USING THE CRITICAL TEMPERATURE TO IMPROVE THE SPEED OF GEOSTATISTICAL APPLICATIONS OF SIMULATED ANNEALING

K. Norrena
University of Alberta, Edmonton, Alberta, CANADA (knorrena@ualberta.ca)
C. V. Deutsch
University of Alberta, Edmonton, Alberta, CANADA (cdeutsch@civil.ualberta.ca)

Reliable numerical geological models must honor all available information, including complex non-linear geologic features and flow-related measurements. Simulated annealing (SA) has seen increasing application in the integration of such complex data in geostatistical models. Nevertheless, the delicate adjustment of the annealing schedule and excessive CPU requirements continue to limit the application of SA in geostatistics. Most difficulties with SA are linked to the specification of how to reduce the temperature control parameter during the SA procedure. Convergence, or a reasonable numerical model, may not be obtained when the temperature is lowered too quickly. A slow reduction of temperature leads to excessive CPU time. From thermodynamics and observation, we know that there exists a critical temperature. Convergence is assured when the temperature is lowered quickly until the critical temperature, held at the critical temperature for some time, and then lowered quickly again. Knowledge of the "critical temperature" permits robust and fast application of annealing. We present a method for the fast determination of the critical temperature for geostatistical applications. A fast SA run is made where the temperature is lowered very quickly, the specific heat is calculated for all temperatures, that is, the derivative of energy with respect to temperature; the specific heat reaches a maximum at the critical temperature. We demonstrate the use of the critical temperature for faster geostatistical modeling.

INTRODUCTION

Annealing in material sciences refers to a heat treatment process where a material is exposed to an elevated temperature for an extended period of time and then slowly cooled. The goal of annealing is to alter the microstructure and hence the global properties of the material by recrystallization. Slowly cooling the material causes the molecules to form a microstructure with greater order; they position themselves to minimize Gibbs free energy. Very slow cooling promotes greater order and low total "energy". The material is cooled until the molecules no longer reposition themselves and the material is frozen.

Simulated annealing (SA) is a numerical analogy to the process of true annealing. A matrix of nodes, each node having any number of attributes (grades, rock type, porosity, permeability), is the material which SA will be applied to. The algorithm perturbs the attribute value at each node in a random path over the matrix. The perturbations mimic thermal vibrations, or heat energy.

WJ Kleingeld and DG Krige (eds.), Geostats 2000 Cape Town, Volume 1, 254-262.

The objective function (O) measures the energy of the matrix; the mismatch between some user defined mathematical function and the state of the matrix. For example, the objective function could measure the mismatch between variogram of the conditioning data and the variogram of the matrix. Optimality is reached when the objective function is minimized. There are sufficient degrees of freedom in most SA problems to permit the objective function to reduce to zero. Moreover, in geostatistical applications of SA, the solution is non-unique. Different solutions are considered a space of uncertainty.

SA successively perturbs the matrix nodes to achieve a matrix with optimal properties. If a perturbation decreases the objective function ($O_{new} \leq O_{old}$) the perturbation is accepted. If a perturbation increases O ($O_{new} > O_{old}$) it is accepted or rejected based on the Boltzmann distribution, see Figure 1 and Equation (1).

$$P\{acceptance\} = \begin{cases} 1 & if\ O_{new} \leq O_{old} \\ e^{-(O_{new} - O_{old})/T} \end{cases} \quad (1)$$

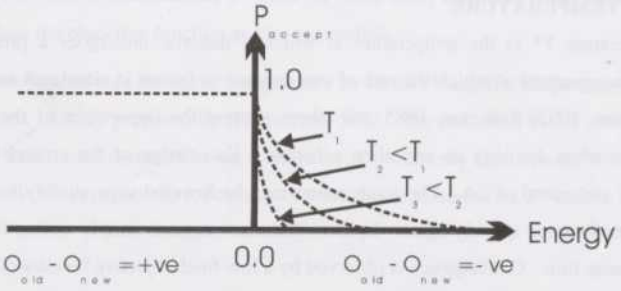

Figure 1: The probability of accepting a perturbation is 1.0 when the perturbation decreases the global energy and <1 when the perturbation increases the global energy. Note that the probability of acceptance for perturbations that increase the global energy decreases with temperature.

