USMS
023565 The Application of Simulated Annealing to Stochastic Reservoir Modeling

C.V. Deutsch, Stanford U.; A.G. Journel, Stanford U.

Copyright 1991 Society of Petroleum Engineers

This manuscript was provided to the Society of Petroleum Engineers for distribution and possible publication in an SPE journal. The material is subject to correction by the author(s). Permission to copy is restricted to an abstract of not more than 300 words. Write SPE Book Order Dept., Library Technician, P.O. Box 833836, Richardson, TX 75083-3836 U.S.A. Telex 730989 SPEDAL.

REC'D

JUL 0 1 1991

SPE
Technical Publications

SPE 23565

UNSOLICITED

The Application of Simulated Annealing to Stochastic Reservoir Modeling

Clayton V. Deutsch and André G. Journel [1]

Summary

Stochastic reservoir models must honor as much data as possible to be reliable numerical models of the reservoir under study. Traditional stochastic imaging techniques[1] are ill-suited to reproduce complex geological/morphological patterns and engineering data from well tests. The simulated annealing technique offers promise as a complementary tool to incorporate such information into stochastic reservoir models.

This paper presents the current state of *annealing* as applied to reservoir modeling. Comparative studies indicate cases when simulated annealing should be used instead of, or as a complement to, stochastic simulation.

Introduction

Stochastic reservoir modeling is becoming commonly used to describe and visualize reservoir heterogeneities.[1-3] The idea is to generate 3-D images of the reservoir lithofacies and rock properties that, ideally, would honor all available data (core measurements, well logs, seismic and geological interpretations, analog outcrops, well test interpretations, ...). Potentially, there are a large number of plausible realizations that

[1]Stanford Center for Reservoir Forecasting, Applied Earth Sciences, Stanford U.

honor such data. The idea is to retain a few realizations and process them through a flow simulator with the envisaged production scheme. The resulting distribution of important production response variables can then be used for decision making.[4–7]

However, there is no single stochastic modeling algorithm that can simultaneously honor all types of available information. Some algorithms are well suited for discrete or categorical information such as lithofacies types; others are suited for information carried by continuous variables like porosity, saturation, and permeability. Certain information like production data or effective properties derived from well tests, cannot be easily incorporated into the reservoir model. Almost always, a stochastic reservoir modeling exercise will involve a hybrid technique combining the best features of a number of available algorithms.

Simulated annealing is an algorithm initially developed for the solution of combinatorial optimization problems. The type of problem typically considered involves finding the optimum ordering of a system with a large number of components. An optimum ordering is one that minimizes some global cost or objective function. In the context of stochastic reservoir modeling, the components could be a reservoir attribute like porosity defined for blocks of constant size. The cost or objective function could be a measure of how close the ordering (spatial arrangement of the block porosity values) reproduces the pattern of spatial correlation (variogram) inferred from an outcrop study. Finding an *optimum* ordering is equivalent to generating a numerical model.

Recent interest in using the simulated annealing technique for reservoir characterization was triggered by a paper written by C.L. Farmer.[8] The technique capitalizes on two new ideas. First, the imaging problem is set up as an optimization problem. Second, the optimization problem is solved with simulated annealing. This formalism allows the ability to account for diverse types of information by building objective functions more complex than merely identifying a variogram model. For example, one

part of the objective function could be to match statistical properties inferred from data available within the actual reservoir or from other control patterns such as those interpreted from outcrops. A second part of the objective function could be to match an average (effective) property inferred from a well test.

Reasonable care must be taken to limit the complexity of the objective function, otherwise, the computational effort may become too large to obtain a solution in a practical amount of time. There is some evidence to suggest that simulated annealing could be associated as a post-processor to existing stochastic simulation methods allowing the reproduction of information of different types, see hereafter and Deutsch, 1990.[9]

The goal of this paper is to document the implementation details and the current ~~state of~~ simulated annealing as applied to stochastic reservoir modeling. The background and physical principles of simulated annealing are explained. The implementation of Farmer[8] is discussed in some detail. The extension of annealing to *traditional* geostatistical conditioning (i.e., reproduction of the histogram, variogram, and conditioning data) is presented with another algorithm based on a Maximum A Posteriori probability approach.[10,11] A number of important implementation steps and limitations of the annealing algorithm are then uncovered. Some examples are presented using a variety of techniques. Finally, in light of its strong and weak points, the future applications of simulated annealing are discussed.

Annealing Background

Posing stochastic simulation as an optimization problem calls first for a translation of the desired geological, statistical, and engineering properties of the reservoir model into some numerical quantities. Next, reference properties and corresponding numerical quantities must be established from data and/or control patterns. An objective function

is defined as a weighted sum of differences between the properties of any simulated image and the previous reference values. The optimization problem consists of lowering the objective function enough so that the image has all or most of the desired properties.

The solution of such an optimization problem is sometimes possible using the simulated annealing technique. The central idea behind simulated annealing is an analogy with thermodynamics, specifically with the way liquids freeze and crystallize, or metals cool and anneal. At high temperatures the molecules can move freely. As the temperature is slowly lowered the molecules line up in crystals which represent the minimum energy state for the system.

Metropolis and his coworkers[12] developed the idea of numerically simulating molecular behavior. From concepts developed in thermodynamics and statistical physics it is known that a system will change from a configuration of energy E_1 to a configuration of energy E_2 with probability $p = e^{\frac{-(E_2-E_1)}{k_b T}}$. The system will always change if E_2 is less than E_1 (i.e., a favourable step will always be taken); however, it may sometimes take an unfavourable step. The application of this probability distribution in the numerical simulation of systems composed of many parts has come to be known as the Metropolis algorithm. More generally, any optimization procedure that draws upon the thermodynamic analogy of annealing is known as simulated annealing.

In the early 1980's Kirkpatrick et al.[13] and independently Cerny[14] extended these concepts to combinatorial optimization, i.e., they formulated an analogy between the objective function and the free energy of a thermodynamical system.[15,16] A control parameter, analogous to temperature, is used to control the iterative optimization algorithm until a state with a low objective function (energy) is reached.

One of the first direct applications to spatial phenomena was published by Geman and Geman[11] who applied the method to the restoration of degraded images. About the same time Rothman[17] applied the method to nonlinear inversion and residual statics

estimation in geophysics. Independent research by C.L. Farmer[8] led to the publication of a simulated annealing algorithm for the generation of rock type models. This triggered considerable interest in the method among geostatisticians.[9,10,18]

The essential components of an annealing algorithm are the objective function, a procedure to update the objective function, a perturbation mechanism, and some empirical procedure for lowering the temperature or reducing the control parameter (the temperature).

