Mathematical Geology, Vol. 38, No. 5, July 2006 (© 2006)
DOI: 10.1007/s11004-006-9031-2

The Necessity of a Multiple-Point Prior Model[1]

Andre Journel[2] and Tuanfeng Zhang[2]

Any interpolation, any hand contouring or digital drawing of a map or a numerical model necessarily calls for a prior model of the multiple-point statistics that link together the data to the unsampled nodes, then these unsampled nodes together. That prior model can be implicit, poorly defined as in hand contouring; it can be explicit through an algorithm as in digital mapping. The multiple-point statistics involved go well beyond single-point histogram and two-point covariance models; the challenge is to define algorithms that can control more of such statistics, particularly those that impact most the utilization of the resulting maps beyond their visual appearance. The newly introduced multiple-point simulation (mps) algorithms borrow the high order statistics from a visually and statistically explicit model, a training image. It is shown that mps can simulate realizations with high entropy character as well as traditional Gaussian-based algorithms, while offering the flexibility of considering alternative training images with various levels of low entropy (organized) structures. The impact on flow performance (spatial connectivity) of choosing a wrong training image among many sharing the same histogram and variogram is demonstrated.

KEY WORDS: variogram; connectivity; training image; pattern reconstruction; multivariate Gaussian model; multiple-point simulation.

INTRODUCTION

This paper suggests that a large enough (for ergodicity reasons) training image can be seen as a visually and statistically explicit *prior* random function model, one that discloses fully the prior statistics actually used by the simulation algorithm, whether single-point (histogram), two-point (covariance) or multiple-point up to the size of the scanning template. More exactly, a training image (Ti) is a realization assumed representative of that prior random function model. Stochastic simulation then consists of drawing alternative realizations from a random function model updated from that prior by both the local data and the specific implementations of the simulation algorithm used. Indeed the conditioning data retained need not be all consistent with the prior model, and any practical implementation algorithm necessarily entails approximations and some departure from

[1]Received 18 April 2005; accepted 8 December 2005; Published online: 9 December 2006.
[2]Department of Geological and Environmental Sciences, Stanford University, California 94305, USA
e-mail: journel@pangea.stanford.edu.

the prior model properties. For example, sequential Gaussian simulation consists of morphing (hence updating) a prior maximum entropy model to fit conditioning data that may reveal patterns and structure beyond the prior covariance input. It will be shown that multiple-point simulation (mps) can handle maximum entropy prior models as well as do traditional multivariate Gaussian algorithms, but can also handle a whole range of prior lower entropy structural models that would be more consistent with either the conditioning data or prior geological expertise.

To make our various points, a reference exhaustively known image interpreted as cracks distribution is sampled, and is given for reconstruction to the well established algorithm of sequential Gaussian simulation (GSLIB program *sgsim*, in Deutsch and Journel, 1998), then to a newly developed mps algorithm (program *filtersim*, in Zhang, Switzer and Journel, 2006). The ideal inference condition is considered for application of the *sgsim* algorithm, that are, the exhaustive reference variogram model and a representative data histogram is made available. As for *filtersim*, a range of training images is considered, from a maximum entropy Ti to a low entropy Ti, the latter reflecting an excellent prior knowledge of the crack structure. Both *sgsim* and *filtersim* algorithms are conditioned to the same sparse data set. The results are compared as simulated realizations and as estimated maps (E-type averages of many simulated realizations). The *filtersim* results based on a maximum entropy Ti reproduce that Ti covariance as well as *sgsim* realizations would reproduce the same covariance, but with a definite advantage in processing speed (CPU). However, if there is any prior inkling of the existence of cracks, the same *filtersim* code with a corresponding crack-type Ti, no matter how poor that Ti, would do better than the traditional *sgsim* approach.

The E-type average of *filtersim* realizations compare favorably with direct kriging estimates, particularly when these realizations capitalize on the additional structural information delivered by an appropriate Ti. Therefore and provided a representative Ti is available, mp simulation would outperform most traditional approaches that are only covariance-based, both in terms of simulation (reproduction of structures beyond the covariance) and estimation (local accuracy).

Interpreting the crack image as a permeability distribution, a water flood flow experiment is applied to the various simulated realizations generated. It is shown that for some well configurations the maximum entropy of *sgsim* realizations does not matter as long as the histogram (single-point statistic) of the reference image is correctly reproduced. However, for other well configurations, it is critical that the crack-type structure be reproduced with its fast flow conduits.

THE REFERENCE IMAGE

A soil crack image was selected from the published literature of computer sciences and image analysis (www.vision.ee.ethz.ch/∼rpaget/ nonparaMRFfast-Contents_Xu_Zhu_Shum_Guo.htm).

The original image gray scale was converted through a rank-preserving transform into a scale that could represent a permeability distribution with the cracks providing the high values tail. The resulting image could be seen as depicting a spatial distribution of low entropy, well structured, high permeability flow paths (the cracks) within a high entropy amorphous background. Figure 1 gives that reference image, its statistics, and the location of 50 samples taken at random yet whose set reproduces reasonably the exhaustive reference histogram. Figure 2 gives the exhaustive variograms calculated from the original values and their normal score transform along the two principal directions of anisotropy, NS and EW. These variograms were fitted by an anisotropic exponential model with a small nugget effect. This reference exponential model is made accessible to the task of reconstruction using the sequential Gaussian algorithm and program *sgsim*. This corresponds to an extremely favorable application situation; indeed, the experimental variograms calculated from the 50 data (not given here) were too noisy to allow any interpretation, a fact not uncommon in practice.

THE GAUSSIAN APPROACH TO SIMULATION

The traditional route to simulation of a continuous variable would call for the *sgsim* algorithm, which ensures reproduction of the data values at their locations, the data histogram and the input normal score variogram model. Beyond the variogram and histogram input the *sgsim* algorithm uses its own prior structural model, that is, its own set of multiple-point (mp) statistics. No simulation algorithm can operate without an mp structural model linking all data and all unknowns together, not only pair wise but all together jointly. In the case of *sgsim*, that model is multivariate Gaussian, a model that imposes maximum entropy that is maximum randomness beyond the input variogram and data histogram. Maximum entropy, far from being a no-model or a safe non-committing model (Journel and Deutsch, 1993), corresponds to a very particular training image, one that displays no pattern as exemplified by any non-conditional *sgsim* realization reproducing only the variogram input (Fig. 3). Of the six equally likely drawn realizations shown in that figure, the upper right one was retained as a high entropy training image for multiple-point simulation using the *filtersim* algorithm. That particular realization was chosen because it provides an excellent reproduction of the 50 data histogram and the input variogram model (check figure not given here).

