



Performance Characteristics of Simulated Annealing to Build Markov Fields

Kevin P. Parks & Laurence R. Bentley

Department of Geology and Geophysics, University of Calgary, 2500 University Drive N.W., Calgary, Alberta, Canada T2P 1E7, email bentley@geo.ucalgary.ca

Abstract

Simulated annealing is a global optimization method which permits stochastic generation of structured fields useful for populating simulator grids. Though computationally expensive, annealing methods have appeal because they allow incorporation of information from disparate sources to reduce the space of uncertainty explored. In this paper, Markov transitional structures are used to inform structure in two-dimensional categorical fields. Markov transitional structures can be derived from geological conceptual models, outcrop study, or core descriptions. They can be encoded into an objective function using the multipoint histogram approach. Some practical performance issues of using simulated annealing to construct Markov categorical fields are considered here. First, point swapping within a discrete field already populated with the necessary proportions of each category is found to produce results superior to drawing new nodal values from the underlying population distribution. Second, the convergence behaviour of annealing is shown to depend on the size of grid relative to the length-scales of categorical bodies implicit in the Markov transition structure. Lastly, in our experience, a combination of true simulated annealing (Metropolis algorithm) and iterative improvement (steepest descent) provides the most efficient approach to constructing complex Markov categorical fields by this method.

1 Introduction

Simulated annealing is a combinatorial optimization method which permits stochastic generation of fields comprised of categorical or continuous variables. Such fields are useful for imparting geologic

realism to aquifer simulator grids, e.g. Jensen et al.¹. Stochastic fields can be produced that honour the length-scale and geometrical information encapsulated in variograms, transition probability matrices, or multipoint histograms. Data relationships associated with geophysical surveys, well-tests, or core data can also be enforced. The attraction of simulated annealing is that it works without limitations associated with using *a priori* defined probability models. This flexibility is obtained at a relatively high computational cost. Also, annealing parameters need to be empirically calibrated to achieve optimum modeling efficiency.

Recently, Parks and Bentley² demonstrated that geological structures captured by Markov probability-transition matrices can be used as annealing constraints in stochastic field generation. These structures have significant influence on flow and transport. Since Markov transition matrices can be derived from geological conceptual models and limited core or outcrop study, they are attractive conveyors of geological information into simulator grids in situations where field data are sparse and likely to remain so. Annealing experiments with two-dimensional Markov fields have revealed several performance issues that need to be addressed in order to satisfactorily employ annealing in stochastic field generation. Some of these issues are discussed here.

2 Constructing Markov Fields by Multi-Point Histograms with Simulated Annealing

A sequence of events wherein the present state of the sequence is contingent on the state of the sequence at some time prior to the present is said to be a Markov chain. The structure of a Markov chain can be summarized in a “transition frequency matrix”, wherein the frequencies of transition from any one state to itself or the other states are tabulated (in columns) by state (the rows). If the frequencies are normalized by the row totals, the matrix provides the probability of transition from any state to any other in a unit step, e.g., Doveton³. A Markov transition matrix can be encoded in an annealing objective function as a series of multipoint histograms as first described by Farmer⁴. The details of encoding Markov transition matrices into an annealing objective function are given in Parks and Bentley².

Simulated annealing, as applied to stochastic simulation, is a global optimization method whereby an image or field that honours an idealized set of control statistics is created from a random field or structured field

generated from some other method. An idealized field is first characterized by a set of statistical or other descriptive measures. A trial field is then generated. The trial field can be a totally random image matching the ideal global histogram or a structured field created by a different algorithm. The same set of measures are calculated for the trial field. An objective function, O , can be computed as the weighted difference or squared difference between the ideal and trial field measures. Weighting is needed to assign equal importance to small as well as large values or components with different units of measurement (Deutsch and Journel⁵). The trial field is then perturbed and the objective function is recalculated, usually by a fast-updating procedure to minimize the number of computations. Perturbations that reduce the value of the objective function are kept. Perturbations that increase the value of the objective function are accepted with a probability that decreases in proportion to the increase in objective function scaled by a parameter called the “temperature”. Provided that the objective function is normalized by its original value, the cumulative probability density function, P_{accept} , governing acceptance of perturbations is of the form:

$$P_{\text{accept}} = \exp\left(\frac{O_{\text{old}} - O_{\text{new}}}{\text{Temperature}}\right) \quad (1)$$

If a predetermined number of perturbations are accepted (usually of the order of $10 \cdot N$ where N is the number of elements in a field) before some maximum number of total perturbations (of the order $100 \cdot N$), the temperature is reduced by some factor less than 1 (for example, 0.1). This procedure is repeated until the objective function falls below a threshold or its value can no longer be reduced. This implementation, which allows hill-climbing to avoid local minima along the objective function trajectory, is known as the Metropolis algorithm. Steepest-descent or iterative improvement variants, which do not allow hill-climbing, are also used, particularly for post-processing fields generated by other applications or Metropolis algorithm. For a starting point on the details of simulated annealing in stochastic field generation, the reader is referred to Jensen et al.¹.