The general algorithm may be described with the following steps:

1. Generate an initial image. This could be the result of a prior simulation algorithm or possibly an image with the nodal values drawn at random from the desired univariate distribution.

2. Establish an initial control parameter c and a schedule for lowering it as the looping progresses. The magnitude of the starting control parameter may be set to the value of the initial objective function.

3. Perturb the image. For example, swap the attribute value assigned to two different grid nodes.

4. Compute a new objective function O_{new}. For example, the objective function could be the squared difference between the variogram of the image and that of a prior model.

5. Establish the acceptance probability distribution:

$$P\{accept\} = \begin{cases} 1, & \text{if } O_{new} \leq O_{old} \\ e^{\frac{O_{old}-O_{new}}{c}}, & \text{otherwise} \end{cases}$$

6. Draw from that probability distribution. If the perturbation is accepted then update the image and reset the objective function $O_{old} = O_{new}$.

7. Return to step 3 until the objective function is low enough or there has not been any significant improvement in many successive iterations.

The control parameter c or *temperature* must not be lowered too fast or else the image may get trapped in a sub-optimal situation and never converge. However, if lowered too slowly then convergence may be unnecessarily slow. The specification of how to lower the control parameter is known as the "annealing schedule". There are mathematically based annealing schedules that guarantee convergence;[11,15] however, they are too slow in practice. The following empirical annealing schedule is one possibility.[8,16]

The idea is to start with an initially high control parameter c_0 and lower it by some multiplicative factor λ whenever enough perturbations have been accepted (K_{accept}) or too many have been tried (K_{max}). The algorithm is stopped when efforts to lower the objective function become sufficiently discouraging. The following parameters describe this annealing schedule:

c_0: the initial control parameter.

λ: the reduction factor $0 < \lambda < 1$.

K_{max}: the maximum number of attempted perturbations at any one value of the control parameter (on the order of 100 times the number of nodes). The control parameter is lowered by λ whenever K_{max} is reached.

K_{accept}: the acceptance target. After K_{accept} perturbations are accepted at any one of its value the control parameter is lowered by λ (on the order of 10 times the number of nodes).

S: the stopping number. If K_{max} is reached S times then the algorithm is stopped (usually set at 2 or 3).

ΔO : a low objective function indicating convergence.

This type of annealing schedule is used for all implementations presented in this paper.

Another implementation corresponds more closely to a Bayesian approach whereby only downhill steps (ones that lower the objective function) are taken; no annealing schedule is needed, see hereafter and, Doyen.[10]

Numerical Rocks

Consider the numerical modeling of integer coded rock types.

The proportion of each rock type is honored at the beginning by randomly assigning rock types at all nodes according to their respective proportions. Wherever the rock type is known, e.g., at the well locations, the rock type is set to its known value and never changed in the subsequent annealing algorithm; thus conditioning to well data is obtained by construction.

A two-point histogram and/or a correlation function for some specified number of lags and directions enter the objective function to control the simulation.

A two-point histogram for a specific lag and direction vector $\mathbf{h}$ specifies the probability of each possible transition. For example, if there are two rock types (1 and 2) there are four possible transitions from the rock type at location $\mathbf{u}$ to the rock type at location $\mathbf{u}+\mathbf{h}$. In general, given n integer coded rock types there are n^2 *classes* in the two-point histogram.

A complete two-point histogram contains all the univariate and bivariate information. More precisely, it contains all the direct and cross indicator covariance information.[6,19] Experimental data may not be enough to infer such a two-point histogram; they could then be computed from *control patterns* obtained from independent measurements (e.g., an outcrop study or a catalog of geological features). The control pattern conditions all spatial features of the simulated realizations; therefore, it must be appropriate.

A possible objective function could be defined as the sum of the squared differences

between the control statistics and the image statistics, for example:

$$O = \left[\sum_{l=1}^{n_{lag}} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \left(t_{ijl} - t^{o}_{ijl} \right)^2}{\sum_{l=1}^{n_{lag}} n_l} \right]$$

where n is the number of rock types; n_{lag} is the number of lags considered in the two-point histogram; t_{ijl} and t^{o}_{ijl} are the numbers of pairs switching from code i to code j at lag vector l for the image and the control pattern respectively; n_l is the number of pairs at lag vector l.

Virtually all of the computational effort is taken to update the objective function; the time needed to choose two locations at random is negligible. Therefore, it is important to code this part of the algorithm with a minimum number of arithmetic operations.

Different realizations, i.e., conditional simulations, are achieved by starting with different random images.

An example control pattern is shown on the upper left of Figure 1. This control pattern represents classes of air permeameter measurements taken on a vertical slab of Berea sandstone of dimensions 2 x 2 foot.[20] The continuous distribution of permeability values was divided into six classes for the purposes of this demonstration. An example random image, 40 by 40 pixels, is shown on the upper right of Figure 1. The two-point histogram, for 15 lags, in two directions aligned with the coordinate axes was computed and retained to control the generation of the simulation shown on the lower left. Although the general features are reproduced, the strong diagonal banding is not well reproduced. The realization shown on the lower right of Figure 1 was created by choosing the two directions parallel and orthogonal to the banding. Choosing suitable directions and lags that enter the objective function is an important aspect of annealing.

Continuous Variable Simulations

Farmer's algorithm may be used to simulate continuous variables by defining classes, as already done for the previous Berea sandstone example. Another alternative objective function is to reproduce *traditional* geostatistical constraints. That is, a conditional simulation is an image that honors a given histogram, variogram model, and the data values at their locations.[6,21]

In the following implementation, the initial image is created by filling all non-data nodes with values drawn at random from a specified histogram. This histogram is preserved in the final images since only the spatial locations of the values are changed, not the values themselves. The conditioning data values are honoured by fixing them in the initial image and never perturbing them in the subsequent optimization. The final requirement of matching the variogram is achieved through the simulated annealing algorithm.

