Reproducing Local Proportions in MPS Simulation

Steve Lyster

Simulation of geologic structures has advanced in recent years to use ever-more sophisticated techniques and account for more complex spatial structure. Multiple-point statistics (MPS) is a developing field that uses spatial relations of order greater than two to characterize the structure of resource deposits such as petroleum reservoirs and vein-type ore deposits. Increasing the complexity of the methods used in geostatistical simulation can lead to difficulties when trying to integrate all available information. In many reservoirs secondary data such as definable vertical trends, seismic surveying, or other geophysical information is available; in these cases the resource estimates and characterization of uncertainty is improved by using these "soft" secondary information sources. Reproducing these secondary information by using a variety of techniques will be explored and implemented in MPS simulation.

Introduction

The family of methods described by the term Multiple-Point Statistics share a common property: all use some form of spatial statistics that are of order greater than two and therefore go beyond the linear estimation paradigm of kriging. There are numerous methods which have been proposed and/or implemented to use MPS. The first, and most well-developed, simulation method using MPS is the single normal equation approach which uses Bayes' Law and direct inference of conditional probabilities from a training image to populate the simulation grid. This method was first proposed by Guardiano and Srivastava (1993) and was further developed in numerous papers including Strebelle et al (2002). Some other methods utilizing MPS are simulated annealing (Deutsch, 1992, Lyster et al, 2004), a Gibbs sampler algorithm (Srivastava, 1992, Lyster 2007), and neural networks (Caers, 2001).

These four methods all simulate facies indicators only and not continuous variables, which is typical of most MPS techniques. This is due to the relative ease of creating a categorical training image (TI) from which the MPS may be inferred; it is very difficult to directly characterize complex spatial structure from sparse data. TIs can be created by outcrop mapping, simulated geologic processes, simulation involving the geometric shapes seen in the deposit, expert interpretation, or any other desired technique (Boisvert, 2007).

Three of the four MPS simulation methods mentioned above are iterative algorithms. There are advantages to an iterative approach, which will be discussed later. Iterative methods have their own particular strengths and weaknesses but share a number of common issues both in theory and in implementation. The particular subject of interest in this paper is the use and reproduction of locally varying proportions of facies determined from secondary data.

Secondary Data and Local Proportions

"Soft" data, also called secondary or likelihood information, can be useful in characterizing subsurface phenomena. This information is faster, cheaper and easier to obtain than "hard" sample data such as drillhole samples or well logs; however, secondary data does not explicitly define which facies occur at any given location, rather giving probabilities which may be of varying usefulness and reliability. Secondary data defines local probability density functions (PDFs) which represent the proportion of each facies occurring at all individual locations. These local PDFs are useful in that they define areas likely to contain the facies of interest, such as petroleum-bearing channel sands or high-grade ore veins.

Methods for Using Local PDFs

There are a number of possible methods for using local PDFs obtained from secondary information. These include: Bayesian updating or the assumption of full independence between data sources; the permanence of ratios framework; tau and lambda models which account for data redundancy; and the servosystem or additive method. All of these methods determine the probability of a facies, A, given the hard data, B, and the secondary information, *C*, to form a single conditional probability. Regardless of the method utilized one of the most important considerations for local PDFs is honouring values equal to zero or one, which correspond to full certainty in the absence or presence of facies.

Bayesian updating (Deutsch, 2002) is a very simple multiplicative approach that is fast and easy to apply. The Bayesian updating method, in a general sense, is used to integrate prior and likelihood distributions into an updated posterior distribution. Assuming there is no variance for the likelihood PDF determined from secondary data, the equations simplify to:

$$\frac{P(A|B,C)}{P(A)} = \frac{P(A|B)}{P(A)} \cdot \frac{P(A|C)}{P(A)} \tag{1}$$

The Bayesian updating equation may be used to integrate any number of primary and secondary data; however, as shown here there is the implicit assumption of full independence between the data (Journel, 2002) and this assumption is often unrealistic. The updated distribution as shown in Equation 1 is often more extreme than it should be when the local PDF is higher or lower than the global distribution.

