MPS Simulation with a Gibbs Sampler Algorithm

Steve Lyster and Clayton V. Deutsch

Complex geologic structure cannot be captured and reproduced by variogram-based geostatistical methods. Multiple-point statistics have been introduced in recent years to improve the modeling of facies to account for prior geologic knowledge and produce results that look more like real geology. There are two difficulties in using multiple-point statistics: the first is determining what statistics to use; the second is how to reproduce these complex statistics in a simulation. This paper puts forwards a simulation method using multiple-point statistics in a Gibbs sampler algorithm for improved geostatistical modeling of facies.

Introduction

In cases where rock types or facies control the continuous variables of interest, modeling of rock types as indicators is often undertaken in a geomodeling workflow. The spatial structure of the facies is different than the continuous variables and is often very complex; curvilinearity, long-range connectivity and complex relations between facies are common. Traditional variogram-based geostatistical methods do not adequately reproduce the high-order structure present in real geology.

Multiple-point statistics (MPS) use spatial moments of order greater than two to model the structure of geology. These high-order statistics cannot be derived directly from available data except in exceptional cases; a representative model of geology called a training image (TI) is used to infer MPS for simulation. What statistics to take from the TI to capture the "essence" of the geology and how to reproduce these statistics in simulation are two related questions that must be asked to use MPS methods.

A number of algorithms using MPS have been developed: SNESIM (Guardiano and Srivastava, 1993, Strebelle, 2002); FILTERSIM (Zhang et al, 2006); GROWTHSIM (Wang, 1996, Eskandaridalvand, 2008); SIMPAT (Arpat and Caers, 2007); neural networks (Caers, 2001); simulated annealing (Deutsch, 1992); and a Gibbs sampler (Srivastava, 1992, Lyster et al, 2006, Lyster, 2007, Lyster and Deutsch, 2008). The Gibbs sampler will be discussed further here.

MPS-GS Algorithm

A Gibbs sampler is an iterative algorithm that starts with an existing "state" and perturbs all of the variables in the state space using conditional distributions until convergence is reached (Geman and Geman, 1984, Gelfand and Smith, 1990, Casella and George, 1992, Robert and Casella, 2004). For geostatistical purposes, the state space is a grid populated with values, the variables are the facies values at all locations, and the simulation proceeds by changing one location conditional to the current state of the surrounding variables.

The original concept of using a Gibbs sampler in geostatistics was put forth by Srivastava, 1992; in that implementation ordinary kriging was used to calculate the conditional probability of each facies. The use of MPS in the form of Bayes' Law was touched upon with a brief example and not explored further. An MPS simulation algorithm using a Gibbs sampler framework, called MPS-GS, has been developed at the CCG for the last few years. This algorithm uses the following workflow:

1. Start with a randomly-populated field on the coarsest grid
2. At a random location:
    a. Calculate the conditional probability density function (cPDF)
    b. Adjust the cPDF to account for secondary information and global facies proportions
    c. Apply noise reduction or cleaning and correct cPDF to sum to 1.0
    d. Draw a new facies value from the corrected cPDF

e. Move to another unsampled location and repeat Step 2

3. After every location has been visited, check for convergence
    a. If there is no convergence yet, repeat Step 2 over all locations
    b. If convergence has been achieved, populate the next-coarsest grid
4. After the final grid has been simulated write out the results and start the next realization

The workflow outlined above is a general Gibbs sampler framework for geostatistics. Any method could be used to calculate the conditional probabilities in Step 2a; the results will reflect this modeling decision. The MPS-GS algorithm uses the concept of multiple-point events (MPEs) to determine the conditional distributions in Step 2a. A MPE is a discrete set of points in space, denoted E_i, i=1,…,M. At the points making up a MPE there are a=1,..,K^N possible combinations of facies, or classes, where K is the number of facies and N is the number of points in each MPE. A single class α within a MPE is denoted as E_i^a. The MPEs contain lower-order information than a full MPS template but are easier to infer from a TI. An example of this is shown in Figure 1. The central MPS template contains 24 points (not counting the central location) and the six MPEs cover the same 24 points when taken together. The major difference for an iterative algorithm is that a single large template will have many classes that are not found in the TI; using a random initial image for simulation a large proportion of the simulated locations will have patterns that cannot be used for inference of conditional probabilities. MPEs reduce this problem.