In our implementation, a flexible, tripartite objective function is enforced that allows different weighting on the Markov transition matrices’ diagonal elements and the off-diagonal elements:

$$\begin{aligned}
 O = \frac{1}{O^o} \{ & \sum_{idir=1}^3 \sum_{i=1}^{im} [w_1 \frac{[d_{idir,i}^{train} - d_{idir,i}^{real}]^2}{d_{idir,i}^{train}}] \\
 & + \sum_{idir=1}^3 \sum_{i=1}^{im} \sum_{j=1, j \neq i}^{im} [w_2 \frac{[od_{idir,i,j}^{train} - od_{idir,i,j}^{real}]^2}{od_{idir,i,j}^{train}}] \quad (2) \\
 & + \sum_{i=1}^{im} w_3 [p_i^{train} - p_i^{real}] \}
 \end{aligned}$$

where O is the value of the objective function, $idir$ is the number of orthogonal directions to enforce the Markov structure, im is the number of states, d^{train} is the number of transitions on the diagonal of the model or training Markov transition frequency matrix, d^{real} is the number of transitions calculated for the image being annealed, od stands for the values of the off-diagonal elements of the Markov transition frequency matrix, and p is the proportion of each state. The calculated value of the objective function is normalized by the initial value of O , O^o . The fast updating scheme is that of Deutsch and Journel⁵ wherein only the contributions of the perturbed nodes to the multi-point histograms are updated. The motivation for the tripartite partition of objective function is discussed below.

3 Approaches to Field Perturbation

Two simple approaches to field perturbation were applied in annealing Markov fields. The first was to perturb the field through redrawing the value of a randomly selected node from the underlying categorical proportions. The second approach was to seed the initial grid with the correct proportions of each category in the Markov structure and then swap the values of two randomly selected nodes. The first approach is easier to implement in the updating scheme.

Our experiments found that the proportions drifted away from the desired proportions dictated by the Markov transition structure. Use of a penalty term, the third term in eqn. (2), improved the reproduction of desired categorical proportions in the annealing results but was never entirely satisfactory. When pair swapping within a fixed population of nodes is employed instead, there is by definition no drift in the field's categorical proportions.

4 Sensitivity to Length Scales on Finite Grids

The performance of the annealing is generally poor when strongly diagonally-dominant Markov transition structures are enforced. Diagonal dominance in Markov transition structures can result from small sampling intervals being used to characterize real sequences. The diagonal elements in a Markov transition matrix encapsulate length-scale information. On finite grids, the diagonal terms will determine the ratio of length scales in each category to grid scale. As the length-scale increases relative to the grid scale, the number of possible geometries honouring the ideal Markov structure decreases. Consequently, it may become more difficult for annealing to find the global minimum by standard perturbation approaches of random pair swapping or node replacement. To demonstrate the length-scale effect, four three-state Markov structures were annealed. The matrices had the same relative transition probabilities between different states. The categorical length scales (as defined by the median body influence length discussed in Rosen and Gustofsen⁶) were the same for each state within the matrices but different between the four cases.

Each structure was enforced on a 100x100 node grid using iterative improvement to a maximum of 2 million perturbations. As the ratio of grid dimension to median body-influence length decreases, the performance of the annealing degrades as measured by the slope of the objective-function trajectory (Figure 1). The remedy for the length-scale effect is to use a larger grid and subsection that grid accordingly for model parameterization.

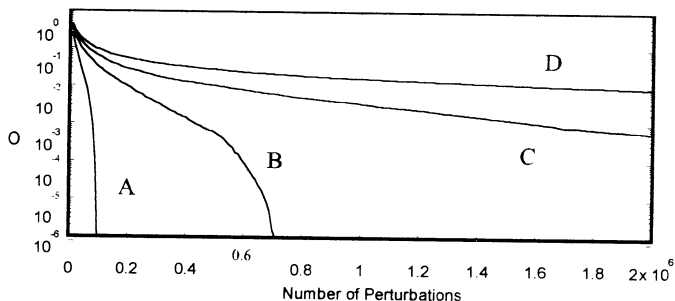


Figure 1: Effect of length scale versus grid size on annealing performance. The ratios of grid size to body influence lengths are A-41.5, B-23.4, C-15.2, and D-7.4 units

5 Improving Annealing Performance with Iterative Improvement

Neither iterative improvement nor conventional annealing were found to be particularly successful in generating complex Markov fields, especially those with five or more states and complicated, non-random interrelationships between categories. The best annealing results for complex Markov fields were attained when true simulated annealing with random swapping of values on a grid seeded with prescribed proportions of each state was followed by iterative improvement using the output of the conventional annealing.

