# Variability of Multiple-Point Statistics

Steven Lyster

Centre for Computational Geostatistics
Department of Civil and Environmental Engineering
University of Alberta

*The use of multiple-point statistics in simulation of mineral deposits has been gaining popularity in recent years. In order to infer complex spatial statistics, conceptual training images are used which are deemed fully representative of the area under study. Multiple-point measures may then be used to characterize the quality of the resulting simulated realizations, similar to how variogram reproduction is used to deem realizations acceptable or unacceptable. However, some uncertainty in multiple-point statistics may arise from the limited size of the training images used. There is inherent variability to any statistics based on finite models, and this may confuse the multiple-point measures such as the histogram. This paper explores the variability underlying complex spatial structure and how this relates to the expected variability arising from multiple-point simulation methods.*

## Introduction

Multiple-point statistics (MPS) are being increasingly used for geostatistical simulation. These methods are intended to reproduce the complex features of conceptual training images (TIs) and to go beyond second-order moments such as covariance or the variogram. While high-order structure may be enforced in simulation algorithms (Liu, 2006 and Strebelle, 2002 among others) determining the quality of reproduction of MPS largely involves visual inspection of the resulting realizations.

The use of MPS as a diagnostic tool for ranking realizations has been explored. Connectivity functions (Ortiz, 2003 and Liu, 2006), which have also been referred to as runs (Boisvert, 2007), measure the probability of finding consecutive sequences of the same, or similar, facies. These functions may be plotted similar to a variogram and inspected for closeness to either strings of sample data or the connectivity of the TI used for simulation. Similar to variogram reproduction, this comparison is largely based on expert judgment as to what constitutes “good” or “bad”. Efforts have recently been made to quantify this relation numerically (Boisvert, 2007).

The multiple-point histogram (Deutsch, 1992) is another tool for characterizing the high-order structure in a field, whether a TI or simulated realization. Unlike the distribution of runs, a MP histogram has no meaningful units which may be used to plot it as a function. Therefore the only feasible way to compare MP histograms is to take some sort of numerical difference, such as the sum of absolute differences between MP histogram classes (Boisvert, 2007). This is a fast and convenient metric for quantifying the differences in MP structure between fully-populated grids. However, it is difficult to tell how much of the variation in MP histograms is due to the simulation and how much is naturally inherent to the underlying random variable.

## Variability of Statistics in Training Images

To test the intrinsic variation in MPS, one hundred unconditional realizations each of four different sizes were created using the `FLUVSIM` program (Deutsch and Tran, 2002). There are three facies present in the realizations and sufficient complexity that simple statistics cannot fully capture the fundamental spatial structure. These realizations will be treated as training images which could be used for simulation using MPS; Figure 1 shows one of each of the four different sizes of TIs.

