of the M perturbations $m = 1, \ldots, M$ arises from the perturbation mechanism that will be employed for the annealing simulation.

The overall objective function is finally written as,

$$O = \frac{1}{O^{(0)}} \cdot \sum_{c=1}^{C} w_c \cdot O_c \tag{11}$$

The objective function O is normalized by its initial value, $O^{(0)}$, so that it always starts at 1.0 and a dimensionless annealing schedule may be used (see earlier section).

In addition to weighting each component on the basis of its average change ($\overline{|\Delta O_c|}$) one could also consider the relative importance of each component. For example, it may be more desirable to have a good semi-variogram match than to reproduce partial quadrivariate information obtained from a questionable training image. Unequal weighting is particularly useful in cases where the objective function cannot be lowered close to zero, e.g., due to conflicting components.

Another question is whether the components should receive the same weight for the entire duration of the optimization procedure. At the beginning, coarse features could be established, e.g., by preferentially weighting large-scale two-point statistics, then, toward the end, small-scale multiple-point statistics could be given more weight to establish the small scale details.

Numerical precision problems should be avoided. Each component objective function O_c and statistic entering the objective function (e.g., $\gamma(\mathbf{h}_l)$) must be stored with enough precision so that round off errors do not accumulate with incremental changes. The authors have found that double precision is necessary (only for updating the statistics, not for the grid of values) on most machines when simulating grid networks with more than one million grid nodes.

As an example of annealing with a two component objective function, consider the case where a secondary variable is available. The secondary variable could be geophysical measurements that are correlated with the attribute being mapped. Any *good* conditional simulation technique must integrate all relevant information. Integrating information from a secondary attribute is possible in an interpolation mode with cokriging approaches (Doyen, 1988; Zhu, 1991) or some type of trend model (Marechal, 1984); it is also possible to perform such integration with annealing in a stochastic simulation mode.

An essential piece of information needed to account for a Y-secondary attribute is a measure of correlation between the Z-primary and Y-secondary attribute. When the Z-Y point to point correlation is very good, stochastic images of the primary Z-attribute will resemble the map of the secondary attribute rescaled to the Z-units.

In many cases, the Z-Y relationship is summarized by the linear correlation coefficient:

$$\rho = \frac{\text{Cov}\,(Z, Y)}{\sigma_Z \cdot \sigma_Y} \tag{12}$$

This correlation coefficient could then be considered as a component in the objective function of an annealing simulation, i.e.,

$$O_c = [\rho_{\text{calibration}} - \rho_{\text{realization}}]^2 \tag{13}$$

Recall that the key criterion for a quantity to enter the objective function is that it must be locally updateable. The correlation coefficient meets this criterion. Consider the following computational formula,

$$\rho = \frac{E\{Z \cdot Y\} - E\{Z\} \cdot E\{Y\}}{\sqrt{(E\{Z \cdot Z\} - E\{Z\} \cdot E\{Z\}) \cdot (E\{Y \cdot Y\} - E\{Y\} \cdot E\{Y\})}} \tag{14}$$

where the expected values are replaced by discrete summations over the N locations in the area of interest, e.g., $E\{Z \cdot Y\}$ is evaluated by $((1/N)\ \Sigma_{i=1}^{N}\ z_i \cdot y_i)$.

The correlation coefficient of the N simulated nodes, $\rho_{\text{realization}}$, is locally updateable by updating each of the five summations involved in Eq. (14). Given a change from z_i^{old} to z_i^{new} the $(1/N)\ \Sigma_{i=1}^{N}\ z_i \cdot y_i$ summation is updated as follows:

$$\left[\sum_{i=1}^{N} z_i \cdot y_i\right]^{\text{new}} = \left[\sum_{i=1}^{N} z_i \cdot y_i\right]^{\text{old}} - [(z_i^{\text{old}} - z_i^{\text{new}}) \cdot y_i] \tag{15}$$

The other four summations are similarly updated. Moreover, if the annealing perturbation mechanism is swapping, only the first term ($\Sigma_{i=1}^{N}\ z_i \cdot y_i$) needs to be updated. The measured y-values (conditioning data) never change and, in the case of swapping, the univariate distribution of the z-values does not change.