More stable than full independence is the permanence of ratios approximation (Journel, 2002). This approach assumes the incremental information provided by the secondary data is the same before and after the hard data is taken into account. Permanence of ratios uses the following definitions:

$$x = \frac{1-P(A|B,C)}{P(A|B,C)} \qquad a = \frac{1-P(A)}{P(A)}$$

$$b = \frac{1-P(A|B)}{P(A|B)} \qquad c = \frac{1-P(A|C)}{P(A|C)}$$

Using these definitions, the probability of a facies that accounts for both hard and soft data can be determined from *x* which can be found by the relation:

$$\frac{x}{a} = \frac{b}{a} \cdot \frac{c}{a} \tag{2}$$

This approach has the benefit of constraining the results between zero and one, as long as the initial PDF and CPDF are both licit. However, there is still the assumption of independence between the data sources.

The reintroduction of data redundancy may be accomplished with what are termed a “tau model” or “lambda model” (Journel, 2002, Hong and Deutsch, 2007). The basic idea of this approach is to reduce the influence of the different data types by adding an exponent to Equation 2 as follows:

$$\frac{x}{a} = \left(\frac{b}{a}\right)^{\tau_b} \cdot \left(\frac{c}{a}\right)^{\tau_c} \tag{3}$$

The parameters may be set to any values desired, with 1.0 not affecting the data influence at all, 0.0 removing the influence entirely, and negative values reversing the influence of the data. There are a number of possible approaches to determining the parameters, and different values may be used for different data types and even for each facies (Hong and Deutsch, 2007). This method retains the licit PDF properties of the permanence of ratios, no matter what the controlling parameters are set to. Often the parameter on the hard sample data is left as 1.0 and the secondary data are scaled on that basis.

Another method for honouring the local distributions involves adding the difference between the local and global PDF values to the conditional probabilities that have been estimated whether these estimates are determined through indicator kriging, Bayes’ law, MPE estimation, or any other method.

$$P(A|B,C) = P(A|B) + P(A|C) - P(A) \tag{4}$$

This approach is similar to a servosytem correction for matching global facies proportions (Strebelle and Journel, 2001) and the two may be easily combined in the following equation:

$$P(A|B,C) = P(A|B) + \mu \cdot \left[P(A|C) - P^{sim}(A) \right] \tag{5}$$

The μ value is a controlling parameter, typically set to 1.0 but it may be lowered to reduce the influence of this correction if the prior statistical model is not properly reproduced.

Using an additive method such as this has the disadvantage of not explicitly honouring local PDF values of zero or one. It is relatively easy to enforce these locations by setting the unsampled locations with a PDF value of one to the appropriate facies, then not visiting these locations; this greatly improves the efficiency of iterative simulation methods in particular as each location is be visited a number of times. To honour values of zero in a local PDF, a check must be carried out to ensure the respective facies are not selected by the algorithm used.

Iterative Simulation Considerations

There are several advantages to iterative simulation methods. The main reason for using a technique where every location must be visited a number of times is the computational efficiency of not having to search or account for nearby samples in an irregular, unknown pattern (Srivastava, 1992). There is, however, difficulty in matching a locally varying PDF in iterative methods. The most direct way to reproduce local PDFs at the beginning of an iterative realization is to draw the initial image from the local, rather than global, distributions. This ensures reproduction initially, although as the algorithm proceeds there will no doubt be large changes made to the entirely random structure of the initial image.

Using the methods outlined above for honouring local information, a significant problem was encountered when using the Gibbs sampler MPS algorithm outlined in Lyster, 2007. The local PDFs were greatly exaggerated due to the repeated application of the correction methods; rather than use Bayesian updating or permanence of ratios only once, as in a sequential simulation algorithm, whichever PDF correction approach was used was applied at every location on each pass through the grid. With the local PDF being applied a number of times it tended to overwhelm the spatial structure of the prior model determined from the TI and give unrealistic results.

A proposed solution to this problem is to apply the servosystem additive correction within a user-defined number of “bins”, $B_k = 1,\ldots,N_B$ for all K facies. The local PDF distributions are divided up into the appropriate number of bins, and each location falls within one bin for each facies. The global simulated proportion of each facies within each bin is then corrected by a servosystem, so for example at locations where a certain facies is expected 90% of the time the global proportions within that bin are in fact 90%. The corrections to the conditional probabilities with this approach are:

$$P^{**}(k) = P^{*}(k) + \mu \cdot \left[P^{LOC}(k) - P^{SIM}(k, B_k) \right] \tag{6}$$

Discretizing the local PDFs in this manner leads to reproduction of the global PDFs within each bin; this is equivalent to honouring the local PDF values. The B_k bin numbers may be calculated easily using the equation

$$B_k = \mathrm{int}\left(\frac{P^{LOC}(k) - P^{MIN}(k)}{P^{MAX}(k) - P^{MIN}(k) + \varepsilon} \cdot N_B \right) + 1 \tag{7}$$

The epsilon parameter in Equation 7 is a very small number, say 10^{-10}, which is used to ensure all local PDF values are assigned valid bins greater than zero and no greater than N_B. This multiple servosystem method shows promise for iterative methods, and the results in practice will be explored.