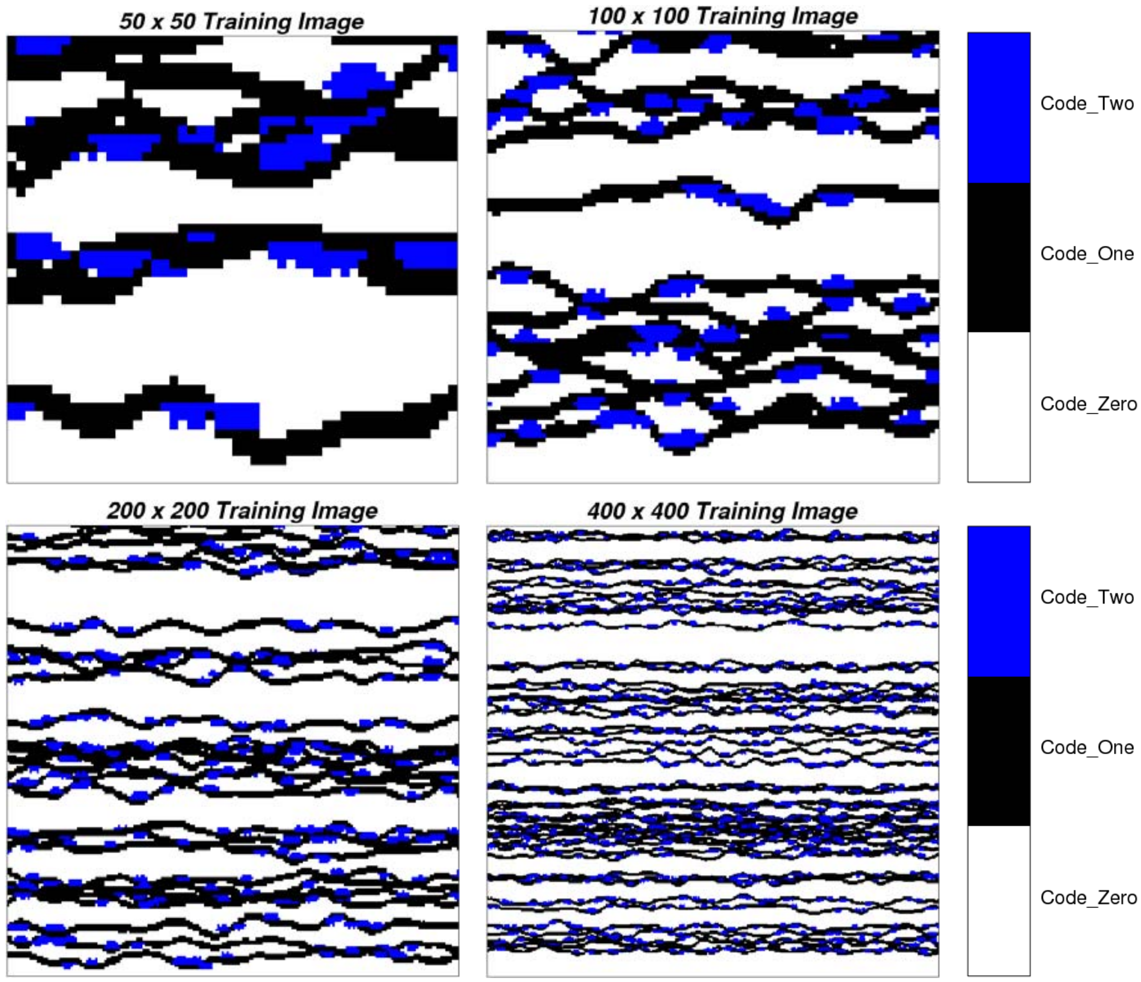


**Figure 1:** Four different sizes of training images with the same underlying spatial structure created using `FLUVSIM` (Deutsch and Tran, 2002).

Each of the four sizes of TI seen in Figure 1 have the same structure with the same scale and univariate proportions. It is easy to see that the 400x400 pixel image contains far more repetitions of the structure (channels, sinuosity, crevasse splays, etc) than the smallest 50x50 pixel TI. It would be reasonable to expect far greater statistical variability for the 100 smaller TIs than for the larger sizes. A (theoretical) infinite TI would have no variability at all, as every TI would contain an infinite number of every class of any MP histogram, and only the ratio between the class frequencies would be calculable as opposed to the numbers of occurrences.

To demonstrate the statistical variability even in similar images, consider Figure 2. This shows the indicator variograms from the smallest TIs for each of the three facies in the X and Y directions as well as the average variograms over all 100 TIs. The average variograms can be assumed representative of the "true" underlying spatial structure, as they are calculated using a very large amount of data and therefore any individual fluctuations cancel each other out. There is quite a bit of variability evident from Figure 2, even for simple two-point statistics and at distances well below the range of correlation. The smallest TIs, at only 50x50 pixels, do not capture the repeating nature of the geology and as such have significant variability.