Application of SA in combinatorial optimization was developed by Kirkpatrick et al. 1983 and independently by Cerny, 1985. Other research followed in a variety of different disciplines but application of SA for use in geology was pioneered by C.L. Farmer, 1992 who used SA to generate rock type models. SA has shown significant promise but little practical application in geostatistics. The promise of SA has not been realized because of the difficulty in setting the temperature reduction schedule and by excessive CPU requirements. Theoretical knowledge of acceptance/rejection of molecular movement has been used in SA, but other theoretical thermodynamic relations have not been exploited. We attempt to make better use of this knowledge.

ANNEALING SCHEDULE

The temperature reduction schedule, also known as the annealing schedule, consists of three stages: (1) a heating stage, (2) a heat soaking stage, (3) and a cooling stage. The idea is to heat the matrix until very "hot," that is, virtually all perturbations are accepted, and cool in a manner that achieves a minimum (or low) objective function with acceptable CPU time. Aggressive reduction of temperature will often lead to a matrix frozen in a suboptimal state.

Certain annealing schedules can be shown to converge (see Geman and Geman, 1984 for proof), but they are prohibitively slow in practice. Often, an empirical scheme is adopted, that is, the temperature starts at a high initial temperature T_0 and is lowered by some reduction factor • whenever enough perturbations have been accepted K_{accept}, or too many have been tried K_{max} for any one temperature, and stops when the convergence criteria (O_{min}) has been met or when the algorithm has attempted too many perturbations with no change in the objective function ($K_{max} \cdot S$).

THE CRITICAL TEMPERATURE

The critical temperature T^* is the temperature at which a material undergoes a phase change in true annealing and the temperature at which the rate of convergence is fastest in simulated annealing. Basu and Fraser, 1990, Farmer, 1992, Rothman, 1985 and others express the importance of the knowledge of the critical temperature when devising an annealing schedule. Knowledge of the critical temperature would improve efficiency and speed of SA. The temperature may be lowered very quickly to T^*, held at T^* for some time, and then lowered quickly again. In fact, Rothman suggests simply setting the temperature just below T^* for the entire time. Convergence is observed by a low final objective function.

Determination of the critical temperature requires calculation of the thermodynamic property known as specific heat, defined as the derivative of energy with respect to temperature:

$$C(T) = \frac{\partial E}{\partial T} \qquad \text{where } E=E_{old}-E_{new} \tag{2}$$

The critical temperature T^* is at a maximum in specific heat. Analytical determination of T^* would require more CPU effort than we are trying to save; however, we show that it may be determined quickly by numerical means.

PROPOSED METHOD

The SA problem is first run using a fast annealing schedule. There are $i=1,\ldots,n$ steps in this fast run. The specific heat is calculated at each step i using:

$$\frac{\partial E_i}{\partial T_i} = \frac{|E_{i-1} - E_i|}{|T_{i-1} - T_i|} \tag{3}$$

where • $E/• T$ is specific heat, E_{i-1} is the average energy over temperature T_{i-1}, E_i is the average energy over temperature T_i.

The specific heat is plotted against the temperature. The maximum specific heat is at the critical temperature T^*. The maximum could be detected automatically. We show this plot for many different cases later.

The critical temperature T^* depends on many factors including the number of nodes in the model and the details of the objective function. The dependency of T^* on some factors will be demonstrated later in this paper; however, we propose that T^* be determined separately for each problem. The two-step procedure is to (1) determine T^* by a fast run and visual inspection of a plot, and (2) fix annealing schedule to account for T^*. This will have to be repeated for each problem.

A number of schemes could be considered to use T^* in the annealing schedule. Rothman, 1985 suggests that the temperature be held just below the T^* for the entire duration of the annealing run. We considered this option and others including a three part annealing schedule where an initial high temperature is lowered by a fast rate λ_1 until T^*, the rate is reduced near T^* until for some time, and finally the temperature is lowered quickly again to reduce the objective function as low as possible.