For example, consider the annealing performance of a complex five-state Markov structure (Figure 2). In this example, the structure was derived from actual sediment cores taken from a hazardous waste site. The transition matrix captured bedding structure as well as a noisy, nearly random spatial structure between lithologic categories. Figure 2 shows the objective-function trajectory during a single annealing experiment with this structure. Due to the effects of relative scale of bedding and grid size, the number of states, and the complexity of the structure, the annealing would not converge within a reasonable number of perturbations. However, by post-processing the realization with iterative improvement, a satisfactory value of the objective function was reached. This behaviour suggests the surface of the objective function is relatively smooth. Using iterative improvement alone is not as effective because the annealing becomes trapped within whatever local minimum on the objective function surface encompasses the starting point (e.g., Figure 2, trajectory C). Only if the local minimum at the starting point is fortuitously deep enough will a satisfactory simulation be achieved.

6 Conclusions

Markov fields offer the groundwater modeler the opportunity to bring the power of stochastic simulation to situations where field data are sparse and likely to remain so. Simulated annealing can be used to generate Markov fields by enforcing multipoint histograms. Several performance issues are discussed here. Pair swapping is found to be a more effective approach to perturbation than randomly drawing single nodes as far as reproducing desired categorical proportions. Convergence difficulties associated with large categorical body length-scales relative to grid dimensions can be resolved by using larger grids. Finally, a combination

of true simulated annealing by Metropolis algorithm post-processed with iterative improvement is found to be the most effective way to anneal complex Markov fields.

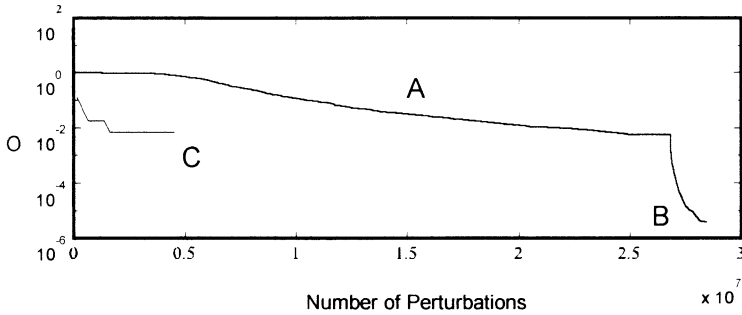


Figure 2: Improving annealing performance by post-processing with iterative improvement. A is the trajectory of O during true simulated annealing. B is the trajectory of O during post-processing of A by iterative improvement. C is the trajectory of O during annealing by iterative improvement alone, becoming trapped in a local minimum.

7 Acknowledgments

This work has been supported through the Amoco Canada Petroleum Company Ltd. Graduate Fellowship in Geology. Additional support has been provided by the Canada Centre for Inland Waters, the NSERC Grant No. OGP0122023, and the Department of Geology and Geophysics, University of Calgary.

8 References

- [1] Jensen, J.L., L.W. Lake, P.W.M. Corbett, and D.J. Goggin, *Statistics for Petroleum Engineers and Geoscientists*. Prentice Hall, Upper Saddle River, New Jersey, p. 318-326, 1997.
- [2] Parks, K.P., and Bentley, L.R., Capturing conceptual models of stratal architecture with Markov chains and simulated annealing, *Proc. of the 11 Int. Conf. on Computational Methods in Water Resources*, eds., A.A. Aldama, J. Aparicio, C.A. Brebbia, W.C. Gray, I. Herrera, and G.F. Pinder, Computational Mechanics



Publications, Southampton and Boston, pp. 655-662.

- [3] Doveton, J.H., 1994. Theory and Application of Vertical Variability Measures from Markov Chain Analysis, Chapter 6 *Stochastic Modeling and Geostatistics - Principles, Methods and Case Studies*, eds., J.M. Yarus and R.L. Chambers, American Association of Petroleum Geologists, Tulsa, pp. 55-64, 1994.
- [4] Farmer, C.L., Numerical Rocks, *Proceedings of the First European Conference on The Mathematics of Oil Recovery, 1989*, ed. P.R. King, Oxford University Press, New York and Oxford, pp. 437-448, 1992.
- [5] Deutsch C.V. and Journel, A.G., *GSLIB Geostatistical Software Library and User's Guide*. Oxford University Press, New York and Oxford, pp. 154-160, 1992.
- [6] Rosen, L., and Gustafson, G., A Bayesian-Markov geostatistical model for estimation of hydrogeological properties. *Ground Water*, vol. 34, no. 5, pp. 865-875, 1996.