Figure 4 gives six conditional *sgsim* realizations. These represent the very best one could get from a multivariate Gaussian algorithm, in that these realizations identify the 50 data values at their locations and reproduce well the data histogram and the input reference variogram, see the leftmost column of Fig. 9. Yet, none of these *sgsim* realizations could be considered a fair reproduction of the reference image of Fig. 1. The limitation is not in the implementation of the *sgsim* algorithm

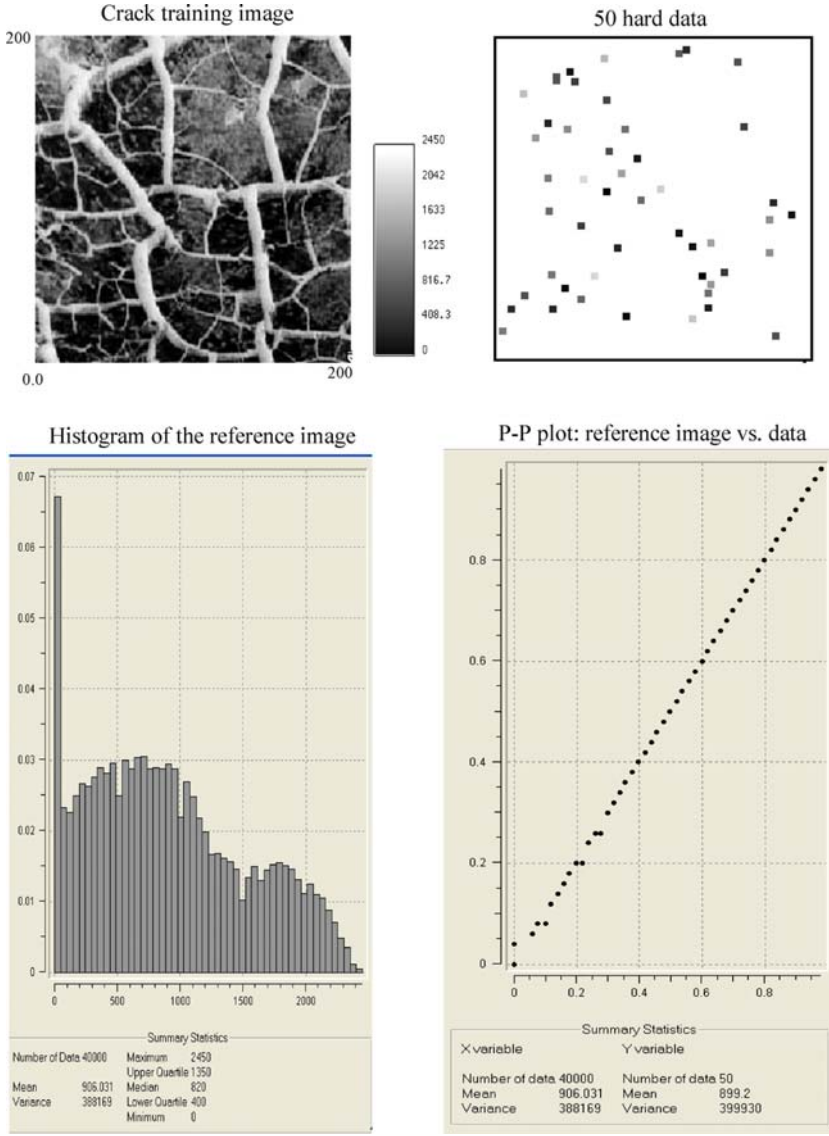

Figure 1. Reference crack training image, 50 hard data and their histograms distance.

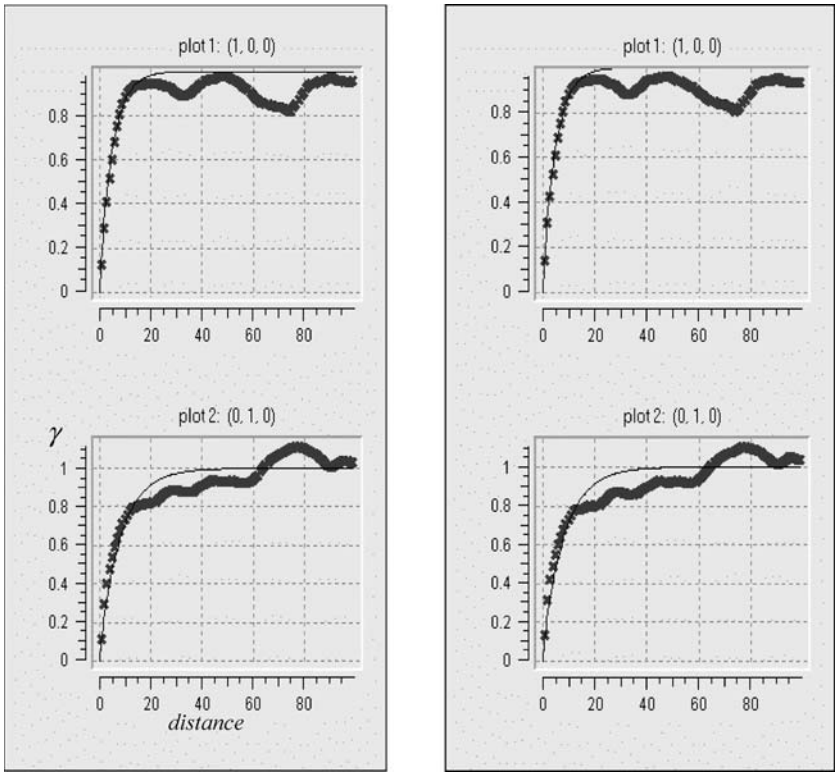

Figure 2. Reference semivariograms and their models fit, in original space (left column) and in normal score space (right column); upper plots are EW, lower plots are NS.

or in the input variogram model, it is in the inappropriateness of the training image implicit to the *sgsim* Gaussian model: that training image (anyone of Fig. 3) fails to render the patterns present in the reference image of Fig. 1.