In conclusion, we recommend that the initial temperature be set to T^* and the temperature lowered with the same empirical scheme initially proposed by Kirkpatrick, 1983. The reduction factor should be set to about 0.5, which is higher than would normally be used; however, T^* is already significantly lower than the default starting temperature. We investigate the robustness of this scheme to imprecise T^* values, see below.

IMPLEMENTATION

The proposed methodology for determining T^* was implemented using `sasim`, a simulated annealing program available in `GSLIB`. The `cluster.dat` data set from GSLIB was used. A location map, target histogram, and target variogram are shown in Figure 2. The modeled variogram has a 20% nugget effect and spherical type structure with an isotropic range of 15 nodes.

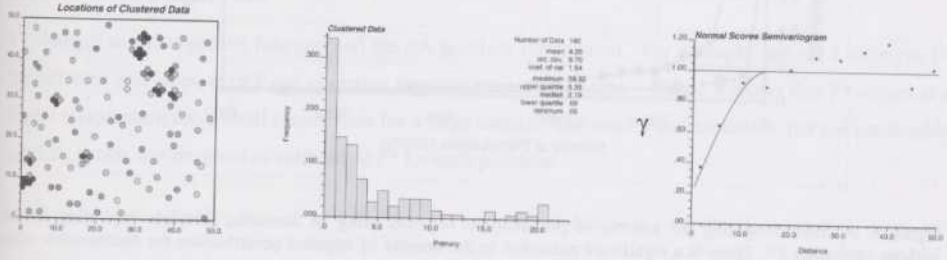

Figure 2: A location map, histogram, and variogram of the data set used in the paper.

A fast run using `sasim` and the default-annealing schedule took approximately 11 seconds on a Pentium II 350. The output from `sasim` was post-processed and plotted in Figure 3. The initial temperature T_0 is 1.0 and the final temperature is approximately $1x10^{-6}$. The critical temperature occurs when the specific heat peaks, that is, at a temperature of approximately 0.0007.

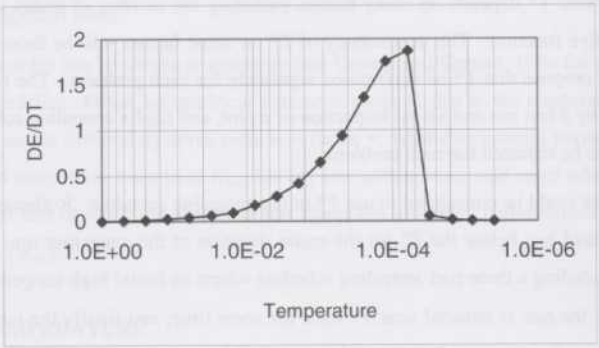

Figure 3: Post-processed `sasim` results from one realization. The peak indicates T^*. T_0, the initial temperature occurs at 0.1, and T_{final} the temperature where the algorithm stopped occurs at $1x10^{-6}$.

The annealing schedule in `sasim` was modified to use T^* as the initial temperature (as described above). Figure 4 shows that there is a significant reduction in the number of perturbations before convergence when one exploits T^*. The CPU time is reduced from about 4 minutes to about 1 ½ minutes on a PC. We see a 3:1 improvement for relatively small grid networks (1000s of nodes). This ratio improves to 5:1 as the grid size increases to more than a million grid nodes.

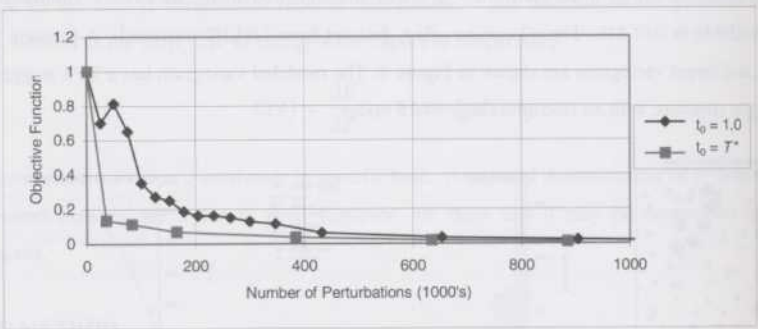

Figure 4: A chart comparing the number of perturbations required using an annealing schedule exploiting T^* and without exploiting T^*. There is a significant reduction in the number of required perturbations for optimization when $T^*$is used in the annealing schedule.