The objective is for the actual variogram $\gamma_{actual}(\mathbf{h})$ to match the variogram model $\gamma_{model}(\mathbf{h})$. The following objective function should approach zero to achieve such a conditional simulation:

$$O = \sum_{\mathbf{h}} \left[\frac{\gamma_{actual}(\mathbf{h}) - \gamma_{model}(\mathbf{h})}{\gamma_{model}(\mathbf{h})} \right]^2$$

There is no requirement to standardize each deviation by the model variogram value at each lag. This is done here to give more weight to closely spaced (small variogram) values.

The initial image is modified by swapping pairs of nodal values z_i and z_j chosen at random, where neither node i nor node j are conditioning data. The annealing schedule documented earlier is used. Note that a control pattern is not needed; the variogram model can be inferred from the available conditioning data.

The more directions and lags considered in the objective function the longer it will

take to calculate and update; consequently, the algorithm will be slower. An efficient coding of the objective function is necessary to speed up the algorithm. For example, one should update rather than recalculate the variogram at each step, i.e., if the value z_j is getting swapped into a specific node and z_i is getting swapped out, the lag at $\mathbf{h}$ is updated by subtracting the previous contribution from z_i and adding the new contribution due to z_j:

$$\gamma_{new}(\mathbf{h}) = \gamma_{old}(\mathbf{h}) + \frac{1}{2N(\mathbf{h})}\left[(z - z_j)^2 - (z - z_i)^2\right]$$

where z is the value at a lag distance $\mathbf{h}$ from the node being altered, and $N(\mathbf{h})$ is the number of pairs separated by lag (vector) $\mathbf{h}$. Of course this updating has to be carried out at two locations: the location where z_i is coming from (where z_j is going) and the location where z_j is coming from (where z_i is going). Further, note that this must be done for all lags and directions called for in the objective function.

When all the pairs are considered without distinction in a one-part objective function, the original local conditioning data are somewhat overwhelmed. That is, although the conditioning data and overall variogram are matched, there is no specific condition to match the variogram involving conditioning data. Experience has shown that the results are more realistic and aesthetically pleasing when a two-part objective function is considered where all the pairs which involve at least one conditioning datum are considered as a second part of the objective function. This is illustrated on Figure 2 where realizations generated with a one part objective function present discontinuities near the wells. Realizations generated with the objective function separated into two components show improved continuity near the wells.

Annealed images can be obtained with a very low objective function, that is, the variogram model can be reproduced very closely. Figure 3 shows the variogram at the beginning, half way through, and at the end of the simulation shown at the lower left of Figure 2.

The indicator formalism[3,21] is useful because it allows specification of the continuity of the extreme values. The objective function could be set up as the cumulative deviation of the actual and modeled indicator variograms. The implementation of Srivastava[18] considers such an objective function. As before, the initial image is generated with the correct histogram and the conditioning data are set in the image, never to be changed. Then, all the nodes are visited in a path spiraling away from the conditioning data. At each simulated node the probability distribution of the attribute value is constructed with indicator kriging.[6] A Monte Carlo realization from that conditional distribution then replaces the node value. The objective function is recalculated only after all nodes have been visited. If the objective function is low enough, or if some maximum number of loops is exceeded, then the image is finished; otherwise, the spiraling path is repeated.

The Maximum A Posteriori Variant

Bayesian classification schemes provide a variant to the simulated annealing algorithm.[10,22,23] While the approach is suitable for discrete or categorical variables, the application to continuous variables is not straightforward. The method described below for generating rock type images is due to the recent work of Doyen.[10]

As in annealing, the algorithm starts with an image that has the conditioning data fixed and the correct proportion of each rock type. A random path is established that sequentially visits each node, much the same as Srivastava's approach.[18] At each node location all the rock types are considered and the one that lowers the objective function the most is kept. This cycling over all the grid nodes is continued until the objective function does not change significantly.

The resulting image is called a Maximum A Posteriori (MAP) lithologic model. The primary difference between this approach and Farmer's algorithm is that the MAP

approach cannot jump out of a local minimum, i.e., it corresponds to a steepest descent-type approach. The rationale is that in stochastic simulation the goal is not a global minimum; the goal is to make a census of images (realizations) that have the statistical, geological, and engineering properties that are considered important.

Figure 4 shows MAP models based on the same control pattern as the numerical rocks example presented earlier (Figure 1). The MAP realization attempting to match the two-point histogram in directions parallel to the coordinate axes fails to reproduce the general banding aspect of the control pattern. Once again the resulting image is much better when the directions are chosen to capture the essential features of the control pattern.

An interesting aspect of Doyen's paper[10] is the incorporation of seismic information. A two part objective function was considered; the first part was the deviation of the rock simulated types from that indicated by a seismic impedance profile and the second part was the deviation of the two-point histogram from that inferred from a control pattern. This two-part objective function yields images that are simultaneously constrained by the seismic data and the statistical properties of the reservoir.

The method is impractical with continuous variables because finding the continuous value which most significantly reduces the global objective function would be prohibitively expensive.

Implementation

As already mentioned, the annealing schedule is quite important. If the image is cooled too quickly then it may *freeze* before convergence, that is, before all conditioning statistics are correctly reproduced. Conversely, if the image is cooled too slowly, besides the issue of computer time there is that of convergence toward a single global optimum which would defeat the objective of providing alternative realizations.

Research is not advanced enough to present an automatic method of obtaining such a schedule. Note that the schedule parameters cannot be changed independently of one another. For example, if the reduction factor λ or *cooling rate* is lowered the stopping number S must be increased, otherwise, the algorithm may stop iterating before convergence.

Experimentation is the only method currently available to establish an appropriate annealing schedule. This can be tedious and time consuming. The main focus of researchers[11,15] is on achieving a more parallel computing approach in hardware rather than fine tuning the annealing schedule. That is, if computers with hundreds or even thousands of processors could be employed the optimum could be reached no matter how slow the cooling schedule.

There are properties analogous to the thermodynamic properties of specific heat and entropy that could be measured to assist in customizing the annealing schedule. For example, the specific heat can be used to recognize when the system is beginning to converge. The temperature should not be lowered when the convergence is rapid.[13]

A rather difficult synthetic example was constructed to test the annealing algorithm. The 200 by 100 pixels control pattern, shown at the top left of Figure 5, consists of many superimposed *geological-like* structures. Most simulation algorithms would have difficulty in simultaneously reproducing all aspects of that image, e.g., the *fault like* structures, the *alteration caps* on the elliptical shapes, and the anisotropic discontinuous bodies entirely within the elliptical shapes.