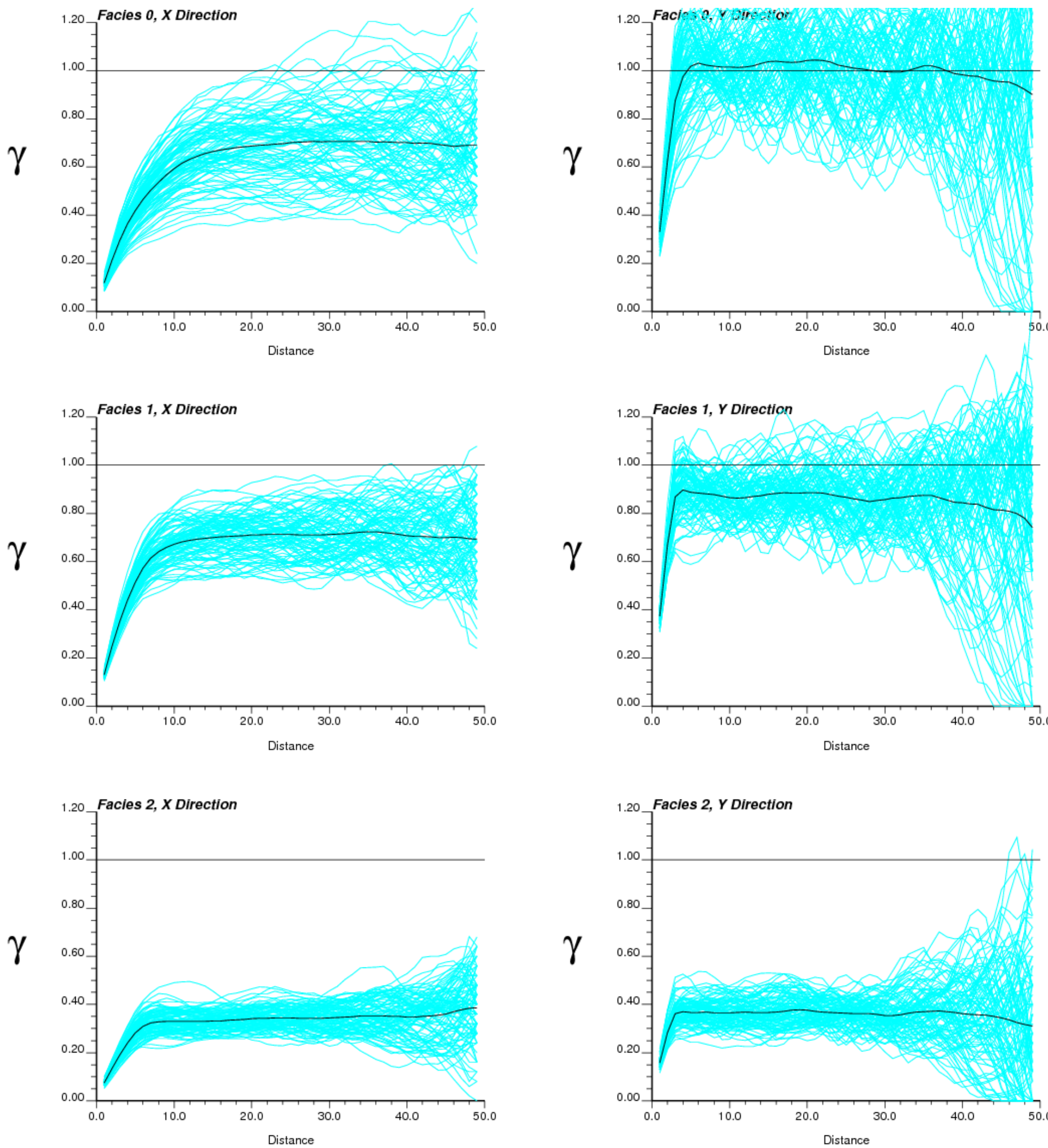


**Figure 2:** Indicator variograms of the 100 50x50 training images (blue) and the average (black).

Figure 3 shows the variograms calculated from the largest set of TIs, which are 400x400 pixels. The average variograms are not significantly different from the averages from the smallest set of TIs, which reinforces the idea that the average variograms are representative of the “true” spatial structure. However, the variograms for individual TIs display far less variability for the larger field size. Each 400x400 TI is more reasonable to use as “fully” representative of the underlying structure than the 50x50 TIs.