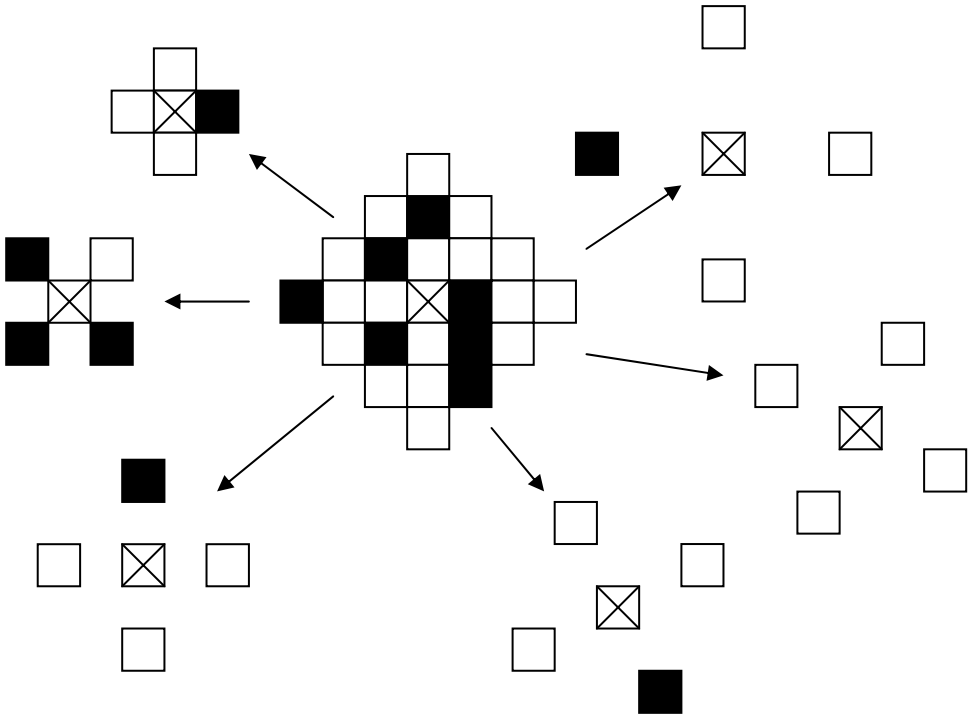

Figure 1: An example of a 24-point MPS template broken into discrete four-point MPEs.

MPEs can be expressed as indicators similar to single-point facies data. Using these indicators in a kriging-like estimator, the conditional probability P^* of facies k is found by the equation

$$P^*(k) = \sum_{i,\alpha} \lambda_{i,\alpha}^k \left[I\left(E_i^\alpha\right) - P\left(E_i^\alpha\right) \right] + P(k) \quad (1)$$

where $P(k)$ is the global probability of k, $I(E_i^a)$ is the indicator of class α in MPE i, $P(E_i^a)$ is the global probability of that class, and $\lambda^k_{i,a}$ is the weight assigned to the indicator. The weights are linear estimation

weights, as Equation 1 is a linear estimate that uses non-linear data (MPEs). The optimal weights can be found by minimizing the estimation variance and solving the resulting system of equations:

$$\sum_{j,\beta} \lambda_{j,\beta}^{k} \cdot Cov\left\{E_i^{\alpha}, E_j^{\beta}\right\} = Cov\left\{E_i^{\alpha}, k\right\} \tag{2}$$

Equation 2 uses covariances between MPEs; these covariances can be found by scanning a TI and using the definition of an indicator covariance:

$$\begin{aligned} Cov\left\{E_i^{\alpha}, E_j^{\beta}\right\} &= E\left\{I\left(E_i^{\alpha}\right) \cdot I\left(E_j^{\beta}\right)\right\} - E\left\{I\left(E_i^{\alpha}\right)\right\} \cdot E\left\{I\left(E_j^{\beta}\right)\right\} \\ &= P\left(E_i^{\alpha} \cap E_j^{\beta}\right) - P\left(E_i^{\alpha}\right) \cdot P\left(E_j^{\beta}\right) \end{aligned} \tag{3}$$

It is notable that if single-point events are used, Equation 1 becomes the full indicator cokriging estimate, Equation 2 becomes the indicator cokriging system, and Equation 3 defines indicator covariances at a defined offset. The MPEs contain more information, however. For the 24-point template shown in Figure 1, a full indicator cokriging system of two facies would be a MNK=48-order system, while the MPE estimator has an order of MK^N=96. The order of the system is related to the information content. If the number of facies is increased to three, the orders of the systems are then 72 and 486; for two facies and three eight-point MPEs the kriging system is still a 48x48 system of equations but the MPE system is 768x768.

Implementation Considerations

There are a number of factors to consider when using the MPS-GS algorithm in a practical simulation. These include the singularity of the system, a lack of inference from the TI for large MPEs, the population of multiple grids, honouring global and local PDFs, noise reduction, and stopping criteria.

The system in Equation 2 is singular because of the properties of indicators. MPE classes are both exhaustive and exclusive, meaning every location has exactly one class for each MPE with an indicator value of one, and the rest are zero. This introduces redundancy; if one class has an indicator of one then the rest must be zero, or if K^N-1 classes have indicators of zero then the remaining class must be one. This redundancy, and the resulting singularity, means the system has an infinite number of solutions with the same value, ie, estimation variance. It is suggested here to use the Moore-Penrose pseudoinverse to find the optimal solution that minimizes the norm of the weights, to avoid extreme weights (Anderson et al, 1999). Any method could be used to find an optimal solution and the results would be numerically correct, although care should be used when dealing with this type of a system.