THE mps APPROACH

If prior information exists about the type of structures to be reconstructed, that information should be used to improve on the random appearance of the realizations of either Fig. 3 or Fig. 4. Multiple-point simulation (mps) algorithms use such prior information as delivered by training images. Consider three such different Tis:

- The first Ti is precisely the maximum entropy *sgsim* non-conditional realization given in the upper right of Fig. 3, reproduced in the upper right

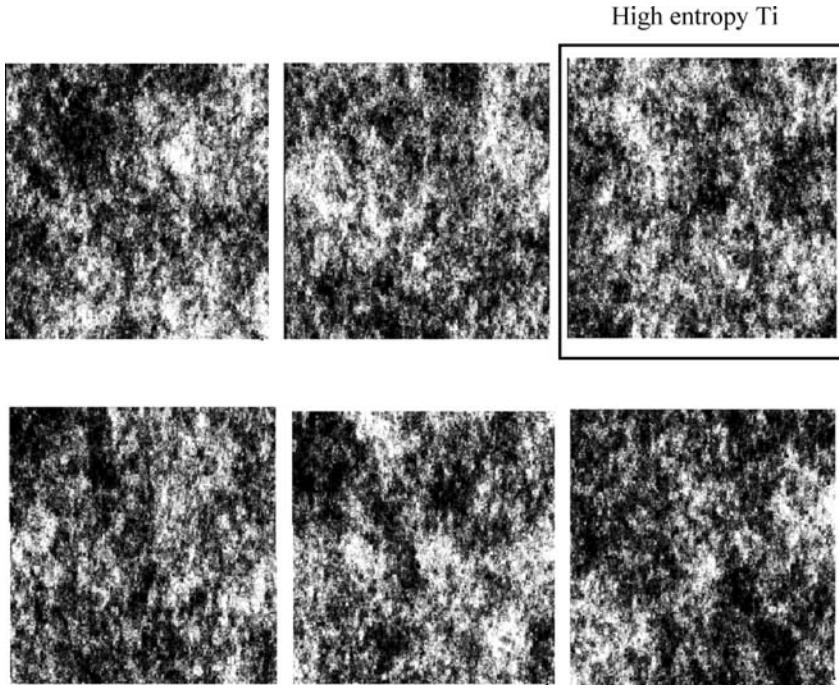

Figure 3. Six non-conditional *sgsim* realizations; the top right realization is taken as the high entropy training image.

of Fig. 5. The idea here is to compare the performance of mps and *sgsim* algorithms under the same Ti, that used implicitly by *sgsim*.

- The second Ti provides a low entropy excellent depiction of the reference image structures. It was generated as a non-conditional realization of an mps algorithm using for Ti the reference image itself, see the upper left image of Fig. 5.
- The third Ti was generated as a mixture of the two previous Tis, where at each location the attribute value is taken from one or the other Ti depending on a spatially distributed simulated indicator variable shown in the middle upper row of Fig. 5. That medium entropy Ti is shown at the bottom of Fig. 5; it could represent a structural model built from an imperfect prior knowledge of the type of cracks to be reproduced.

All three Tis share approximately the reference image histogram and variogram (check figures not given here). These Tis are only structural models, and as such they need not have any local accuracy: none honor any of the 50 data values. The task of any stochastic simulation, and for that matter of any estimation

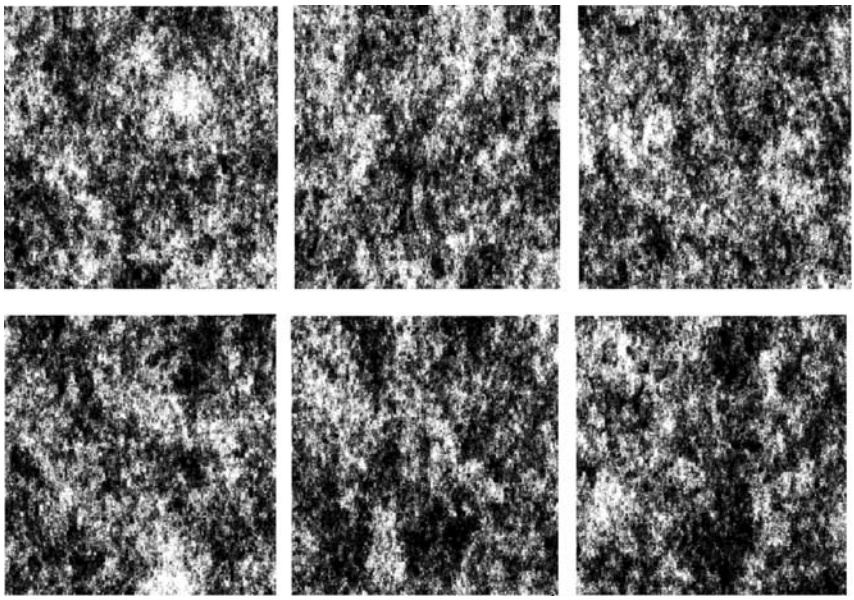

Figure 4. Six conditional *sgsim* realizations.

algorithm, is to produce maps which fill-in the voids between the local conditioning data. Estimation, however, is not required to reproduce the structural model used in the interpolation process.

A case can be made here for the construction of a digital catalog of "basic" training images from which to retrieve or build up any specific Ti reflecting one's prior vision of the structures deemed to prevail over the study field. Linear transforms, such as rotations followed by affinity (Caers and Zhang, 2004), and combination of these basic Tis as done in Fig. 5, would make that catalog far reaching.

The Filtersim Algorithm

The specific mps algorithm used for this demonstration is *filtersim* (Zhang, Switzer and Journel, 2006). In a nutshell, the filtersim algorithm starts by decomposing any training pattern into a few (6–9) linear filter scores which are weighted averages of the Ti values over moving windows of a specific template size. These Ti patterns are then classified in their multidimensional filter score space; the classification aims at regrouping similar patterns, where similarity is defined by a specific distance measure. That classification is done only once per Ti, prior to the

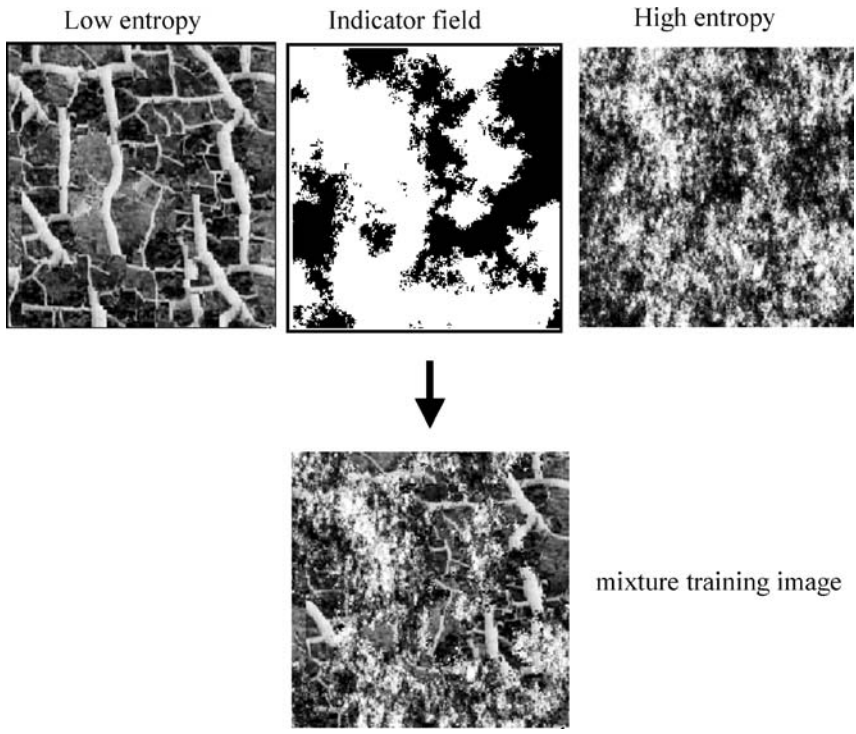

Figure 5. Mixture training image obtained by combining low and high entropy *images* with a simulated indicator field.