Examples

To illustrate the problem of exaggeration of high and low local PDF values, several of the methods for utilizing secondary data were used in unconditional MPS simulation using a Gibbs sampler algorithm (Lyster, 2007). The training image used for these simulations is shown in Figure 1. There are four facies, ranging in reservoir quality from background shale to high-quality channel sands, with low- and medium-quality lobes. This TI and data set may also be seen in Boisvert, 2007, and Hoffman et al, 2005.

The model of the reservoir is 78 by 59 by 116 cells, with cell sizes of 75 by 75 by 6 feet. The TI was created to characterize the good reservoir portion of the model, and for this reason the global univariate statistics are greatly different than the data set. The top and the bottom of the model have little good reservoir facies, with the majority of the net sands concentrated in two layers towards the middle of the model vertically. The highly nonstationary facies proportions make this data set ideal for this type of study.

A locally-varying PDF model was generated by creating a simple vertical trend from the well data. The trend models are shown in Figure 2. The curves for facies 2 and 3 are the noisiest, as they have low overall global proportions. The trend for facies 1 clearly shows the two regions of good net reservoir, one at about 300 feet and the other at about 500 feet; these elevations have the lowest proportion of background shale and higher proportions of the better facies. The distributions of the local PDFs are shown in Figure 3.

Figure 4 shows an unconditional realization which was simulated using the MPS Gibbs sampler and no trend, with the target global facies proportions derived from the average of the trend model. The accuracy plots over ten realizations for the four different facies does not reproduce the local proportions, which is expected.

The simulation was repeated, this time accounting for the trend and using the Bayesian updating method in Equation 1; Figure 4 shows slices of a single realization. It is obvious that the vertical trend is grossly overstated in the resulting unconditional realization as the Bayesian updating is applied a number of times at all locations. The local proportions are not reproduced.

In similar examples, Figure 4 shows realizations that use permanence of ratios and the additive method, respectively, to account for the local PDF information. While there are some subtle differences the same overall results may be seen as for the Bayesian updating. The accuracy plots confirm that the local facies proportions are not reproduced.

Figure 4 also shows a realization which uses the multiple servosystem approach to reproduce the vertical trend. This realization is clearly better than those using the other three local PDF methods; the accuracy plots in Figure 5 verify that the local facies proportions are reproduced reasonably well. The multiple servosystem method is somewhat ad-hoc, but does not affect the speed of the algorithm at all compared to the other varying PDF methods.

For comparison, an unconditional sequential simulation was performed using the SISIM_LM program; a realization is also shown in Figure 4. The accuracy plots are similar to those of the multiple servosystem iterative method (Figure 5). This helps support the practicality of the proposed approach.

To explore how the complexity of MPS may make integrating secondary data more difficult, a simulation was performed using the Gibbs sampler algorithm but only second-order statistics derived from the TI. This is more robust than the simple kriging in the SISIM_LM program but contains less information than the MPS used previously. A realization with this technique is shown in Figure 4, and the corresponding accuracy plots for ten realizations are similar to those of the multiple servosystem iterative method (Figure 5). The vertical trend is reproduced with acceptable variation from the target proportions and the results are similar to the MPS and sequential simulations.

Discussion

Traditional methods for reproducing local distributions, which are useful in sequential methods, are not appropriate for iterative methods due to the nature of these techniques. Many MPS simulation algorithms use an iterative framework and as such need a different approach to incorporate secondary data, trends, and local PDFs. The multiple servosystem approach proposed here reasonably reproduces a vertical trend derived from hard well data. Multiple servosystems with a user-defined number of bins has been incorporated into the MPS Gibbs sampler algorithm under development.

While a smooth vertical trend is easily reproduced, seismic-derived PDFs with many values equal to zero or one has not been explored yet. The peculiarities of hard boundaries may present problems with calculation of conditional probabilities and needs to be investigated further. As of yet the assumption has been that the user will incorporate all available data into a single PDF model before commencing simulation.

A potential problem may arise when several distinct regions with similar PDFs exist within a model. In these cases, the multiple servosystem approach does not distinguish one region from another and if the local PDFs are in the same bins they will be treated as a single region to be simulated with one global PDF. This may or may not lead to problems and could be remedied by allowing the user to explicitly define zones, or by simulating each distinct geologic area separately.