When starting with a random image as the initial state, there are many mismatches from the TI; for larger MPEs many will be uninformed and thus provide no information. Larger MPEs may be necessary in some cases to capture the complex structure seen in the TI, but this early mismatch can cause convergence problems. A possible solution to this is to use single-point events as additional data for estimating the conditional probabilities:

$$P^*\left(k\right) = \sum_{i,\alpha} \lambda_{i,\alpha}^{k} \cdot \left[I\left(E_i^{\alpha}\right) - P\left(E_i^{\alpha}\right)\right] + \sum_{u,k'} \chi_{u,k'}^{k} \cdot \left[I\left(k';u\right) - P\left(k'\right)\right] + P\left(k\right) \tag{4}$$

In the early iterations many MPEs will be uninformed (assuming large MPEs) and the single-point data will provide additional information to help "guide" the Gibbs sampler towards a final realization. The optimal weights $\chi_{u,k'}^{k}$ can be found by expanding the system in Equation 2 to include covariances between MPEs and single-point indicators.

Honouring global and local PDFs is an important simulation consideration for any algorithm. The MPS-GS algorithm can accomplish this by using a multiple-servosystem approach (Lyster, 2008) in Step 2b. The estimated conditional probability of each facies from Equation 1 (or 4) is updated using the equation:

$$P'\left(k\right) = P^*\left(k\right) + \mu \cdot \left[P^{LOC}\left(k\right) - P^{SIM}\left(k, B_k\right)\right] \tag{5}$$

where μ is a control parameter, P^{LOC} is the local PDF distribution, P^{SIM} is the current simulated PDF, and B_k is the local PDF bin, calculated using the following equation:

$$B_k = \text{int}\left(\frac{P^{LOC}(k) - P^{MIN}(k)}{P^{MAX}(k) - P^{MIN}(k) + \varepsilon} \cdot N_B \right) + 1 \tag{6}$$

The multiple-servosystem method "bends" the probability of each facies towards the correct local proportions, or towards the correct global PDF if there is no local information. This method is used in other MPS algorithms, albeit only for global proportions (Liu, 2006).

The results of MPS-GS can often be very noisy in some regions; different parts of the simulated field will converge at different times. Locations close to hard data can be expected to converge quickly and not change as much as locations completely isolated from any data. Because of this, noise reduction is implemented as mentioned in Step 2c. The noise reduction is accomplished by using the short-range connectivity of facies. This connectivity, denoted C_k, is defined as the number of immediately adjacent cells that have the value of facies k. An example of this is shown in Figure 2.

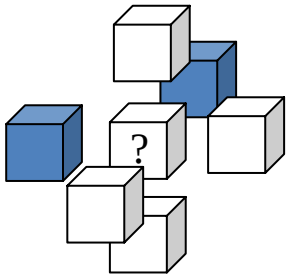

Figure 2: An example of the connectivity C_k at a location. C_{blue}=2, C_{white}=4.

Using this concept, the conditional probabilities from Equation 5 are modified by the equation:

$$P''(k) = P'(k) + \eta \cdot \left[P(C_k) - E\left\{ P(C_k) \mid P(C_k) > 0 \right\} \right] \tag{7}$$

where η is a control parameter and $P(C_k)$ is the probability of facies k having connectivity C_k as determined from the TI. The connectivity values range from zero to four (2D) or six (3D). Facies that have the correct connectivity have increased probabilities in this scheme, while those with inappropriate connectivity values at the current location have decreased probabilities. This reduces the chances of random noise appearing in a realization, or for random facies to appear in the middle of a well-ordered geobody that has already been formed.

Another method used to reduce noise is to freeze individual cells that have converged. Those cells that have modified conditional probabilities equal to 1.0 (after normalization) are frozen and removed from the random path. This speeds up convergence and reduces simulation time significantly by removing many locations from the path. This freezing is only done after several iterations have been completed, to allow structure to begin to form in the field.

As the MPS-GS algorithm proceeds, it is necessary to check for convergence in Step 3. Convergence can be defined in a number of different ways; making the decision of convergence too early could result in poor results, while allowing a simulation to proceed unnecessarily long wastes time and can cause artifacts. Three convergence criteria are currently used in MPS-GS:

1. The correlation between the iteration number and the number of changes over the last five iterations becomes positive.
2. A minimum proportion of cells to be changed in each iteration is not reached.

3. A set maximum number of loops over all unsampled locations is reached.

The number of cells remaining in the path affects the main stopping criterion (#1). Figure 3 shows a graph of the number of unfrozen points remaining in the random path and the number of changes over 50 iterations of MPS-GS. The number of changes drops drastically and then stabilizes. These stopping criteria have been found to work adequately.