The two images shown on the bottom of Figure 5 were created by retaining the two point histogram for 25 lags in four directions. The images succeed in capturing many of the control pattern features.

In one of the simpler examples presented earlier (Figure 4) the MAP simulation method worked almost as well as the annealing algorithm. However, in the present

example there are many conflicting goals and the MAP algorithm gets stuck in various local minima before converging to reasonable images. Figure 6 shows two such MAP realizations. This indicates that if the control pattern is complex it may be necessary to implement the full simulated annealing algorithm rather than the simpler and faster MAP simulation method.

The addition of more information into the objective function is one of the major advantages to the annealing algorithm. For example, it is possible to consider all configurations of a 2 by 2 pixel configuration. This 4-point histogram or partial quadrivariate information would add important details about the short scale structure. Such quadrivariate information has been extracted from the control pattern shown on Figure 5. A realization that honors this statistical control is shown on Figure 7. The general character of the control pattern is well reproduced even though only very short scale information has been used.

Comparison of Annealing to Other Methods

The Berea sandstone example[20] presented on Figure 1 may also be approached with a Gaussian simulation and an indicator simulation technique[3,21] Table 1 summarizes the computer requirements for a single 40 by 40 simulated image.

- The sequential Gaussian simulation considered a maximum of 16 previously simulated grid nodes for the derivation of each conditional distribution and subsequent simulation of a grid node. All other parameters do not affect the speed.
- The sequential Indicator simulation used nine threshold values discretizing the range of variability of the permeability and again a maximum of 16 data for the simulation of a grid node.

Sequential Gaussian Simulation	10.9
Sequential Indicator Simulation	41.0
Simulated Annealing: *Farmer's Algorithm*	2332.9
Simulated Annealing: *Farmer's Algorithm* - stripped down	336.6
Simulated Annealing: *MAP algorithm*	302.0
Simulated Annealing: *MAP algorithm* - stripped down	39.1

Table 1: The computer requirements for one 40 by 40 simulated image (seconds on a DEC 5000-200 workstation).

- Farmer's algorithm considered 4 directions and 15 lags in the objective function. The algorithm was executed using the annealing schedule presented earlier.
- The *stripped down* Farmer's algorithm considered only 4 lags in each of the 4 directions. Thus, 16 actual lags were used in the objective function corresponding more closely to the 16 data used with the sequential Gaussian and sequential indicator algorithms. However, the resulting images are not as good as when using the full 15 lags in each direction.
- The MAP algorithm considered 4 directions and 15 lags when selecting the optimum classification at each point.
- The *stripped down* MAP algorithm considered only 4 lags in each of the 4 directions to more closely vorrespond to the sequential Gaussian and sequential indicator algorithms.

In all cases the statistics required, whether variograms or two-point histograms, were derived from the Berea control pattern. No initial data was used.

Figure 8 shows one realization from each of the algorithms using standard softwares.[24] From this figure we would either choose the simulated annealing simulation or the MAP simulation. The advantage of the Gaussian and indicator simulations is that a prior exhaustive control pattern is not necessary.

One can argue that the images on Figure 8 are not directly comparable because they have been generated with two different objectives:

1. The Gaussian and Indicator images are generated with a *stochastic simulation* algorithm which aims at sampling from random function models constrained by specific statistics. These images are intended to visualize the spatial uncertainty left once these statistics have been honoured.

2. The annealing and MAP images are generated with an *optimization* algorithm which aims specifically at identifying the conditioning properties or statistics. It is possible that, whatever the initial random image, these algorithms may zoom toward the same class of final "optimal" images, thus failing to reflect spatial uncertainty, in particular that associated with the reliability of the control pattern used.

Clearly, if best reproduction of one specific control pattern is the goal then optimization techniques are preferable. In reservoir modeling, any control pattern provides but a general mold for the actual reservoir; the goal is *not* to reproduce exactly that image but to evaluate the various realizations possible around that mold. This important distinction between simulation and optimization is being actively investigated.

Further Applications of Annealing

The annealing objective function allows the flexibility to impose constraints that cannot easily be expressed in terms of bivariate statistics. The biggest potential is for annealing

to be used in conjunction with a tailored simulation algorithm, i.e., initially capture as much information as possible with a faster more traditional simulation algorithm and then *finish* the simulations with annealing to capture difficult information like complex geological patterns or production-type data.

One outstanding potential is associated with the inability of stochastic simulation techniques to handle effective properties inferred from well tests. It may be possible to formulate a post-processing annealing algorithm that will locally alter initial simulated images such that their statistical properties are minimally modified while forcing it to match the effective properties inferred from a well test.

Annealing could also be used to extend dimensionality, e.g., geologists can conveniently draw 2-D cross sections; however, a consistent 3-D image from such cross-sections is tedious to hand draw. Annealing could be used to in-fill the complete 3-D distribution conditional to the geological interpretation and statistics available from several 2-D cross sections.

The application of simulated annealing to stochastic simulation is a good example of cross fertilization between different scientific disciplines. There are many other combinatorial optimization problems, outside the realm of stochastic modeling, that may benefit from its application.

Conclusions

The application of simulated annealing to stochastic reservoir modeling has been presented in some detail. A number of examples have shown that the method generates high quality realizations albeit with considerably more computer time that conventional techniques.

The major advantage of annealing is that additional information can be accounted for by simply adding a contribution to an objective function. One example of matching

partial quadrivariate statistics was presented. This work also points the way to other applications such as the incorporation of seismic data, and effective properties inferred from pressure transient well tests.

The current concept is to start with a fast tailored simulation algorithm such as Boolean, Gaussian, or indicator methods to capture the broad features. Then, annealing can be applied as the *finisher* to incorporate information and details not possible with tailored algorithms.

There remain questions about whether simulated annealing realizations succeed to image the spatial uncertainty prevailing beyond the properties (data) contained in the objective function. It could be that these realizations, being based on an optimization principle, are too close to each other offering a biased and non-conservative image of actual uncertainty.

Nomenclature

c	=	control parameter in simulated annealing
O	=	objective function
ΔO	=	a low objective function indicating convergence.
$\gamma(h)$	=	variogram value for direction vector h
$N(h)$	=	number of pairs corresponding to direction vector h
z	=	reservoir attribute (e.g., permeability)
t_{ijl}	=	two-point histogram
T	=	temperature
E	=	Gibbs free energy
k_b	=	Boltzmann constant
c_l	=	correlation for lag l
c_0	=	the initial annealing control parameter.
λ	=	the control parameter reduction factor.
K_{max}	=	maximum number of swaps at given control parameter.
K_{accept}	=	acceptance target number of swaps.
S	=	the stopping number.