actual simulation process. Sequential simulation then proceeds along a random path visiting the unsampled grid locations: at each node the conditioning data event available within that node neighborhood template is related to the training class closest to it according to the previous distance; a training pattern is drawn from that class and patched onto the simulation grid centered on the node. The patched pattern consists of a template of point values; the few central values (how many is specified by an input parameter) are frozen not to be re-simulated, that is their locations are removed from the visiting random path; the other values will be replaced later by new simulated values as their locations are visited along the random path. Hard data conditioning is obtained by freezing the corresponding grid node values never to be changed during the course of simulation; these hard data contribute preferentially to the similarity distances thus ensuring that no training pattern would be retained if it creates discontinuities.

One could see the training part of *filtersim* as taking a training image, cutting it into puzzle pieces of the size of the scanning template, and then classifying similar pieces into different piles. Simulation proceeds like a puzzle reconstruction

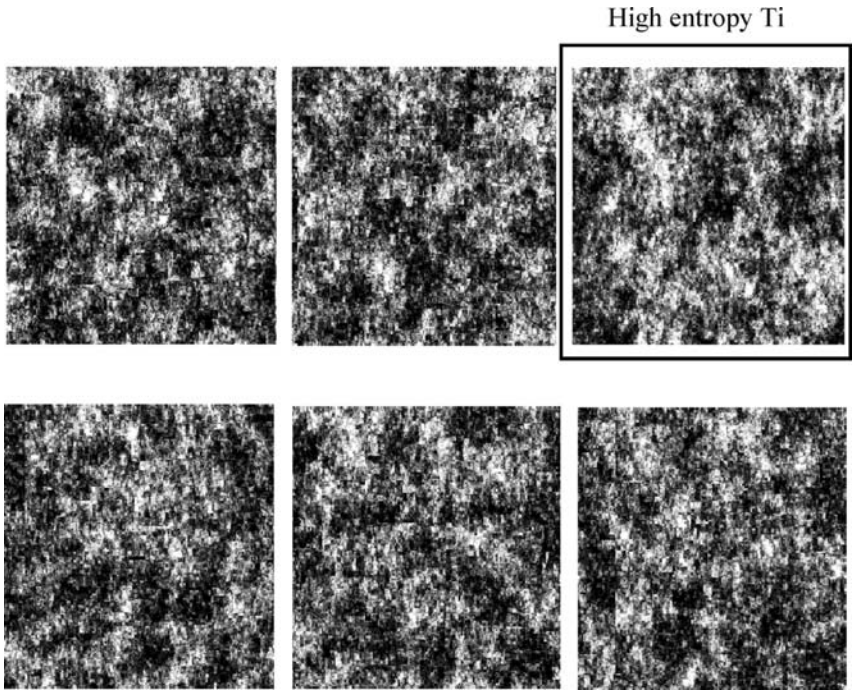

Figure 6. Five conditional *filtersim* simulations, using the low entropy training image given at right top.

by a sequence of digs into the appropriate pile, the pile that matches best the neighboring puzzle pieces already laid on the image, then patching the new piece onto the image being built. The piles never get exhausted because any piece taken out is immediately replaced by an identical one.

From a probabilistic angle, one could see these piles or classes of similar patterns as a distribution of multiple-point patterns built from the Ti. Simulation consists first of retrieving whatever conditioning data informs the local neighborhood to be simulated, then searching for the class matching best these data, last drawing a pattern from that class or conditional distribution.

Note that *filtersim* simulates patterns from pattern data; this is as opposed to the original mps algorithm, *snesim*, introduced by Strebelle (2000, 2002), which simulates single point values from pattern data. Also as opposed to *snesim*, the *filtersim* algorithm is not limited to simulation of categorical variables, it can handle continuous variables: its coding does not require large RAM memory because of the dimension reduction brought by the reduction of any mp pattern to a few (6 or 9) filter scores. In many regards, *filtersim* is closer to the *simpat* algorithm (Arpat

and Caers, 2005) which also patches patterns conditional to pattern data with, however, one major difference: *simpat* does not perform any prior classification and it looks for the single pattern most similar to the conditioning mp data.

FILTERSIM SIMULATION RESULTS

High Entropy Training Image

The mps *filtersim* algorithm was applied first using for Ti the non-conditional *sgsim* realization shown in the upper right of either Fig. 5 or Fig. 6. The resulting *filtersim* conditional simulations shown in Fig. 7 reproduce the maximum entropy character (lack of well structured patterns) of the conditional *sgsim* realizations of Fig. 4. RAM-wise *filtersim* is no more demanding than *sgsim*. CPU-wise, *filtersim* does not call for any kriging system construction and solution; at each node it requires only a data search (as does *sgsim*) then a distance calculation to find the closest class. The prior task of classification of training patterns is common to all *filtersim* realizations, and could be fully discounted if such classification is already part of a Ti catalog.

Medium Entropy Training Image

The *filtersim* algorithm was then applied using the second Ti shown at the bottom of Fig. 5. That Ti is a mixture of high and low entropy structures. The five filtersim conditional realizations shown in Fig. 7 indicate that the mps algorithm has succeeded to reproduce the style of variability and patterns depicted by the training image used. Detailed observation would reveal, however, artifact discontinuities due to the frozen inner part of the training patterns patched onto the simulation. Their occurrence can be reduced by decreasing the number of central values frozen when transferring a Ti pattern onto the image being built; such decrease leads to greater CPU cost and also to poorer reproduction of the Ti patterns. These artifacts do not affect, however, reproduction of the target statistics which are the 50 data histogram and the Ti variogram, as shown in the second column of Fig. 9. The target statistics, data histogram and Ti variogram, are tightly reproduced, see the third column of Fig. 9.

Note that this new run of the *filtersim* algorithm did not call for any re-coding: it suffices to change the input Ti file, everything else remains the same.