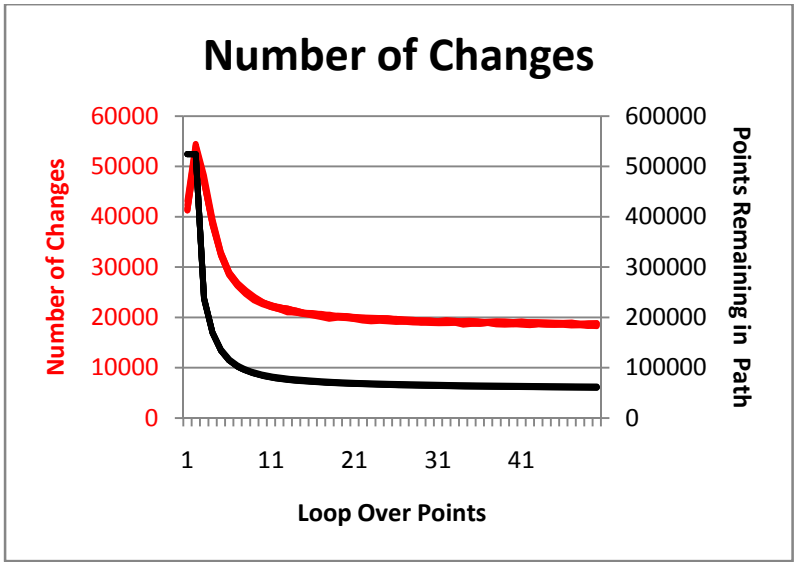

Figure 3: Number of unfrozen points remaining in the random path and the number of changes over 50 iterations of MPS-GS.

After convergence has been reached, the next-finer grid must be populated for the simulation to proceed. This could be accomplished by randomly populating the grid; however, the randomness injected into the image will disrupt much of the structure that was developed at the coarser grid and can lead to artifacts. In the MPS-GS algorithm finer grids are populated by discretizing the existing locations to four (2D) or eight (3D) new grid cells. An example of this is shown in Figure 4. The coarse grid cells are discretized to four cells each and only one out of every four (eight in 3D) is frozen as conditioning data for finer grids.

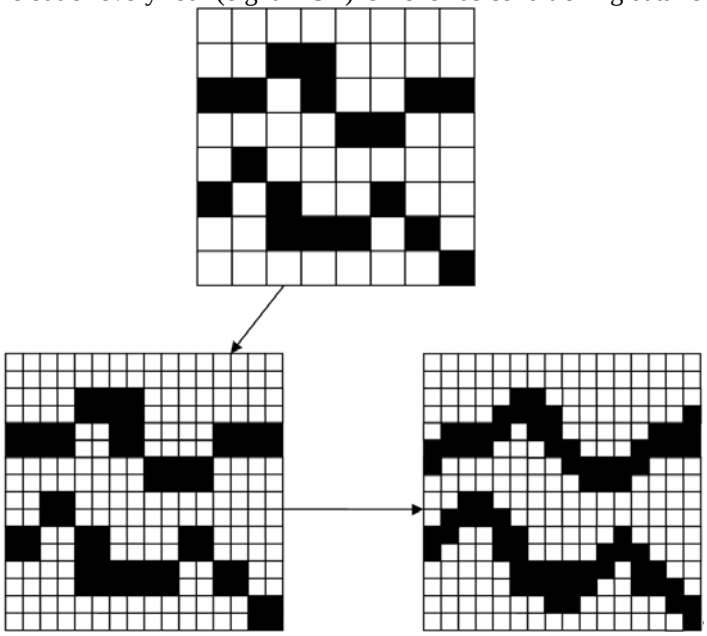

Figure 4: An example showing the method for populating multiple grids in MPS-GS.

Example – 2D Braided Channels

A small 2D TI featuring braided channels is shown in Figure 5. The TI is 250x250 cells and has about 30% channel facies with 70% background facies. The MPS-GS algorithm was used to create realizations with MPS derived from this TI; SISIM (Deutsch and Journel, 1998) was also used for comparison. Ten four-point MPEs were used in MPS-GS, with five grids. The optimal weights in Equation 4 were solved for and stored in a text file to save CPU time; solving the system took 36 seconds.

Twenty unconditional realizations were simulated using MPS-GS, two of which are shown in Figure 6. Simulation of these realizations took 27 seconds. The realizations show long-range connectivity and channel structure similar to the TI; there is not as much braiding as in the TI. Figure 7 shows east-west and north-south variograms of the 20 realizations compared to the reference TI variogram. The variograms are matched well.

A multiple-point histogram (MPH) was also used for measuring the "goodness" of the realizations. A 3x3 template, for a total of 512 classes, was used. Figure 8 shows a comparison of the 64 most-common TI class frequencies with the P5, P50, and P95 results from MPS-GS simulation. At the relatively small scale of 3x3 cells the MPS-GS results match the most common MPH classes very well. There is more randomness towards the lower end of the graph, but at frequencies of 0.0001 and 62,500 cells there are only about 6 expected occurrences of a given class.