Subscripts

i	=	grid node location/index
j	=	grid node location/index
model	=	prespecified model value
actual	=	experimentally calculated value
new	=	function at a new level of iteration
old	=	function at an old level of iteration

Acknowledgments

Funding for this study was provided by the Stanford Center for Reservoir Forecasting (SCRF) and the Alberta Research Council (ARC).

References

[1] Haldorsen, H. and Damsleth, E.: "Stochastic Modeling," *JPT* (April 1990).

[2] Hewett, T.: "Fractal distributions of reservoir heterogeneity and their influence on fluid transport," paper SPE 15386 presented at the 1986 Annual Technical Conference and Exhibition, New Orleans, Oct. 5-8.

[3] Journel, A. and Alabert, F.: "New method for reservoir mapping," *JPT* (Feb. 1990).

[4] Ballin, P.: "Quantile-Preserving Approximations of Complex Transfer Functions: A Waterflood Example," Report 3, Stanford, CA (May 1990).

[5] Emanuel, A., Alameda, G., Behrens, R., and Hewett, T.: "Reservoir performance prediction methods based on fractal geostatistics," *SPE Reservoir Engineering* (August 1989) 311–318.

[6] Journel, A.: *Fundamentals of Geostatistics in Five Lessons*, Short Course in Geology Series, American Geophysical Union, Washington, D.C. (1989) **8**.

[7] Srivastava, R.: "An Application of Geostatistical Methods for Risk Analysis in Reservoir Management," paper SPE 20608 presented at the 1990 Annual Technical Conference and Exhibition, New Orleans, Sept. 23-26.

[8] Farmer, C.: "Numerical Rocks," Joint IMA/SPE European Conference on the Mathematics of Oil Recovery, Robinson College, Cambridge University (July 1989).

[9] Deutsch, C.: "Relative Influence on Flow Performance of Spatial Heterogeneities versus Uncertainty in Relative Permeability Curves," Report 3 - Stanford Center for Reservoir Forecasting, Stanford, CA (May 1990).

[10] Doyen, P. and Guidish, T.: "Seismic Discrimination of Lithology: A Bayesian Approach," Geostatistics Symposium, Calgary, AB (May 1990).

[11] Geman, S. and Geman, D.: "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence* (November 1984) **PAMI-6**, No. 6, 721–741.

[12] Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and Teller, E.: "Equation of state calculations by fast computing machines," *J. Chem. Phys.* (June 1953) **21**, No. 6, 1087–1092.

[13] Kirkpatrick, S., Gelatt Jr., C., and Vecchi, M.: "Optimization by simulated annealing," *Science* (May 1983) **220**, No. 4598, 671–680.

[14] Černy, V.: "Thermodynamical approach to the travelling salesman problem: an efficient simulation algorithm," *Journal of Optimization Theory and Applications* (1985) **45**, 41–51.

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

[16] Press, W., Flannery, B., Teukolsky, S., and Vetterling, W.: *Numerical Recipes*, Cambridge University Press, New York (1986).

[17] Rothman, D.: "Nonlinear inversion, statistical mechanics, and residual statics estimation," *Geophysics* (1985) **50**, 2784–2796.

[18] Srivastava, R.: *INDSIM2D: An FSS International Training Tool*, FSS International, 3900 Quebec Street, Vancouver, BC, Canada (1990).

[19] Isaaks, E. and Srivastava, R.: *An Introduction to Applied Geostatistics*, Oxford University Press, New York (1989).

[20] Giordano, R., Salter, S., and Mohanty, K.: "The effect of permeability variations on flow in porous media," paper SPE 14365 presented at the 1985 Annual Technical Conference and Exhibition, Las Vegas, Sept. 22-25.

[21] Journel, A. and Alabert, F.: "Non-Gaussian data expansion in the Earth Sciences," *Terra Nova* (1989) **1**, 123–134.

[22] Besag, J.: "On the statistical analysis of Dirty Pictures," *J.R. Statistical Society B* (1986) **48**, No. 3, 259–302.

[23] Andrews, H. and Hunt, B.: *Digital Image Restoration*, Prentice Hall Inc., Englewood Cliffs, N.J. (1989).

[24] Deutsch, C. and Journel, A.: *GSLIB: Geostatistical Software Library*, Accepted for Publication by Oxford University Press, New York (1991).