Public domain data published with the GSLIB software (Deutsch and Journel, 1992) have been used below to demonstrate the implementation of such a two-component objective function. There are 29 data informing a 2-D square area and 50 by 50 (2500) secondary data over the same area (Fig. 5). The actual underlying distribution of the primary variable is shown on Fig. 6; this is used to check the simulation results. The scatterplot of the 29 collocated primary and secondary data are shown on Fig. 8. The corresponding linear correlation coefficient ($\rho_{\text{calibration}}$) is 0.493. The sample normal scores semi-variogram and its model are shown on Fig. 9.

The conditional simulation exercise is to generate realizations of the primary Z-variable that honor locally the 29 primary data and the previous correlation coefficient. The stochastic realizations should also honor the Z-sample histogram and semi-variogram model. Conventional simulation without accounting for the Z-Y correlation would yield realizations like the two shown on Fig. 10.

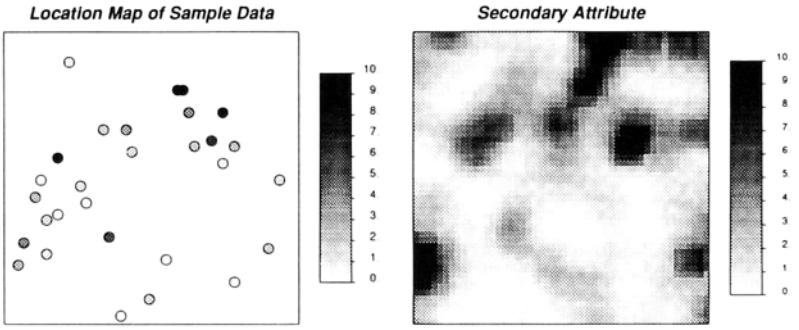

Fig. 5. A map of the 29 primary data locations (left side) and a gray scale map of the 2500 secondary data values.

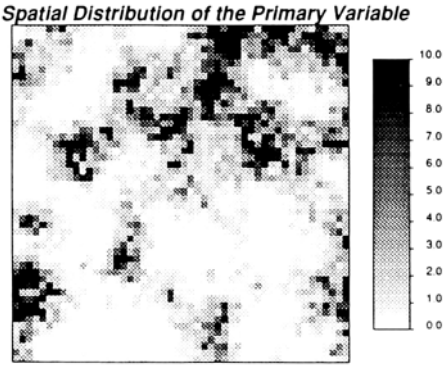

Fig. 6. A gray scale map of the truth.

The *sasim* program was set up to simulate the primary attribute with the objective function:

$$O = \lambda_1 \sum_{i=1}^{N} [\gamma_{\text{model}}(\mathbf{h}_i) - \gamma_{\text{realization}}(\mathbf{h}_i)]^2 + \lambda_2 [\rho_{\text{calibration}} - \rho_{\text{realization}}]^2 \quad (16)$$

The $N = 40$ separation vectors $\mathbf{h}_i$ are defined by the most compact arrangement of 40 lags (see Fig. 7). The $\rho_{\text{calibration}} = 0.493$ is taken from the calibration scatterplot shown on Fig. 8. The weights were established on the basis of the average changes of each component objective function. For the semi-variogram the weight $\lambda_1 = 24.6$ was based on an average change of $\overline{|\Delta O_1|} = 0.0407$; for the correlation coefficient the weight $\lambda_2 = 2481.4$ was based on an average change of $\overline{|\Delta O_2|} = 0.000403$.

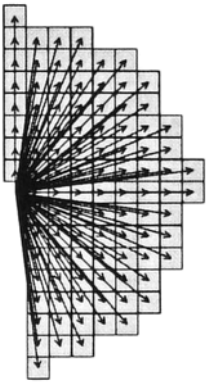

Fig. 7. A polar plot of the lag vectors entering the annealing objective function.