For comparison, two SISIM realizations are shown in Figure 9; the variogram reproduction is shown in Figure 10 and Figure 11 shows the MPH results. Visually the SISIM realizations are much worse than MPS-GS, showing no channel structure or long-range connectivity. The variogram reproduction of SISIM is no better than MPS-GS. The MPH reproduction is much worse for SISIM than MPS-GS, mainly because of the differences at the higher-frequency classes; the absolute difference from the TI MPH averages 0.07 for MPS-GS and 0.26 for SISIM.

Another 20 realizations were simulated using MPS-GS and SISIM, this time using hard conditioning data. The data locations are shown in Figure 12. An interesting feature of this hard data is that the facies proportions are reversed from the TI; 70% of the hard data are channel facies with only 30% background. The TI proportions were used at the target for both algorithms.

Figure 13 shows one MPS-GS realization and the simulated probability of sand from 20 realizations. The realizations itself looks similar to the unconditional ones in Figure 6; the probability of sand map shows channel and non-channel structure near the conditioning data, demonstrating that some of the TI structure is preserved over all realizations. Figure 14 shows one SISIM realization and the simulated probability of sand over 20 realizations. Again, the realization is very similar to the unconditional case; however, the univariate proportions are slightly skewed due to the data. The probability of sand from 20 SISIM realizations shows no channel structure on average, but rather looks similar to a kriging map with blobby ellipses and no realistic-looking geology.

For a third example case, soft data in the form of local sand probabilities were used. Figure 15 shows the local probabilities of sand. The structure shown in the local PDFs is not stationary, but has varying channel thicknesses and orientations; using a stationary TI will give a measure of how the MPS-GS algorithm can reconcile prior geologic knowledge and the likelihood distributions from sources such as seismic.

Figure 16 shows one MPS-GS realization and the simulated probability of sand at all locations. The distinct braiding of the channels in the TI may be seen in the realization, and the average map shows that the soft probabilities are reproduced as closely as possible while still using the TI for structure. Figure 17 shows an SISIM realization and the probability of sand from 20 realizations. The local probabilities are honoured, but there is no apparent channel structure resembling the TI or local PDFs.

Conclusions

The presented MPS-GS algorithm has the potential to reproduce complex geologic structure, characterized by a TI, while also honouring hard and soft data. The speed of the algorithm is quite good when compared to many of the other options for MPS simulation. The quality of the realizations is significantly improved over traditional variogram-based facies modeling methods.

There are a number of improvements that could be made to the MPS-GS algorithm. Better selection of the template, optimization of the number of MPEs and points per event, improvement in computational efficiency, using anisotropic multiple grids, and estimating conditional probabilities of a central MPE rather than a single-point indicator are all options for the future potential of MPS-GS. The current state of the algorithm is satisfactory for many purposes.

References

Anderson, E., Bai, Z., Bischof, C., Blackford, S., Demmel, J., Dongarra, J., Du Croz, J., Greenbaum, A., Hammarling, S., McKenney, A., Sorensen, D. (1999) *LAPACK User's Guide, 3rd Ed.* Society for Industrial and Applied Mathematics, Philadelphia, PA, 429 p.

Arpat, G.B. and Caers, J. (2007) Conditional Simulation with Patterns. *Mathematical Geology*, Vol. 39, No. 2, Feb. 2007, pp 177-203.

Caers, J. (2001) Geostatistical Reservoir Modelling Using Statistical Pattern Recognition. *Journal of Petroleum Science and Engineering*, Vol. 29, No. 3, May 2001, pp 177-188.

Casella, G. and George, E.I. (1992) Explaining the Gibbs Sampler. *The American Statistician*, Vol. 46, No. 3, Aug. 1992, pp 167-174.

Deutsch, C.V. (1992) *Annealing Techniques Applied to Reservoir Modeling and the Integration of Geological and Engineering (Well Test) Data*. Ph.D. Thesis, Stanford University, 304 p.

Deutsch, C.V. (2002) *Geostatistical Reservoir Modeling*. Oxford University Press, New York, 376 p.

Deutsch, C.V. and Journel, A.G. (1998) *GSLIB: Geostatistical Software Library and User's Guide, 2nd Ed*. Oxford University Press, New York, 369 p.

Eskandaridalvand, K. (2008) *Growthsim: A complete framework for integrating static and dynamic data into reservoir models*. Ph.D. Thesis, University of Texas at Austin, 199 p.

Gelfand, A.E. and Smith, A.F.M. (1990) Sampling-Based Approached to Calculating Marginal Densities. *Journal of the American Statistical Association*, Vol. 85, No. 410, June 1990, pp 398-409.

Geman, S. and Geman, D. (1984) Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, No. 6, Nov. 1984, pp 721-741.

Guardiano, F.B. and Srivastava, R.M. (1993) Multivariate Geostatistics: Beyond Bivariate Moments. Soares, A., Editor, *Geostatistics Troia '92*, Vol. 1, pp 133-144.

Liu, Y. (2006) Using the Snesim Program for Multiple-Point Statistical Simulation. *Computers & Geosciences*, Vol. 32, No. 10, Dec. 2006, pp 1544-1563.