Although this simulation exercise has been built such that all Tis used share approximately the same histogram (that of the 50 conditioning data) and the same variogram model, the control of any higher order (mp) statistics is beyond the realm of *sgsim*. The *sgsim* realizations of Fig. 4 fail to reproduce the structures

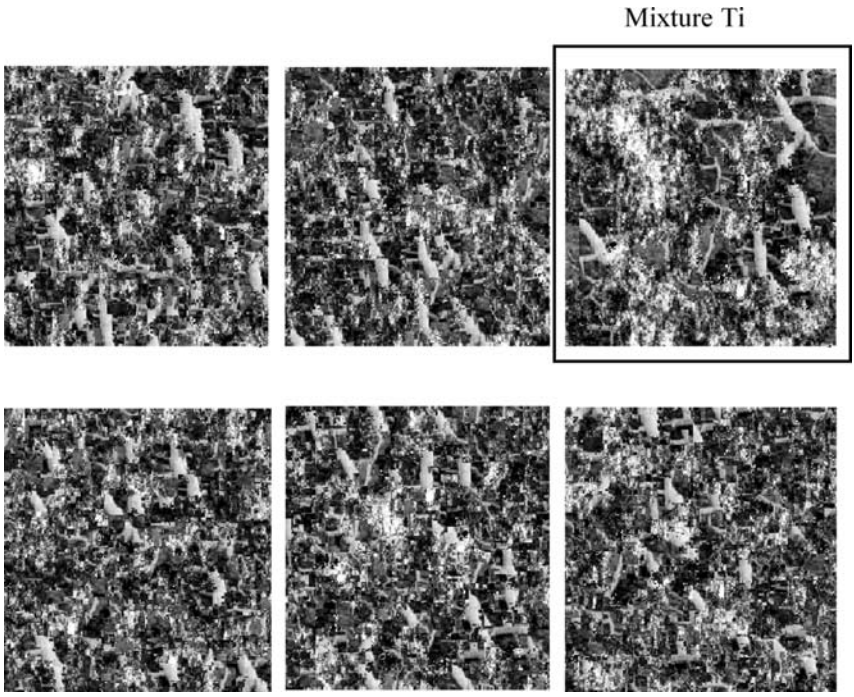

Figure 7. Five conditional *filtersim* simulations, using the mixture training image given at right top.

of the reference image of Fig. 1 even though they reproduce well its histogram and variogram. Even with a poor Ti, but one that indicates presence of cracks, the *filtersim* realizations of Fig. 7 fare much better. That visual evidence is later backed quantitatively in Fig. 10.

Low Entropy Training Image

If the training image depicts correctly the structures present in the reference image, as does the third Ti shown at the upper left of Fig. 5 and repeated at the upper right of Fig. 8, the resulting conditional *filtersim* realizations are excellent. Figure 9 (last column) shows that the 50 data histogram and Ti variogram are tightly reproduced by these *filtersim* realizations.

Multiple-Point Connectivity Comparison

To back the preceding visual comparisons, a specific measure of spatial connectivity was calculated on the reference image and its reproduction checked on

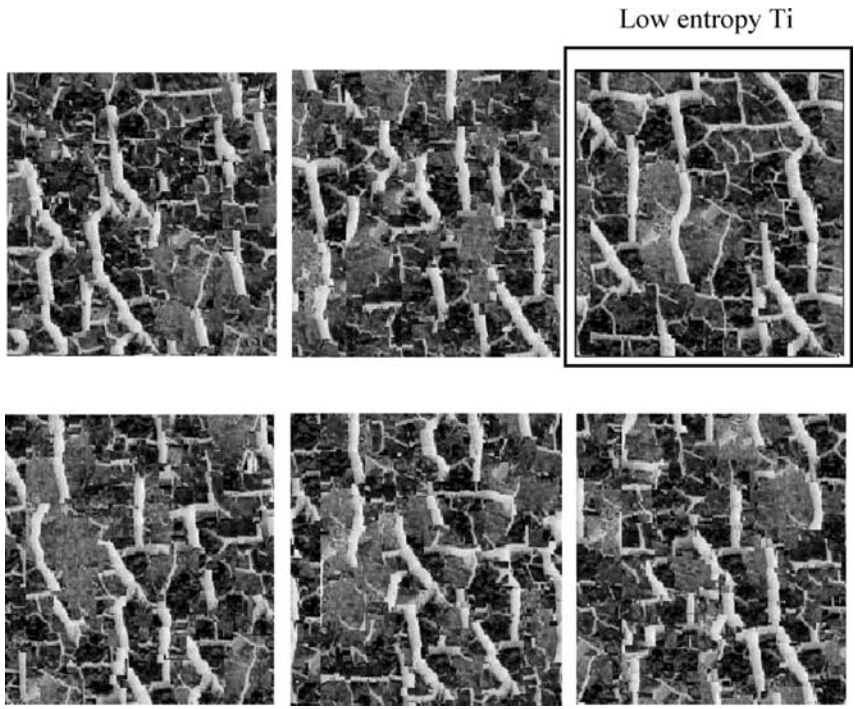

Figure 8. Five conditional *filtersim* simulations, using the low entropy training image given at right top.

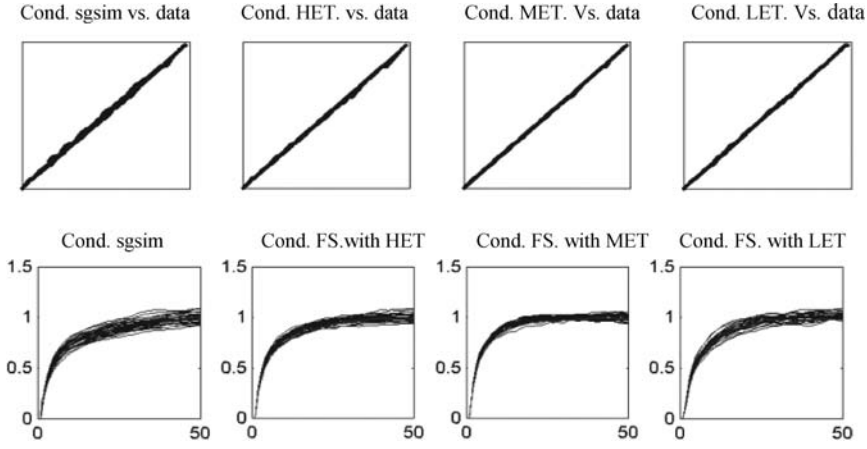

Figure 9. Reproduction of histograms (Q-Q plots) and NS variograms HET: high entropy Ti, MET: mixture Ti, LET: low entropy Ti *sgsim*: sequential Gaussian simulation, FS: *filtersim* simulation.

the various simulations. That measure is the multiple-point rectilinear connectivity proposed by Krishnan and Journel (2003). It is the proportion of rectilinear strings of n contiguous pixels or nodes all valued above a given threshold; that proportion is standardized to be 1 for $n=1$ and decreases as n increases along the particular direction considered, see Fig. 10. The greater these successive proportions, the higher in the graph lies the connectivity function indicating a better alignment in space (lower entropy) of values higher than the threshold considered.