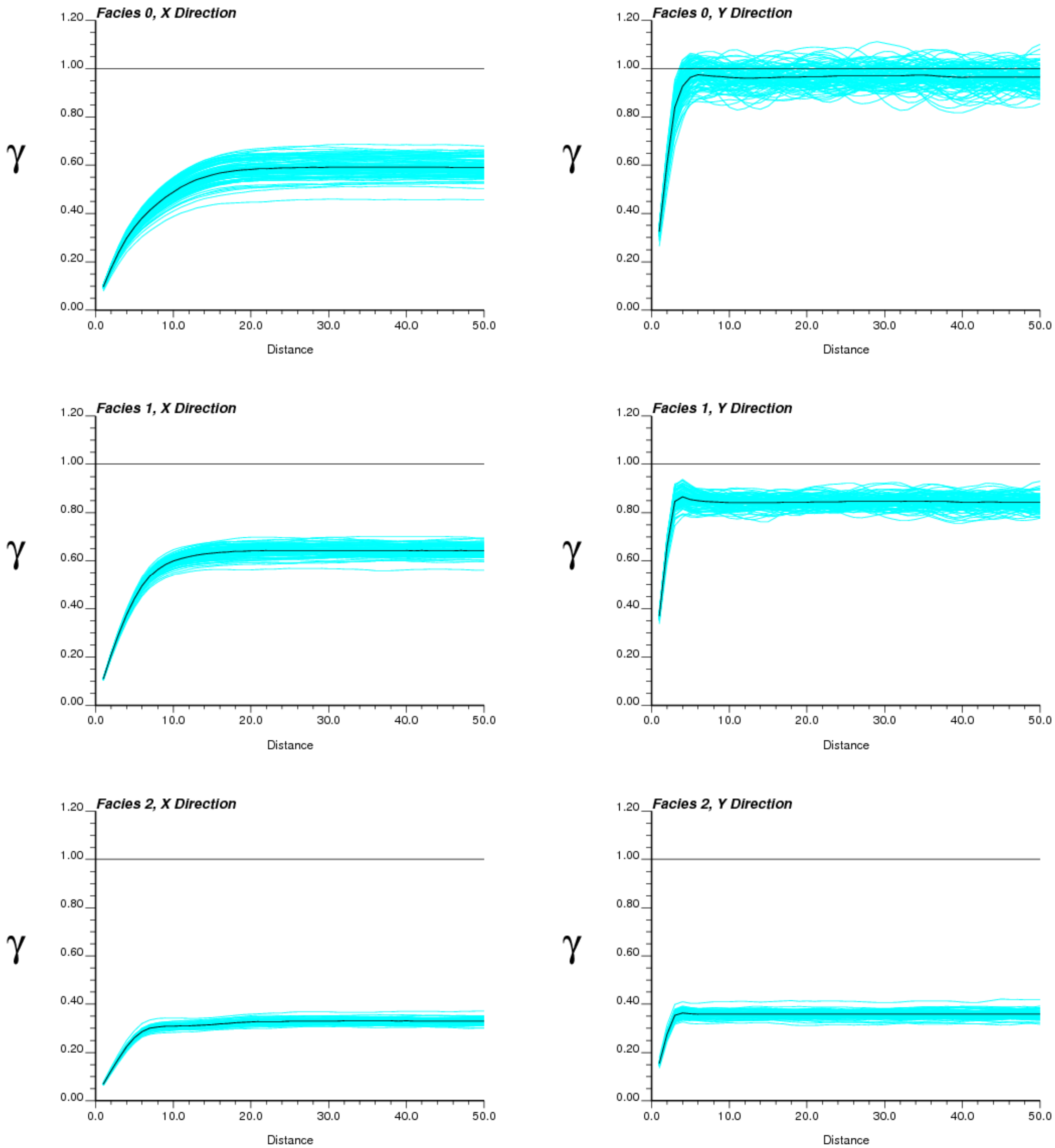


**Figure 3:** Indicator variograms of the 100 400x400 training images (blue) and the average (black).

Variograms are a traditional tool for both geostatistical simulation and evaluation of the quality of realizations. To better understand the variability inherent to MP histograms, four-point statistics were calculated for each of the four sizes of TIs shown in Figure 1. The template used was two pixels square, so there are $3^4 = 81$ different classes in the MP histograms. The average histograms over all 100 TIs of each size were calculated and used as the "true" underlying spatial structure; then, the absolute value differences from the "true" in all 81 classes were summed over all 100 TIs of the same size. This may be summed up in the equation:

$$\Delta = \sum_{j=1}^{K^N} \left| f_j - f_j^* \right| \tag{1}$$

where $\Delta$ is the total difference value for the MP histogram of the TI; $K$ is the number of facies (in this case three); $N$ is the number of points in the template (four in this example); $j$ is the index of the MP histogram classes; $f_j$ is the frequency of class $j$; and $f_j^*$ is the underlying "true" MP histogram value for class $j$.

Equation 1 gives a value which is bounded by zero (for identical histograms) and two (for mutually exclusive histograms). These delta scores were calculated for each of the four sizes; Figure 4 shows histograms of the values for all 100 TIs of each size.