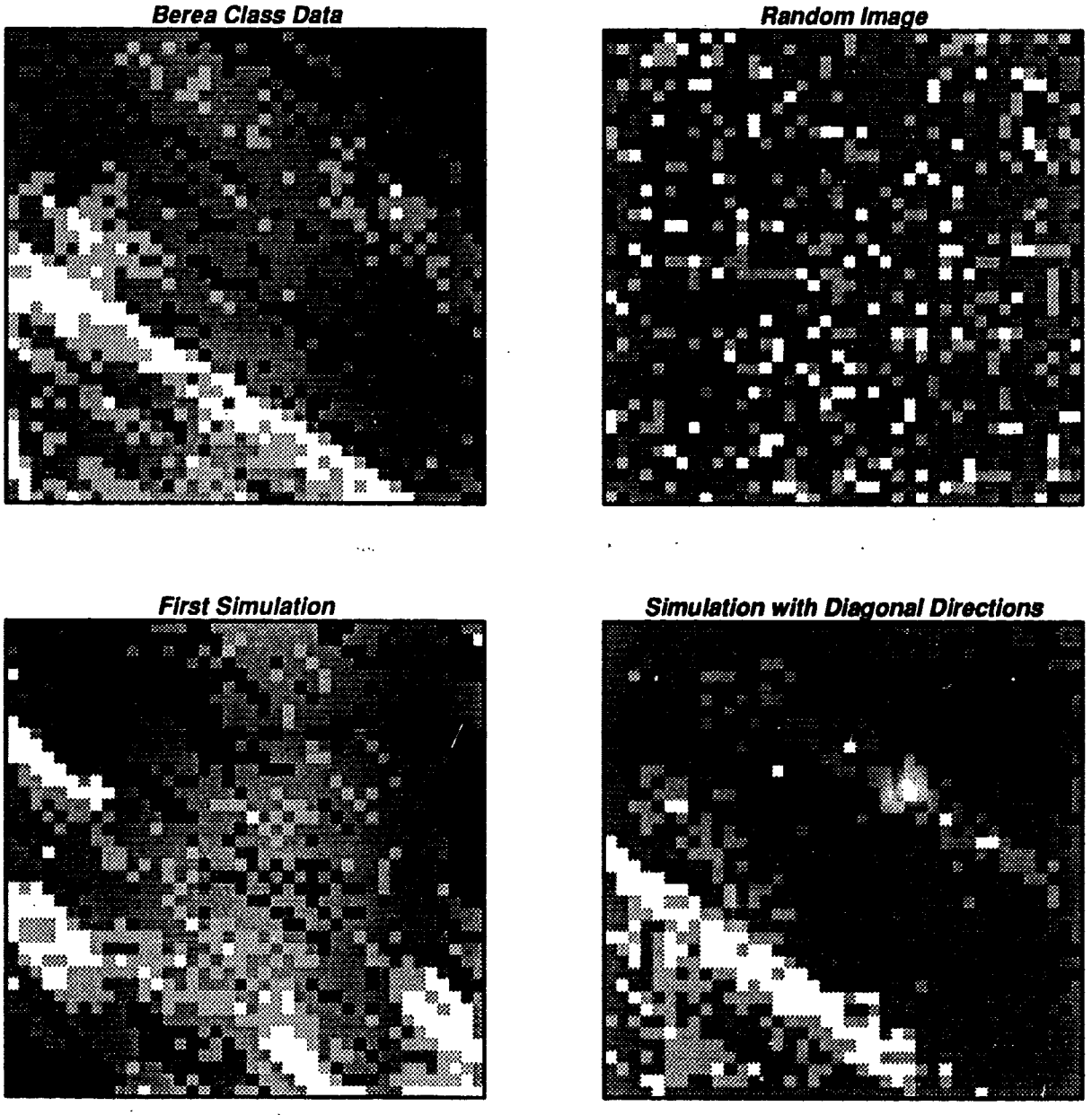

Figure 1: An example application of Farmer's algorithm: the upper left image is a 40 by 40 pixel control pattern; the upper right image is an initial random image; the lower left image was created by retaining the two-point histogram, for 15 lags, in the two coordinate directions; the lower right image was created by choosing the two directions parallel and orthogonal to the banding.

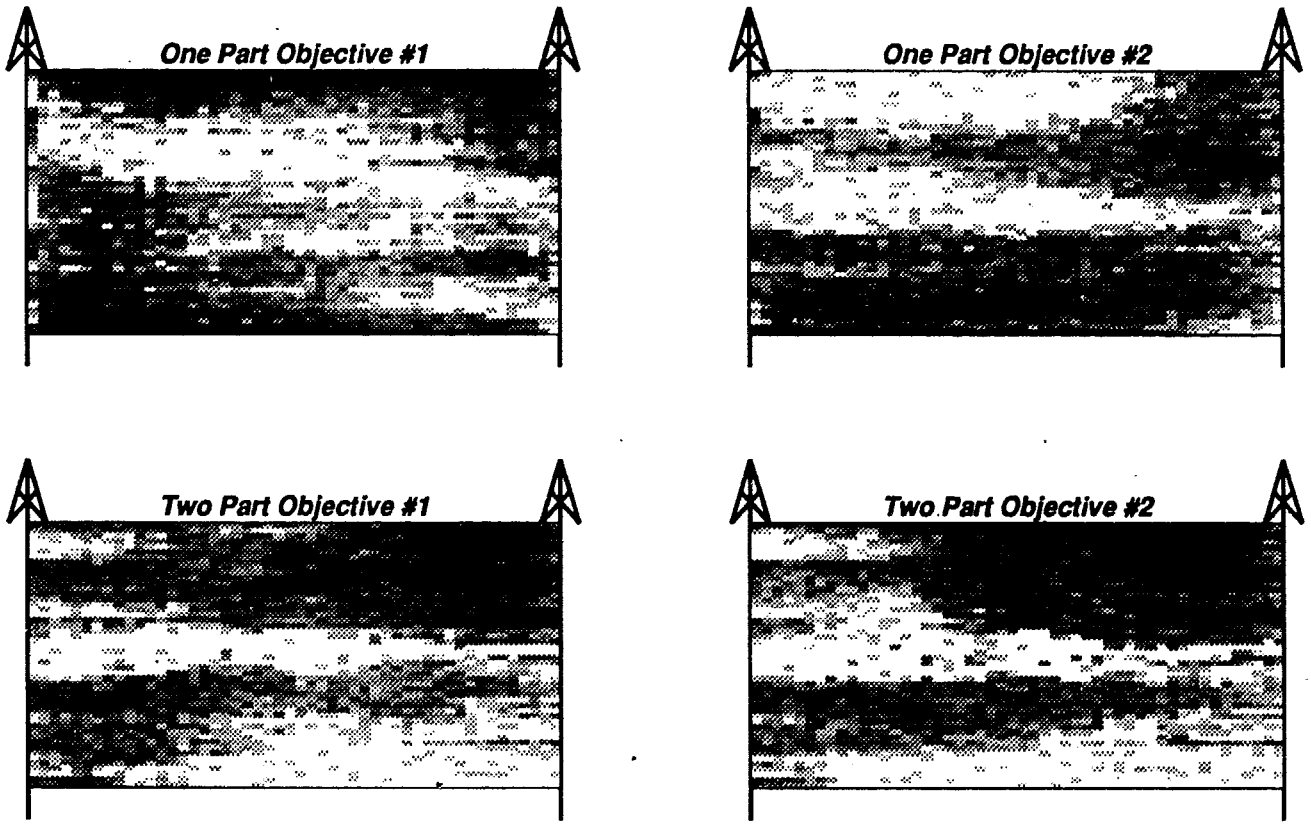

Figure 2: An example of simulating interwell variability using simulated annealing. The top two realizations were generated with a one part objective function. Note the discontinuities near the wells. The bottom two realizations were generated with the objective function separated into two components: the pairs involving at least one conditioning datum and those involving only simulated points. Note the improved continuity near the wells.