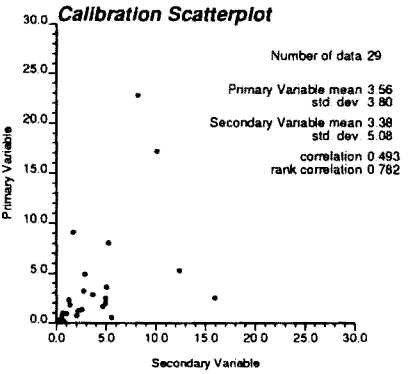

Fig. 8. A scatterplot of the 29 collocated primary and secondary data values.

Two realizations were generated with the objective function (16). Both component objective functions reach zero at the end of the annealing procedure. A useful diagnostic plot is the component objective functions versus the number of attempted swaps (analogous to time). Such a plot for the first realization is shown on Fig. 11. The component objective functions associated to the semivariogram and correlation coefficient start at about 0.7 and 0.3 respectively (due to the scaling). The vertical dashed lines are times at which the temperature is lowered. After about 90,000 swaps both component objective functions reach 0.0.

Figure 12 shows two corresponding simulated realizations. These realizations appear as plausible images of the Z-variable; they are closer to the truth (Fig. 6) than the realizations which did not consider the secondary variable (Fig. 10). These realizations correctly reflect the ring of high values at the top right.

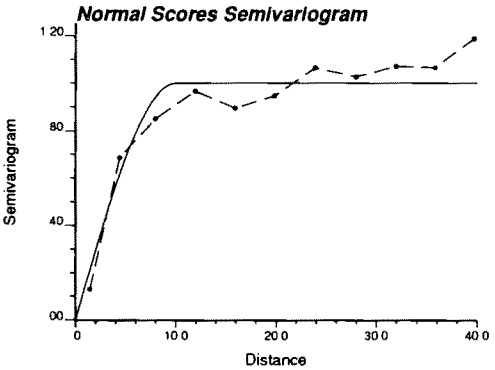

Fig. 9. The normal scores semi-variogram provided with the primary variable (Deutsch and Journel, 1992).

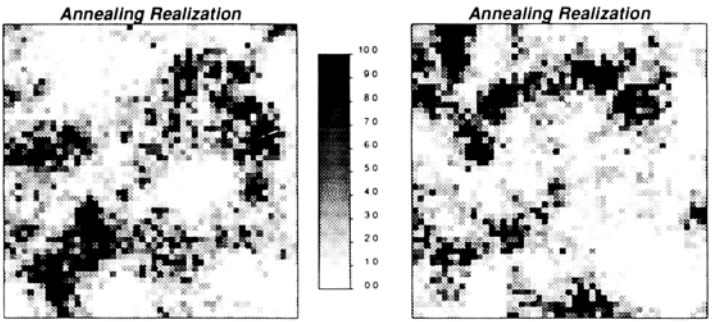

Fig. 10. Two realizations conditional to the primary data only.

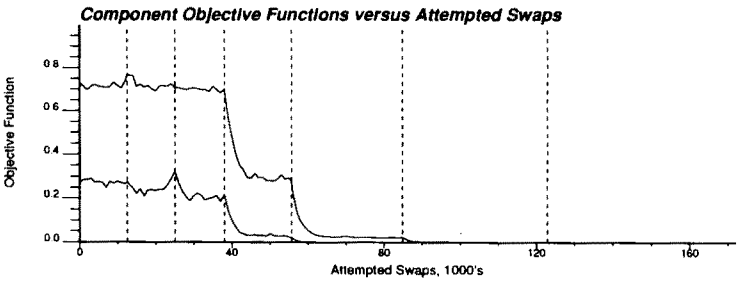

Fig. 11. The component objective functions associated to the semi-variogram (starting at 0.7) and the correlation coefficient (starting at 0.3). The vertical dashed lines denote temperature changes.

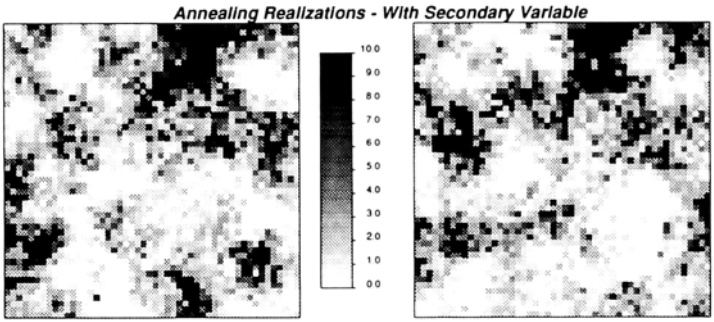

Fig. 12. Two realizations conditional to the primary data and the correlation with the secondary data.