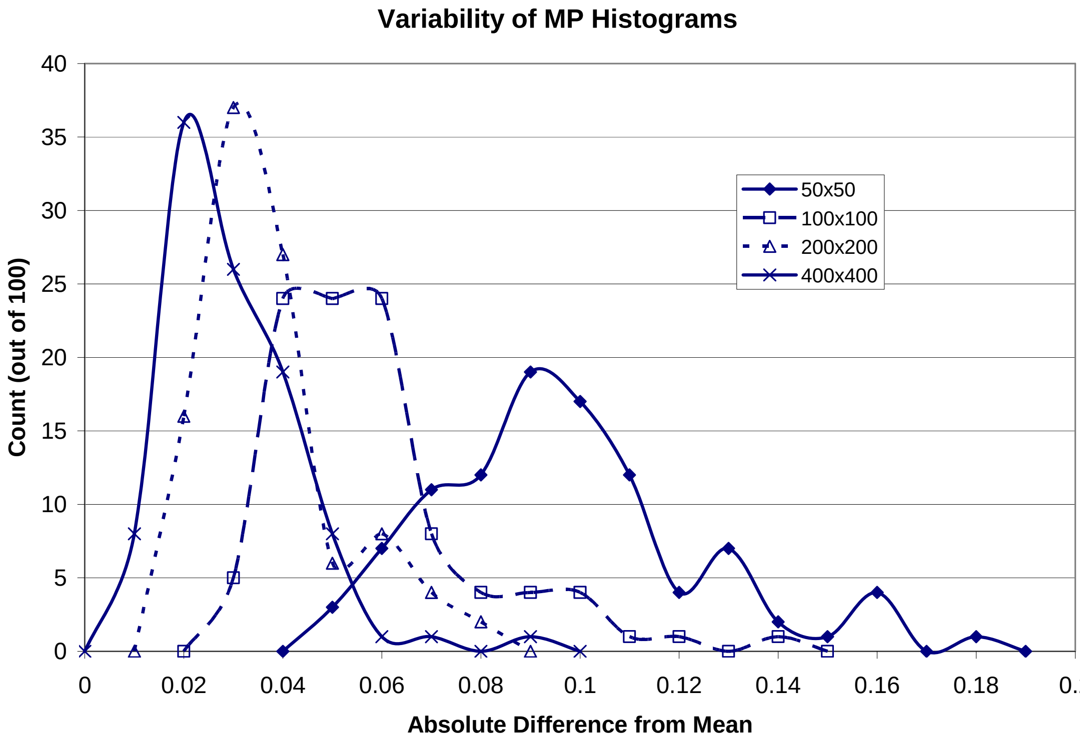


**Figure 4:** Distributions of the absolute four-point histogram differences between individual TIs and the "true" underlying spatial structure.

It may be seen that the delta values for the different sizes of TIs follow similar distributions, each with a single mode and long tails. As the TIs get larger, the average difference from the true structure decreases as does the spread of the variability.

For measuring the quality of simulations, the distributions seen in Figure 4 are likely the minimum attainable MP histogram differences between realizations and the TI. For a 100x100 realization using the TI shown in Figure 1, a difference of 0.05 from the TI would be nearly "perfect" in that this is the inherent variability seen within the TI itself. Simulation methods which begin to approach the underlying variability should produce very good realizations at the scale of the MP histogram being considered. Use of simpler statistics, such as the variogram, will still be necessary to check the long-range structure of the realizations; univariate statistics must also be checked as they may have a profound effect on the response characteristics of the realizations.

## Future Work and Conclusions

To quantify the inherent variability in MP histograms from a single TI it may be possible to divide the TI into several regions, each the size of a realization, if the TI is several times the size of the field to be simulated. This would allow characterization of the variability of the spatial structure contained within the MP histogram and provide a baseline value for the minimum achievable deviation from the true MP histogram. Any realizations produced using a MPS simulation method must still be checked using traditional comparison techniques, but direct use of MPS to quantify and rank realizations should improve the overall robustness of the method and help define the simulation parameters used.

## References


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

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. and Tran, T.T. (2002) FLUVSIM: A Program for Object-based Stochastic Modeling of Fluvial Depositional Systems. *Computers & Geosciences*, Vol. 28, No. 4, May 2002, pp 525-535.

Ortiz, J.M. (2003) *Characterization of High Order Correlation for Enhanced Indicator Simulation*. Ph.D. Thesis, University of Alberta, 255 p.

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

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