Lyster, S. (2007) A Second Generation Gibbs Sampler for Simulation with Multiple-Point Statistics. *Centre for Computational Geostatistics*, No. 9, 26 p.

Lyster, S. (2008) Honouring Local PDFs in MPS Simulation. *Centre for Computational Geostatistics*, No. 10, 15 p.

Lyster, S. and Deutsch, C.V. (2008) MPS Simulation in a Gibbs Sampler Algorithm. *8th International Geostatistics Congress*, 10 p.

Lyster, S., Deutsch, C.V., and Dose, T. (2006) A New MPS Simulation Algorithm Based on Gibbs Sampling. *Centre for Computational Geostatistics*, No. 8, 16 p.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., and Teller, E. (1953) Equations of State Calculations by Fast Computing Machines. *Journal of Chemical Physics*, Vol. 21, No. 6, pp 1087-1091.

Robert, C.P. and Casella, G. (2004) *Monte Carlo Statistical Methods, 2nd Ed*. Springer Science and Business Media LLC, New York, 645 p.

Srivastava, M. (1992) Iterative Methods for Spatial Simulation. *Stanford Center for Reservoir Forecasting*, No. 5, 24 p.

Strebelle, S. (2002) Conditional Simulation of Complex Geological Structures Using Multiple-Point Statistics. *Mathematical Geology*, Vol. 34, No. 1, Jan. 2002, pp 1-21.

Wang, L. (1996) Modeling Complex Reservoir Geometries with Multiple-Point Statistics. *Mathematical Geology*, Vol. 28, No. 7, Oct. 1996, pp 895-907.

Zhang, T., Switzer, P., and Journel, A. (2006) Filter-Based Classification of Training Image Patterns for Spatial Simulation. *Mathematical Geology*, Vol. 28, No. 1, Jan. 2006, pp 63-80.

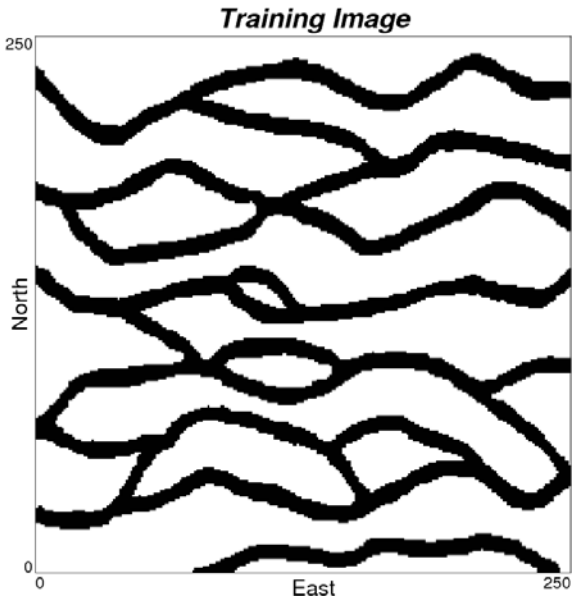

Figure 5: Training Image used for the example.

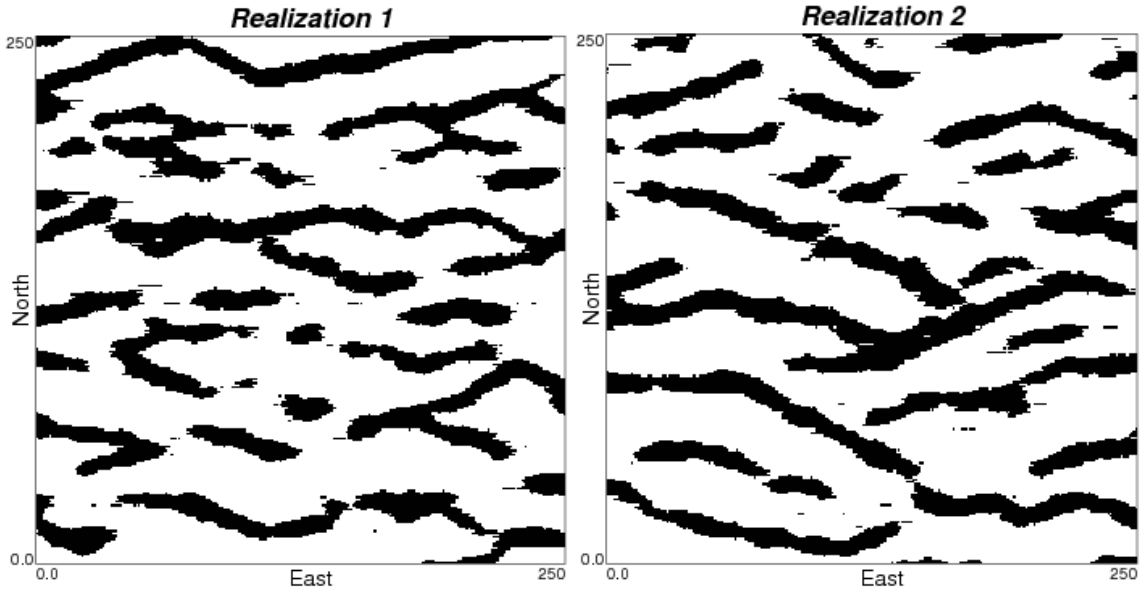

Figure 6: Two unconditional MPS-GS realizations using the TI in Figure 2.