REMARKS AND CONCLUSIONS

The annealing algorithm can integrate many disparate data as long as these data can be quantified to enter a global objective function. Two-point transition probabilities (Farmer, 1992; Deutsch and Journel, 1991), seismic data (Doyen and Guidish, 1990), multiple-point statistics (Deutsch, 1992a), and well test-derived effective properties (Deutsch, 1992b), could be considered as components.

A number of practical problems have been addressed in this paper:

- The statistics included in the objective function should be locally updateable at each step of the perturbation mechanism.
- The perturbation mechanism should be as random as possible. We recommend swapping the values in two randomly selected grid node locations.
- Pair contributions must be weighted near the borders of the grid. This avoids artifact edge effects where extreme values occur more frequently near the edges of the grid.
- Pair contributions where one endpoint is a conditioning datum must also be weighted to ensure that there are no artifact discontinuities near the conditioning data.
- Each component in the objective function must be weighted to allow them an equal opportunity to be reduced to zero in the annealing.

This paper has documented some of the important *tradecraft* issues essential for the successful application of annealing. Although annealing requires some trial-and-error to fine tune the objective function and annealing schedule, it offers unique flexibility in integrating data; a feature not easily achieved with more traditional algorithms.

ACKNOWLEDGMENTS

The authors would like to thank the management of Exxon Production Research Company for permission to publish this paper and Marshall Grant for his helpful comments.

REFERENCES

Aarts, E., and Korst, J., 1989, *Simulated Annealing and Boltzmann Machines*: John Wiley & Sons, New York.

Deutsch, C., 1992a, Annealing Techniques Applied to Reservoir Modeling and the Integration of Geological and Engineering (Well Test) Data: Ph.D. thesis, Stanford University, Stanford, California, 1992.

Deutsch, C., 1992b, Conditioning reservoir models to well test information: Fourth International Geostatistics Congress, Troia, September.

Deutsch, C., and Journel, A., 1991, The application of simulated annealing to stochastic reservoir modeling: J. Pet. Technol., submitted.

Deutsch, C., and Journel, A., 1992, *GSLIB: Geostatistical Software Library and User's Guide*: Oxford University Press, New York.

Doyen, P., 1988, Porosity from seismic data: A geostatistical approach: Geophysics, v. 53, n. 10, p. 1263-1275.

Doyen, P., and Guidish, T., 1990, Seismic discrimination of lithology: A Bayesian approach: Geostatistics Symposium, Calgary, Alberta, May.

Farmer, C., 1992, Numerical rocks, *in* P. King (Ed.), *The Mathematical Generation of Reservoir Geology*: Clarendon Press, Oxford, 1992 (Proceedings of a conference held at Robinson College, Cambridge, 1989).

Geman, S., and Geman, D., 1984, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images: IEEE Trans. Pattern Anal. Machine Intell. PAMI, v. 6, n. 6, p. 721-741.

Haldorsen, H., and Damsleth, E., 1990, Stochastic modeling: J. Pet. Technol., p. 404-412.

Kirkpatrick, S., Gelatt, C., Jr., and Vecchi, M., 1983, Optimization by simulated annealing: Science, v. 220, n. 4598, p. 671-680.

Marechal, A., Kriging seismic data in presence of faults, *in* G. Verly et al. (Ed.), *Geostatistics for Natural Resources Characterization*: Reidel, Dordrecht, Holland, p. 271-294.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and Teller, E., 1953, Equation of state calculations by fast computing machines: J. Chem. Phys., v. 21, n. 6, p. 1087-1092.

Press, W., Flannery, B., Teukolsky, S., and Vetterling, W., 1986, *Numerical Recipes*: Cambridge University Press, New York.

Zhu, H., 1991, Modeling Mixture of Spatial Distributions with Integration of Soft Data: Ph.D. thesis, Stanford University, Stanford, California, 1991.