Figure 10 gives the NS rectilinear connectivity measure of the four sets of simulated realizations, one set for *sgsim* at the far left, and the three sets for *filtersim*, one for each different Ti used. These simulated connectivity measures are all compared to the reference connectivity curve calculated from the reference image of Fig. 1; they are also compared to the connectivity curve of the Ti used. All realizations simulated with *sgsim* underestimate the reference connectivity; this is a consequence of their maximum entropy characteristic or maximum disorder beyond the 2-point statistics controlled by the input covariance model. As expected, the *filtersim* realizations using the correct low entropy Ti reproduce best the reference connectivity. Any set of realizations, whether originating from *sgsim* or *filtersim*, brackets well the connectivity curve of its respective Ti; this latter curve is not seen because it is buried within the corresponding cloud of realizations in Fig. 10.

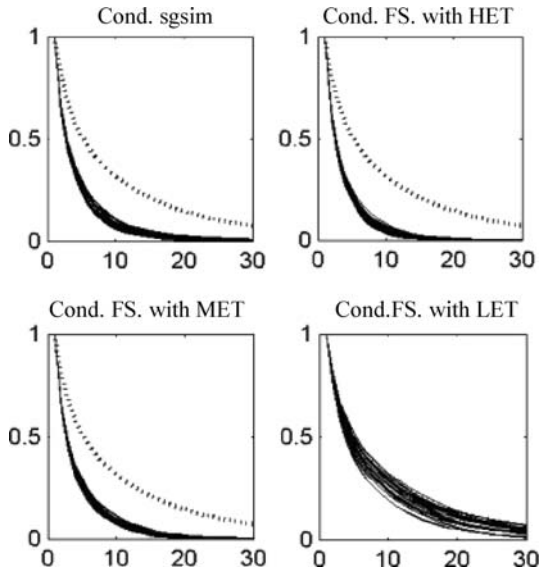

Figure 10. Reproduction of NS rectilinear connectivity (dots: reference image).

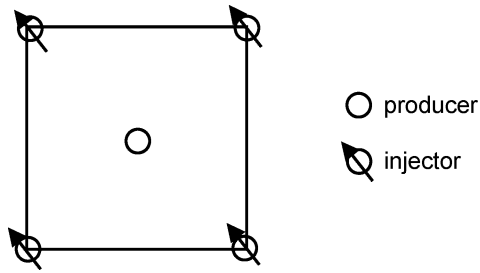

Figure 11. Water injection pattern. (4 corner injectors, 1 central producer; flow simulation run 1000 days). Field dimension: 200×200×1. Block size 15 ft × 15 ft × 500 ft. Layer top depth: 8000 ft. Constant porosity 20%. two phases: oil/water. Oil-water contact OWC: 10000 ft. 4 injectors at corners + 1 producer at the center of the field. Flow rate constraint for each injector: 15000 STB/day. Flow rate constraint 50000 STB/day first. Followed by 1000 PSI BHP constraint for producer.

Flow Experiment Results

To evaluate the impact of the crack pattern reproduction, a water flood experiment was conducted interpreting the reference image as a permeability distribution; see Fig. 11 for the specifics of the flow simulation performed.

Figure 12 gives the oil production rates and water concentrations vs. time calculated from 30 *sgsim*-simulated fields. Figure 13 gives the same results for 30 *filtersim*-simulated fields using the crack-revealing Ti of Fig. 8. These oil/water recovery curves are compared to the reference curves obtained from the reference image of Fig. 1. Somewhat surprisingly the high entropy models provided by *sgsim* bracket correctly the reference recoveries; there appears no bias induced by the absence of cracks (high permeability conduits) in the *sgsim* images. One possible explanation is that the reference high permeability cracks are all over the place with no preferential direction or cluster, a form of high entropy. The central location of the producer combined with the high mobility ratio allows an even sweep of the evenly distributed oil in place associated to the assumed constant porosity. However, had we considered the same 5 wells pattern but with 4 corner producers instead of a single central producer, the scores of the *sgsim* models would not have been so good. In any case, Fig. 13 indicates that the *filtersim*-based recovery predictions are equally unbiased and they bracket more narrowly the reference curves.

This apparently inconclusive flow experiment recalls us that two images can be similar for some global mp statistics (flow recovery values are examples of

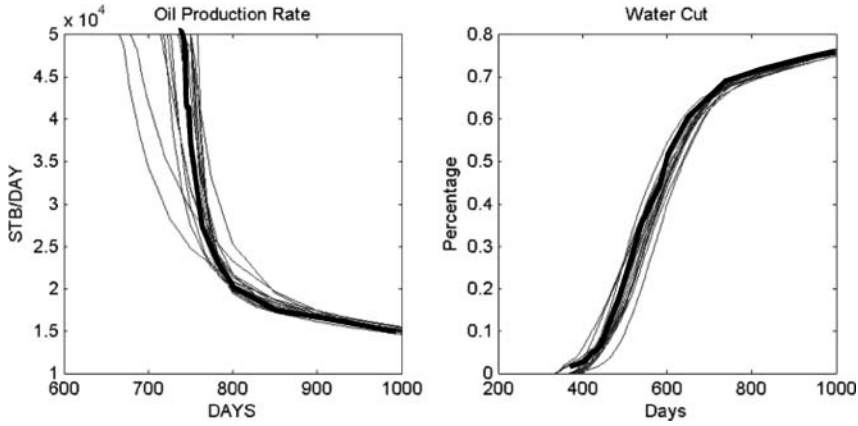

Figure 12. Oil production rate and water cut from 30 *sgsim* realizations (thick curve: reference field).

such statistics), and dissimilar for others. Our visual appreciation of the global difference or similarity of two images is based implicitly on many such global mp statistics, and from one person to another the particular statistics retained may be different. The accuracy of an estimated or simulated map depends on what that map or numerical model is going to be used for, and the human eye at times can be a poor referee.