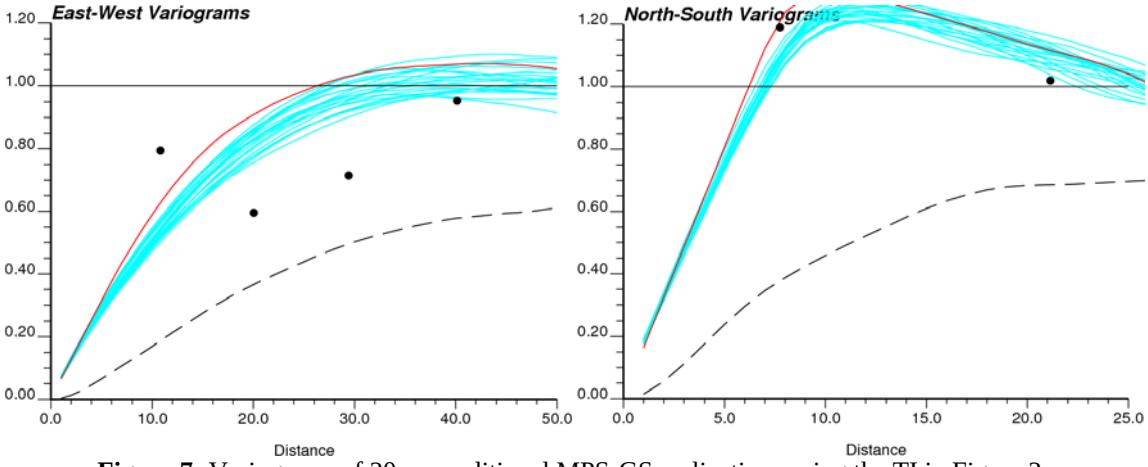

Figure 7: Variograms of 20 unconditional MPS-GS realizations using the TI in Figure 2.

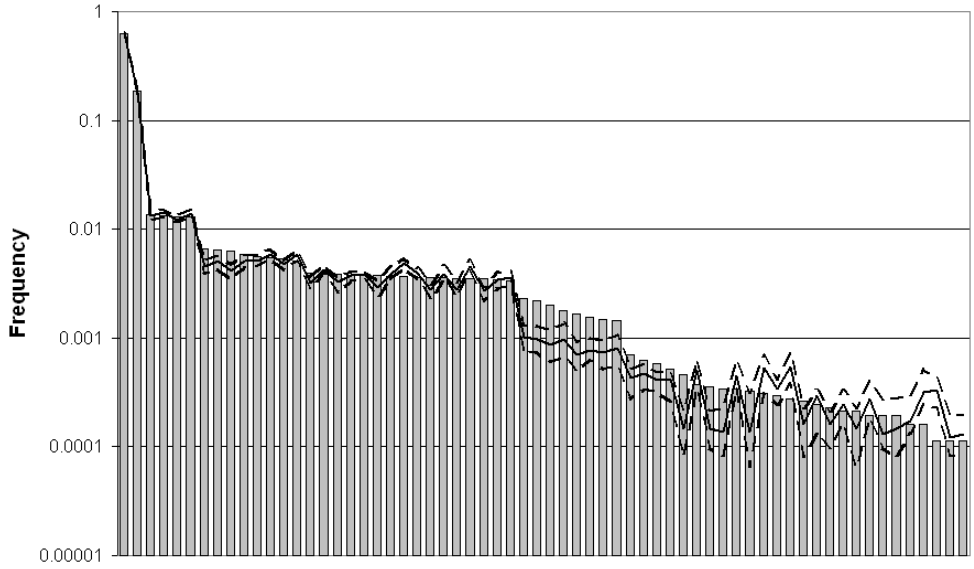

Figure 8: MP histograms of the TI (bars) and P5, P50, P90 values of 20 unconditional MPS-GS realizations (lines).

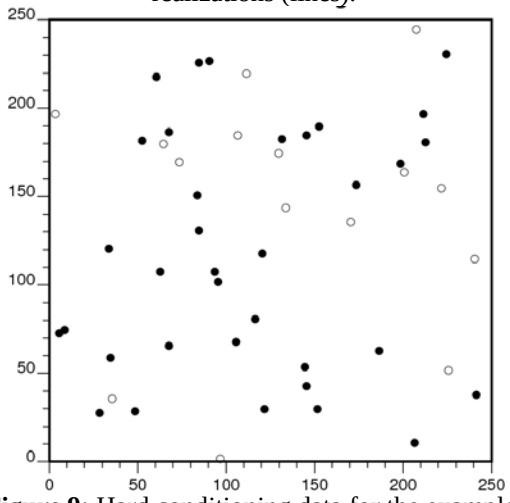

Figure 9: Hard conditioning data for the example.