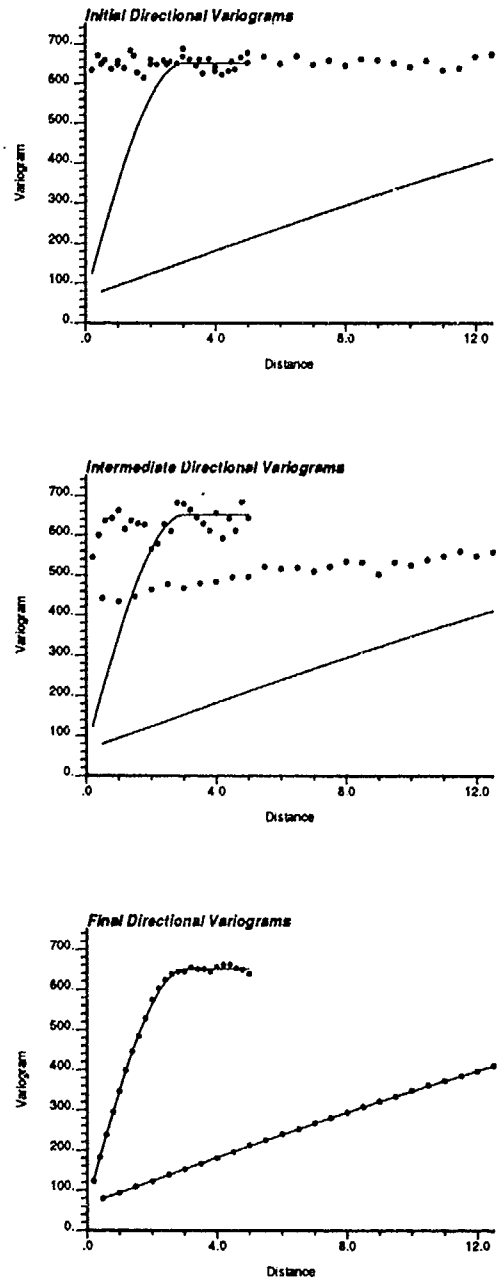

Figure 3: An example of the actual variograms at the beginning, half way through, and at the end of a simulated annealing conditional simulation. These results correspond to the generation of the lower left image on Figure 2.

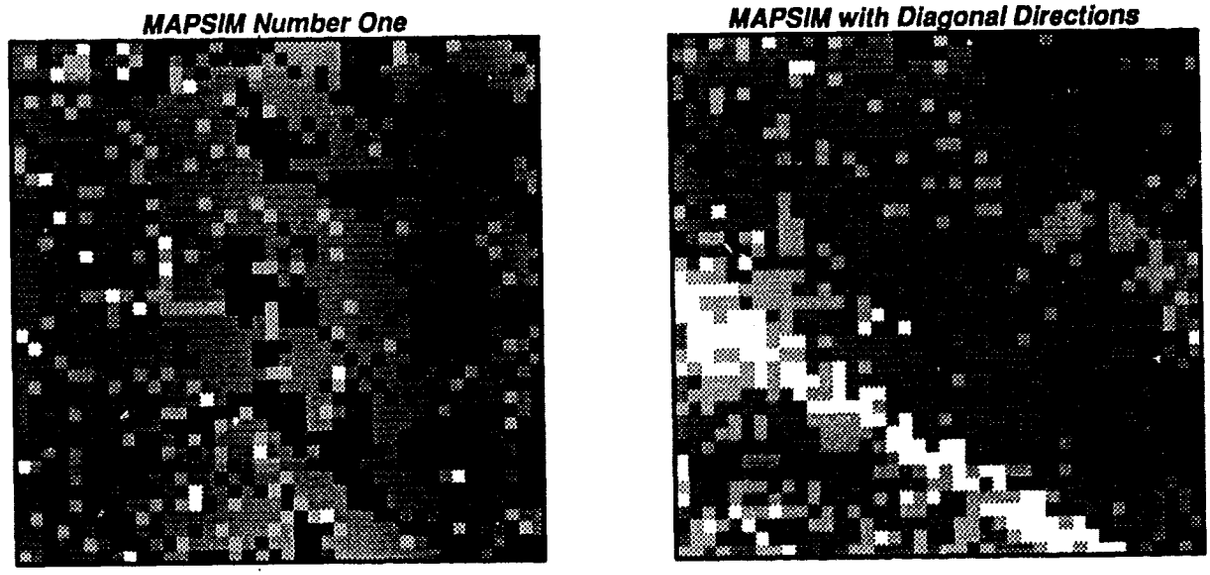

Figure 4: An example application of the Maximum A Posteriori (MAP) algorithm: the left image was created by retaining the two-point histogram, for 15 lags, in the two coordinate directions; the lower right image was created by choosing the two directions parallel and orthogonal to the banding.

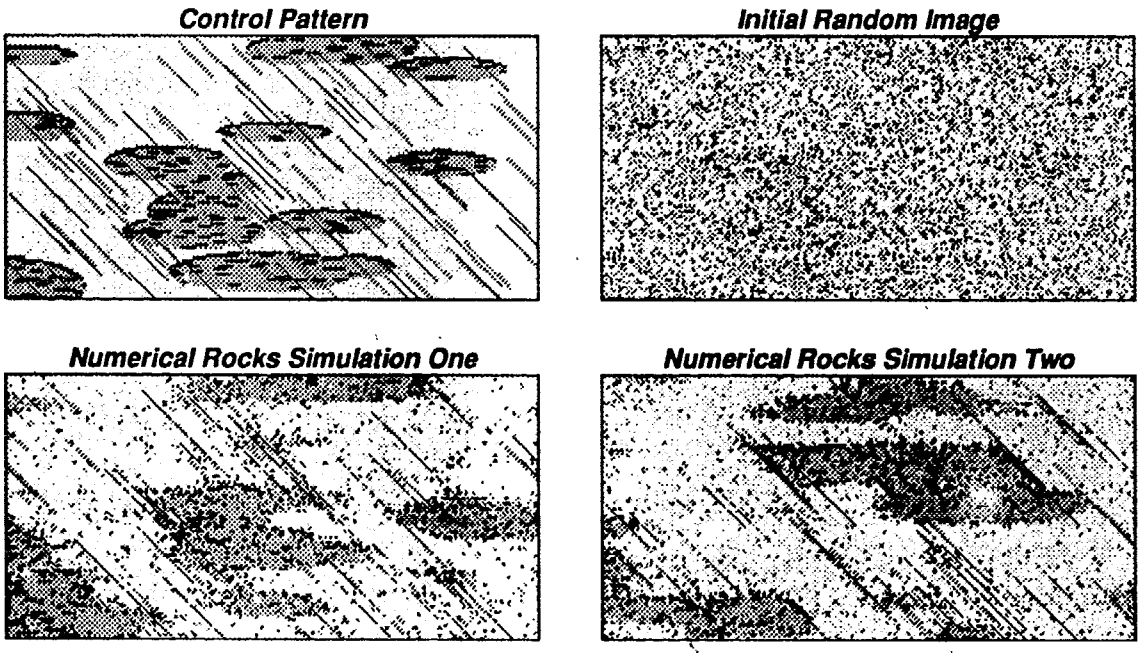

Figure 5: An example of the *numerical rocks* generated using Farmer's algorithm: the upper image is a difficult control pattern from which a two-point histogram is retained. A random image is then modified with annealing to create an unconditional simulation. Two alternate images are shown.