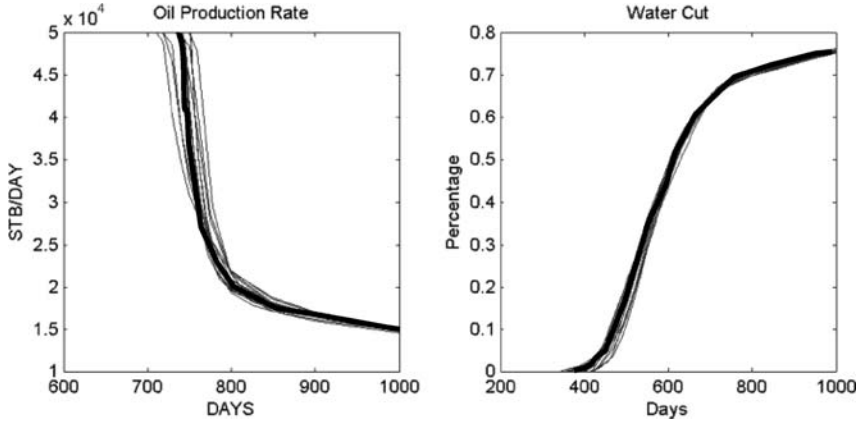

Figure 13. Oil production rate and water cut from 30 low entropy *filtersim* realizations (thick curve: reference field).

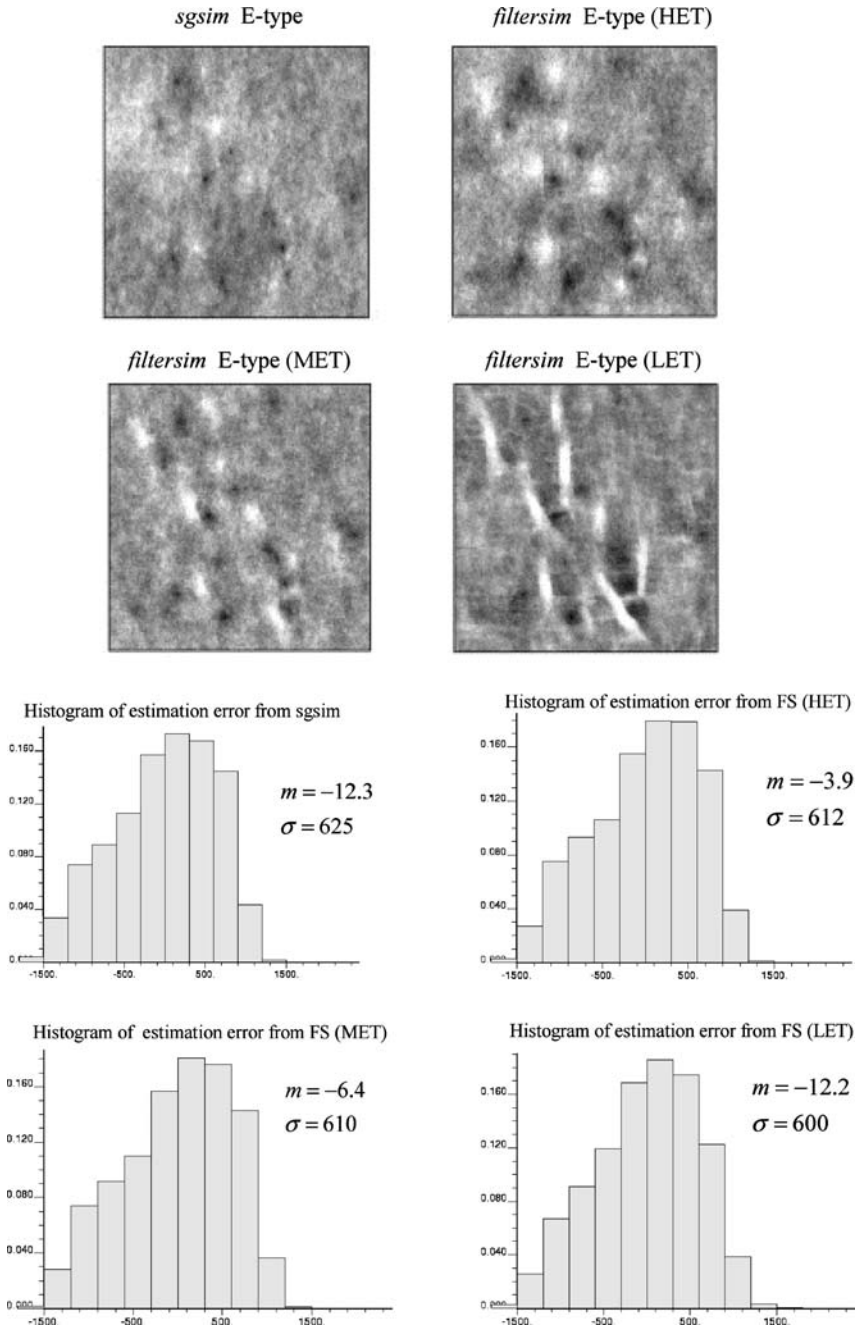

Figure 14. E-type estimated maps and histograms of point-wise estimation errors.

E-TYPE AVERAGES AS LOCAL ESTIMATES

Since all realizations of a given set of simulated values are equally likely to be drawn, their point wise average provides an estimate of the conditional expected value map, also called E-type estimated map. Each value of that E-type map can be seen as a least-squared error estimate of the corresponding unsampled value. The four E-type maps derived from the four sets of simulated values, one for *sgsim* and three for *filtersim*, allows evaluating the contribution of a specific Ti to local estimation.

Recall that in an estimation mode local accuracy is needed, as characterized for example by an estimation variance. In a simulation mode it is reproduction of the structural model which is sought, whether that structural model is a variogram or a training image.

Figure 14 gives the four E-type estimated maps with the corresponding four error histograms. That error is the difference between the E-type estimated value and the reference value at any one of the 40,000 nodes of the reference field. The estimation variance is but the variance of the error histogram. The more centered on zero (no bias) and the less spread the error histogram, the more locally accurate the estimation algorithm in a least square sense.

Figure 14 indicates that the better the Ti the better the estimation, both in terms of bias and error variance. As expected, the *filtersim* E-type using the high entropy *sgsim*-generated Ti has scores close to that of the *sgsim* E-type. The *filtersim* E-type using the best Ti provides the best estimation results: utilizing the correct multiple-point statistics as delivered by an appropriate Ti not only delivers the best simulation but also the best estimated map. In hindsight this is not surprising because both simulation and estimation build on structural information (variogram or Ti) in addition to local data; refer to the dual interpretation of kriging as fitting pieces of a surface or polynomial to the data, thus the more relevant these pieces the more accurate the final surface. Note however that the error histograms do not render full justice to the better E-type map delivered by *filtersim* with the correct Ti: this is because error histograms are but single-point statistics which cannot acknowledge reproduction of the few crack fragments seen on the low entropy estimation map.