Dependence on •

The critical temperature is determined by a "fast" annealing run. We should be concerned about T^* being sensitive to the cooling rate used in the fast run. Figure 5 shows T^* for reduction rates ranging from 0.1 to 0.9 in increments of 0.1. T^* consistently appears at T=0.0007 for each reduction rate, indicating that $T^*$is robust with respect to the cooling rate used to determine it.

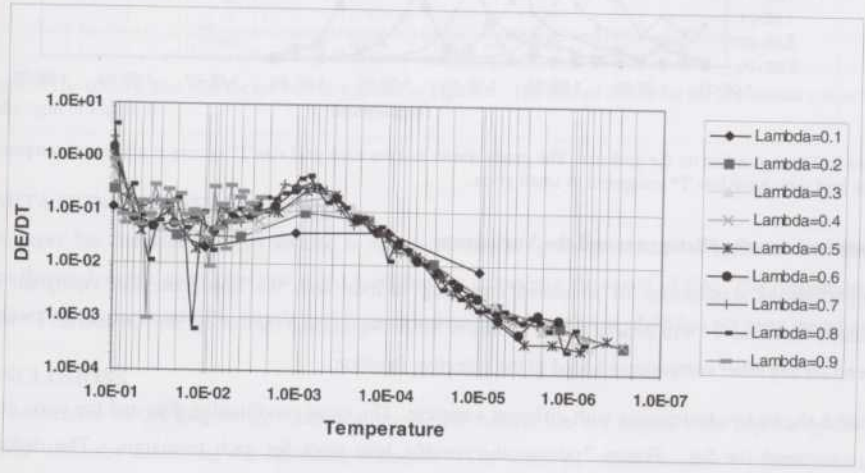

Figure 5: A specific heat versus temperature plot using 9 different reduction factors. The reduction factor has no effect on T^*.

Dependence on the Random Number Seed

The order of visiting the grid nodes and the drawing from the Boltzmann distribution is random. We should be concerned about sensitivity to the sequence of random numbers. A number of different random number seeds were tried and there was no change in T^*.

Dependence on the Grid Size

T^* changes as the objective function and the SA problem formulation. For example, we see a different T^* for different grid sizes, even if the objective function remains the same. Figure 5 shows that T^* occurs at a higher temperature for a small matrix than for a large matrix. The results are consistent, but not predictable a-priori; hence, our proposal of calibrating T^* for each problem.

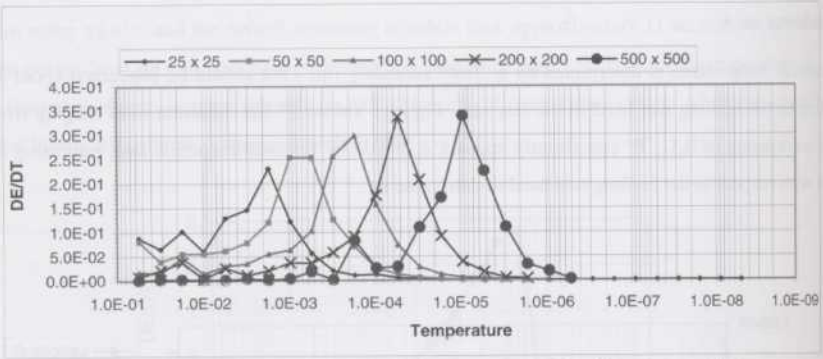

Figure 5: $T^*$is dependent on the grid size. The graph shows that for each grid size T* occurs at a different temperature. Large grid sizes have low T* compared to small grids.

Dependence on the Histogram and the Variogram

Straightforward applications of simulated annealing include both the histogram and variogram in the objective function. T^* will depend on the details of the histogram and the variogram. Of course, T^* will also depend on any other components added to the objective function.

Figure 6 shows two histograms with different variance. The same conditioning data and the same grid size are considered for SA. Figure 7 shows the specific heat plots for each histogram. The dashed line corresponds to the low variance histogram to the right. Note that T^* is greater for this low variance histogram.