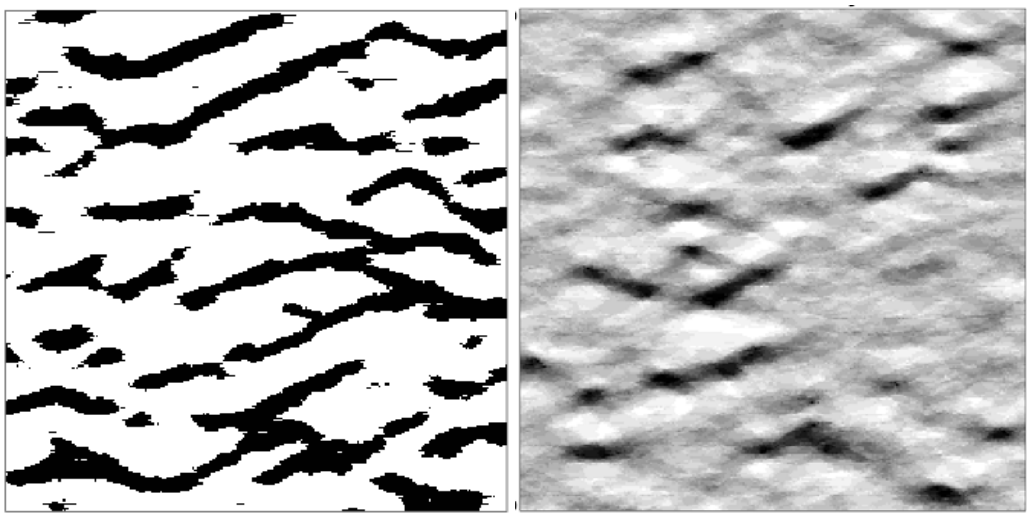

Figure 10: One conditional MPS-GS realization and the probability of sand over 20 realizations

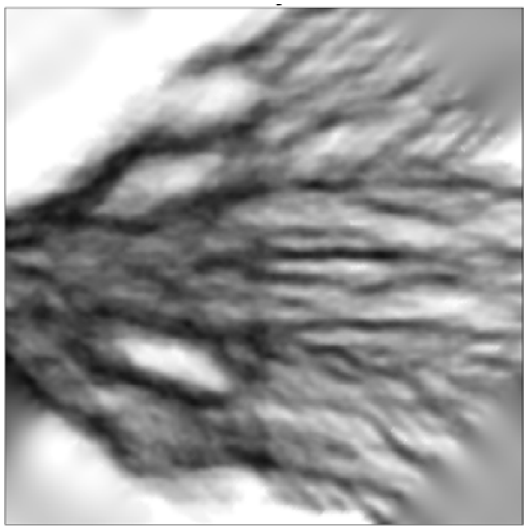

Figure 11: Soft local probability data for the example.

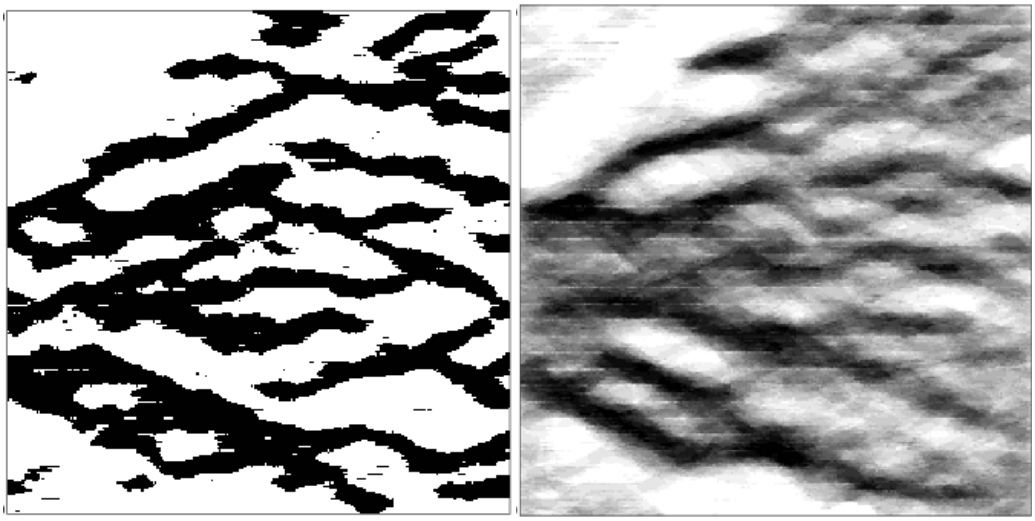

Figure 12: One MPS-GS realization using the local probabilities and the probability of sand all 20