CONCLUSIONS AND DISCUSSION

Least square algorithms, whether called regression or kriging, became practical in the earth sciences with the advent of computers which allowed sifting through large amount of data, setting up and solving the many required systems of linear equations. Stochastic simulations rapidly followed suit. Running tens of realizations of numerical models with millions of pixels or voxels are now

common practice on mere desktop computers. Yet, most theoretical algorithms underlying kriging or stochastic simulation, and for that matter most of geostatistics, are still anchored on covariance-based models which cannot display the full riches of explicit geological drawings.

The multivariate Gaussian model is one such example: prior to computers one could not dream getting from data more than a mean and a covariance matrix, hence the unique properties of the Gaussian model were critically convenient since these few statistics define fully all regressions (linear) and all conditional distributions (Gaussian). Simply said, no practical application was possible outside that Gaussian framework, the convenience of that model took precedence over what the data may say. The Central Limit theorems, maximum entropy principle and other Okham razors gave a veneer of theoretical justification to a model which is only remarkable for its concision thus convenience.

All medals carry two faces, no convenience come free of limitations. In a time of massive computing power, there is no more justification to accept blindly the limitations associated to parameter-poor Gaussian-related models. Data and prior geological concepts, no matter how complex they are, should be let to speak for themselves.

Corollary to the multivariate Gaussian model is the continued reliance on two-point or bivariate statistics and the belief that these may suffice. Indeed, a kriging system or direct sequential simulation (Journel, 1993) need not call for any Gaussian model, but they also deliver remarkably little. In the case of kriging, one gets only an estimate and an estimation variance, not any error distribution, hence no possibility to derive confidence intervals. The case of simulation is even more severe and revealing, since there cannot be any simulation without a multivariate distribution model, a model that links the data to the values being simulated and the simulated values together. That multivariate or multi-point (mp) model is typically not specified explicitly, it is in-built in the simulation algorithm used. Those undisclosed statistics are outside any control from either the operator or the data. The question is which arbitrariness is preferable, accept blindly the mp model in-built into the algorithm or try one's own mp model. We suggest that in many applications prior information does exist which allows building prior mp models more relevant to the phenomenon under study than the traditional assemblage "two-point statistics + some high entropy model". A training image is one vehicle to carry such prior mp models into the simulation algorithm. If that prior mp model is uncertain, alternative Tis could easily be considered. Our present computing ability allows trying many such prior models, there is no more reason to remain constrained by Gaussian-related models with their characteristic lack of spatial patterns. Analytical convenience is, no more, enough justification for accepting unrealistic looking simulated realizations that do not reflect prior knowledge about structures and patterns.

The crack case study presented in this paper demonstrates the flexibility of the training image-based multiple-point simulation (mps) approach. If indeed the image to reconstitute is deemed structure-free beyond an input variogram model, then a Gaussian-based simulation algorithm such as *sgsim* is relevant, but so would be an mps with a high entropy training image. If there is any prior knowledge about existence of specific structures, the mps alternative offers the flexibility to account for such information: it suffices to draw a training image reflecting that structural information (no local accuracy needed), and let mps anchor that Ti patterns to the actual conditioning data. Sensitivity analysis to alternative Tis is immediate, there is no need to change anything in the mps algorithm or code, just replace the Ti by another one and run again. The medium entropy Ti example considered in this paper indicates that even a poor training image, but one that has a minimum structural semblance with the actual phenomenon, would improve considerably the simulations. A severely wrong Ti would, of course, lead to non-representative simulated images, but so would a maximum entropy Ti if equally inappropriate. As earth scientists, we believe that high entropy (little or no spatial organization) models are *a priori* inappropriate; they should not be default models.

Our belief that structures or patterns matter should be mitigated by the flow simulation results presented: visually very different images may share (approximately) the same specific global statistics, a histogram, a variogram or a water rate curve as do the *sgsim* and *filtersim* realizations shown in this study.

An important result is the finding that mps has the potential to improve estimation in addition to simulation. The point wise E-type average of simulated realizations carries the multiple-point structural information of the Ti used; if that Ti is appropriate, the E-type average value at each node, taken as a local estimated value, equals kriging in the very error variance that kriging aims at minimizing and outperforms kriging in reproducing multiple-point structures where there are enough data to reveal them. There is no theoretical contradiction nor should this come as a surprise, since mps and thus its E-type average amount to fit a multiple-point covariance (actually a mp pattern) to the data, whereas kriging (seen under its dual version) is limited to fitting 2-point covariance models. The additional structural information used by mps, if relevant, leads to not only better simulation but also more accurate estimated maps. Of course, with mps one runs into the risk of using a totally wrong Ti; whether this could lead to much worse prediction results than a wrong but amorphous high entropy Ti depends on the application considered.

As for the common criticism about training images being difficult to get and carrying too subjective information it can be easily dismissed in view of the equally common difficulty in obtaining informative experimental variograms and the equally subjective, albeit rarely stated, assumption of maximum or high entropy beyond the uncertain variogram model.

REFERENCES

Arpat, B., and Caers, J., 2005, A multiple scale, pattern-based approach to sequential simulation, *in* Leuangthong, O., and Deutsch, C. V., eds., Geostatistics Banff 2004, Springer, Dordrecht, p. 255–264.

Caers, J., and Zhang, T., 2004, Multiple-point geostatistics: a quantitative vehicle for integrating geologic analogs into multiple reservoir models, *in* Grammer, G. M., Harris, P. M., and Eberli, G. P., eds., Integration of outcrop and modern analogs in reservoir modeling: Am. Assoc. Petrol. Geol., Memoir 80, p. 384–394.

Deutsch, C., and Journel, A. G., 1998, GSLIB: Geostatistical software library and user's guide, 2nd ed.: Oxford University Press, New York, 368 p.

Journel, A. G., 1993, Geostatistics: Roadblocks and challenges, *in* Soares, A., ed., Geostatistics: Tróia '92: Kluwer Academic, Dordrecht, v. 1, p. 213–224.

Journel, A. G., and Deutsch, C. V., 1993, Entropy and spatial disorder: Math. Geol., v. 25, no. 3, p. 329–356.

Krishnan, S., and Journel, A. G., 2003, Spatial connectivity: from variograms to multiple-point measures: Math. Geol., v. 35, no. 8, p. 915–925.

Strebelle, S., 2000, Sequential simulation drawing structures from training images: Unpublished doctoral dissertation, Stanford University, California, 193 p.

Strebelle, S., 2002, Conditional simulation of complex geological structures using multiple-point statistics: Math. Geol., v. 34, no. 1, p. 1–21.

Zhang, T., Switzer, P., and Journel, A. G., 2006, Filter-based classification of training image patterns for spatial simulation: Math. Geol., v. 38, no. 1, p. 63–80.