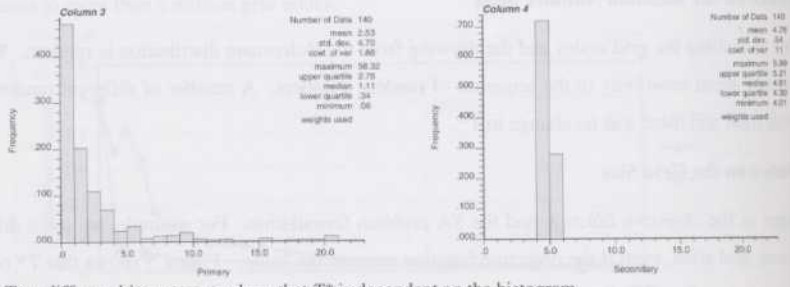

Figure 6: Two different histograms to show that T^* is dependent on the histogram.

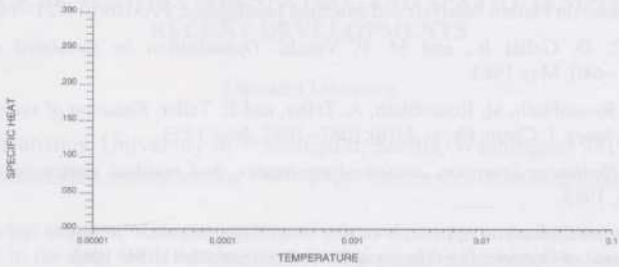

Figure 7: The specific heat plot for the two histograms in Figure 6. The dashed line is for the low variance distribution on the right of Figure 6.

LIMITATIONS AND FUTURE WORK

This paper has considered T^* as applied to simulated annealing for geostatistical applications. There are other thermodynamic principles that could contribute to enhancing the speed of SA. The robustness of the results (convergence and CPU speed) with respect to imprecise knowledge of T^* has not been investigated.

CONCLUSIONS

The application of SA to generate numerical geological models has not gained wide popularity due to the relative difficulty in setting the annealing schedule. The setup of other simulation algorithms such as Gaussian or indicator methods is comparatively straightforward. We present a fast method for determining T^*, which simplifies the annealing schedule and improves CPU speed. . We show that T^* depends on the grid size, the histogram, and the variogram. Although T^* depends on the problem formulation it is insensitive to the details of how we obtain it, that is, it is independent of the reduction factor and the random number seed. The proposed method reduces the complexity associated with devising an annealing schedule and reduces the required number of perturbations for convergence.

REFERENCES

A. Basu and L. N. Frazer. *Rapid determination of the critical temperature in simulated annealing inversion.* Science, 249:1409--1412, September 1990.

C.V. Deutsch, *Annealing Techniques Applied to Reservoir Modeling and the Integration of Geological and Engineering (Well Test) Data.* Ph.D. Thesis, Stanford University, Stanford, CA,1992.

C. V. Deutsch and P. W. Cockerham. *Practical considerations in the application of simulated annealing to stochastic simulation.* Mathematical Geology, 26(1):67-82, 1994.

C. V. Deutsch and A. G. Journel. *The application of simulated annealing to stochastic reservoir modeling.* In Report 4, Stanford Center for Reservoir Forecasting, Stanford, CA, May 1991.

C. V. Deutsch and A. G. Journel. *GSLIB: Geostatistical Software Library and User's Guide.* Oxford University Press, New York, 2nd edition, 1998.

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

S. Geman and D. Geman. *Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images.* IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI6(6) :721--741, November 1984.

S. Kirkpatrick, C. D. Gelatt Jr., and M. P. Vecchi. *Optimization by simulated annealing.* Science, 220(4598):671--680, May 1983.

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

D. H. Rothman. *Nonlinear inversion, statistical mechanics, and residual statics estimation. Geophysics,* 50:2784--2796, 1985.

V. Ÿ Cerny. *Thermodynamical approach to the travelling salesman problem: an efficient simulation algorithm.* Journal of Optimization Theory and Applications, 45:41--51, 1985.