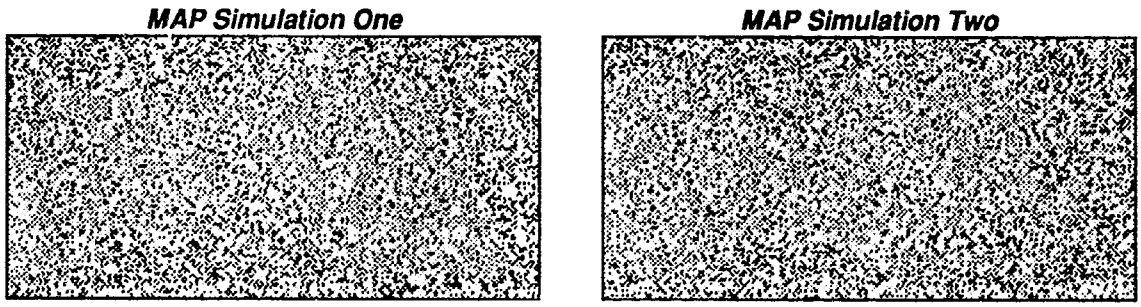

Figure 6: An example application of the Maximum A Posteriori (MAP) algorithm: the two-point histogram was taken from the control pattern shown on Figure 5. Two alternate images are shown. Note that these images have not converged to images that have the desired statistical properties of the control pattern.

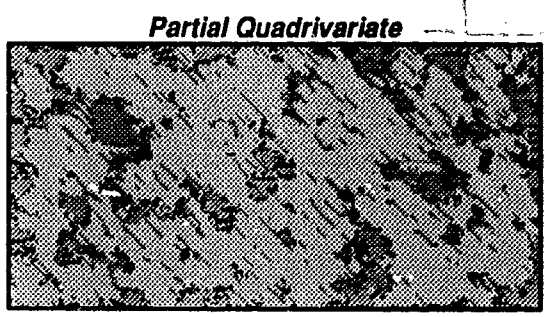

Figure 7: An example application of simulated annealing considering only partial quadrivariate information. This image reproduces the general features of the control pattern shown on Figure 5.

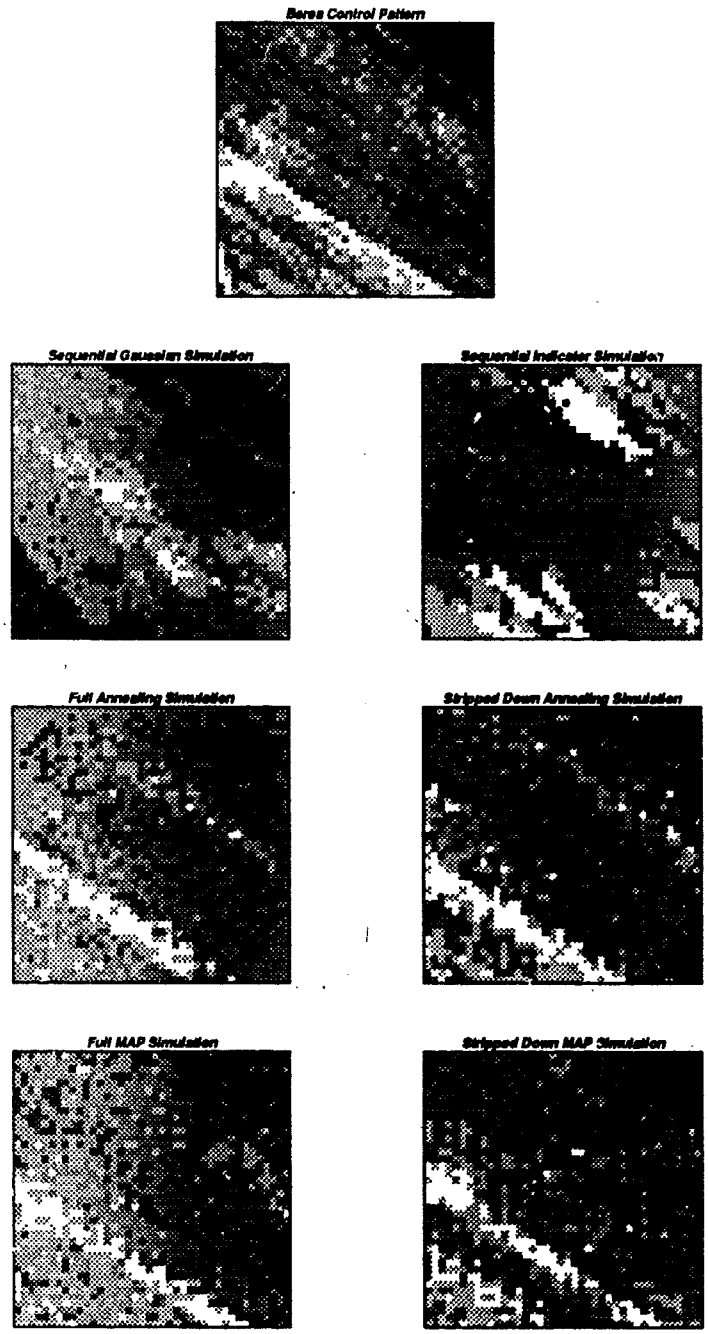

Figure 8: The Berea control pattern, one Gaussian simulation, one indicator simulation, a full simulated annealing simulation, a stripped down simulated annealing simulation, a full MAP simulation, and a stripped down MAP simulation. No initial conditioning data was used.