References

Boisvert, J.B. (2007) *Mineral Deposit Modeling with Pseudo-Genetically Constructed Training Images*. M.Sc. Thesis, University of Alberta, 72 p.

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.

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.

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

Hoffman, B.T., Wen, X.H., Strebelle, S., and Caers, J. (2005) Geologically Consistent History Matching of a Deepwater Turbidite Reservoir. *SPE Annual Technical Conference and Exhibition*, Dallas, TX, October 2005, 12 p. SPE #95557

Hong, S. and Deutsch, C.V. (2007) Methods for Integrating Conditional Probabilities for Geostatistical Modeling. *Centre for Computational Geostatistics*, No. 9, 14 p.

Journel, A. (2002) Combining Knowledge from Diverse Sources: An Alternative to Traditional Data Independence Hypotheses. *Mathematical Geology*, Vol. 34, No. 5, July 2002, pp 573-596.

Lyster, S., Ortiz, J.M, and Deutsch, C.V. (2004) MPASIM: A Multiple Point Annealing Simulation Program. *Centre for Computational Geostatistics*, No. 6, 18 p.

Lyster, S. (2007) A Second Generation Gibbs Sampler for Simulation with Multiple-Point Statistics. *Centre for Computational Geostatistics*, No. 9, 19 p.

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

Strebelle, S.B. and Journel, A.G. (2001) Reservoir Modeling Using Multiple-Point Statistics. SPE Annual Technical Conference and Exhibition, New Orleans, Oct. 2001, 11 p. SPE #71324.

Strebelle, S., Payrazyan, K., and Caers, J. (2002) Modeling of a Deepwater Turbidite Reservoir Conditional to Seismic Data Using Multiple-Point Geostatistics. *SPE Annual Technical Conference and Exhibition*, San Antonio, TX, Sept.-Oct. 2002, 10 p. SPE #77425.

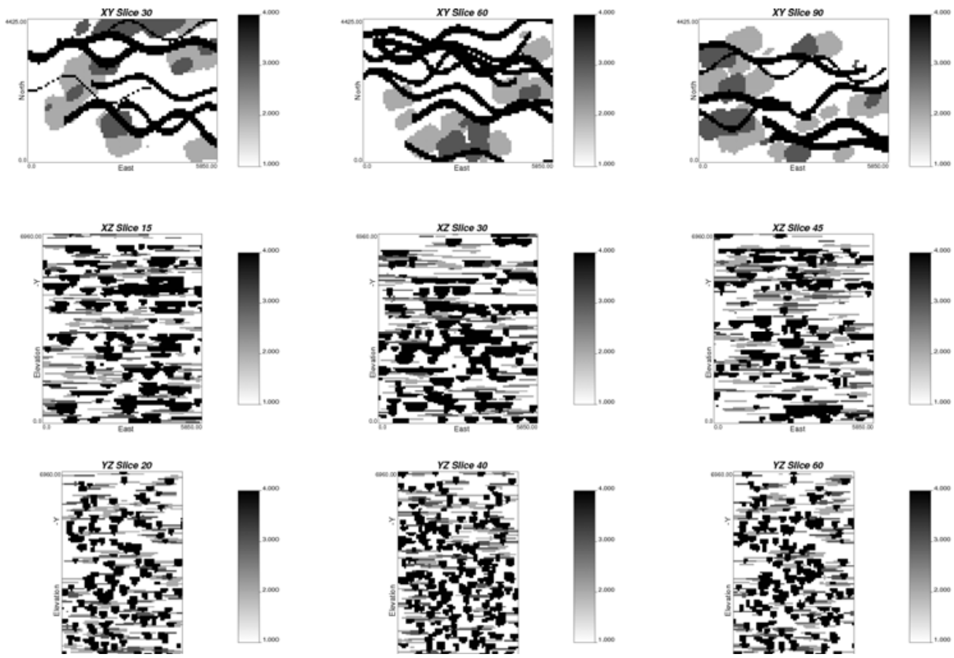

Figure 1: Four-facies turbidite training image used for the example.

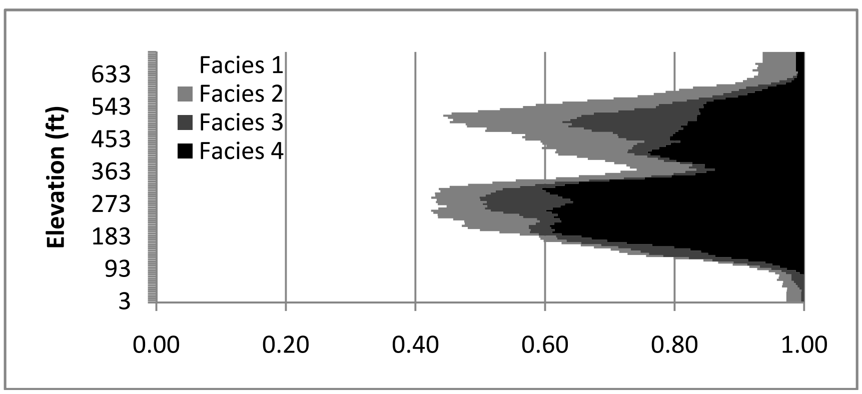

Figure 2: Vertical trend model of the facies proportions.

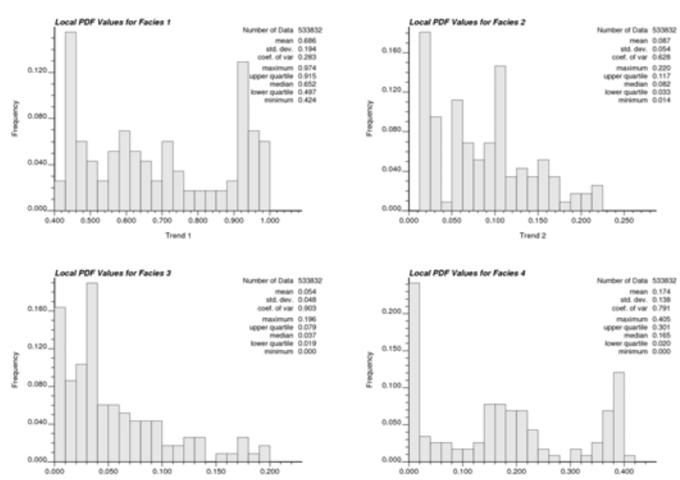

Figure 3: Distributions of the local PDF values for the example.

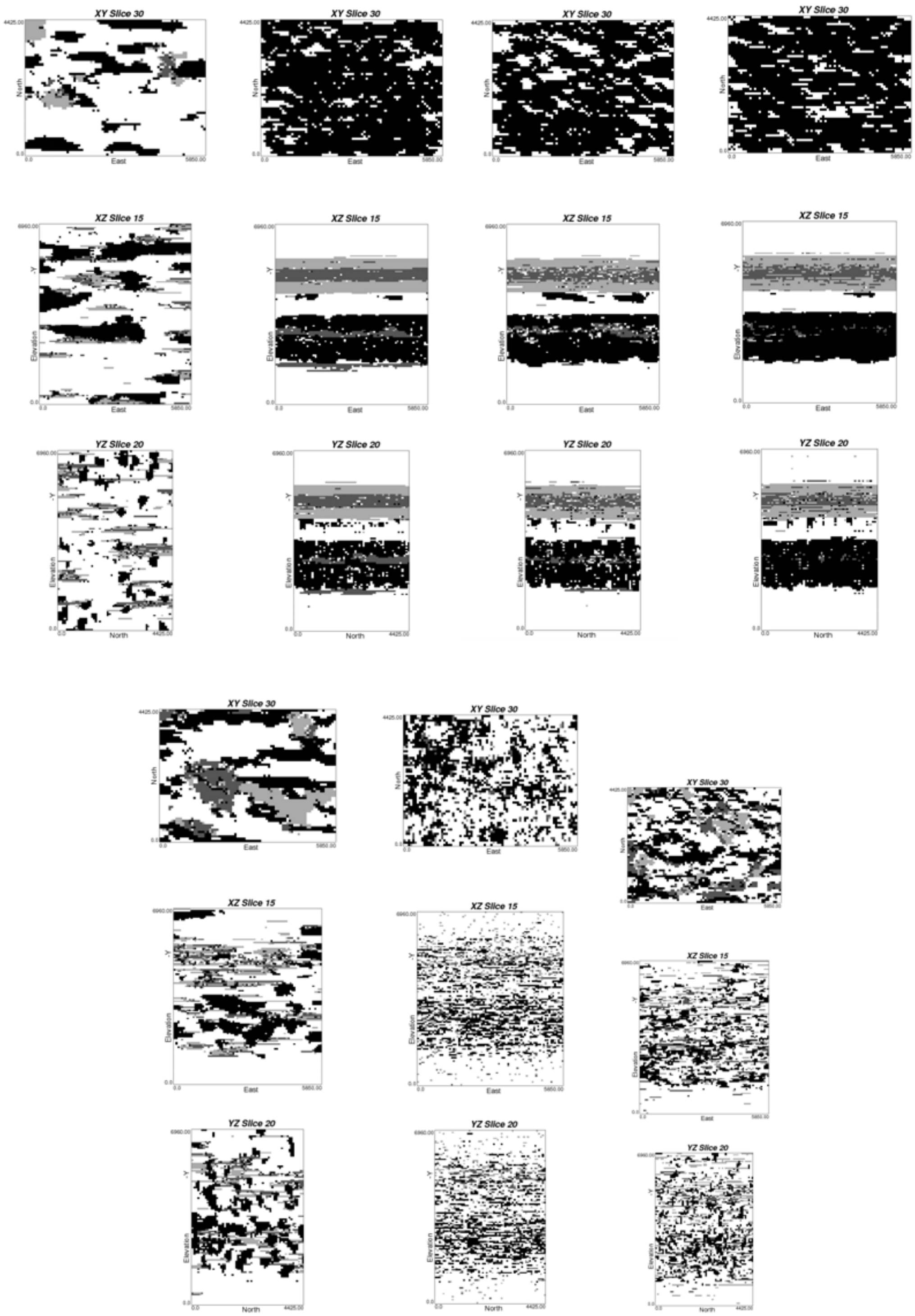

Figure 4: Top row from left to right: unconditional MPS realization without using a trend, Unconditional MPS realization using a trend and Bayesian updating, MPS realization using a trend and permanence of ratios, and MPS realization using a trend and the additive method. Bottom row from left to right: MPS realization using a trend and the multiple servosystem method and SIS realization using a trend and the SISIM_LM program, an iterative realization using covariances, a trend and multiple servosystems.

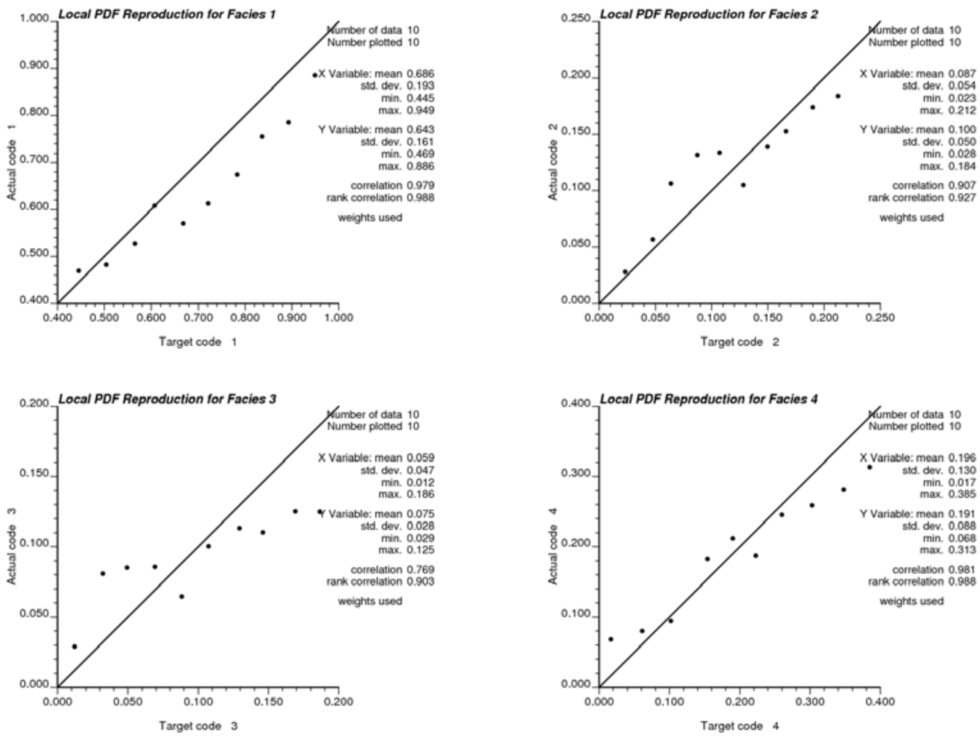

Figure 5: Accuracy plot for the multiple servosystem method realization compared to the input trend.