Excess Variability in Realizations of Sequential Indicator Simulation of Continuous Variables

David F. Machuca Mory, Julian M. Ortiz and Clayton V. Deutsch

Centre for Computational Geostatistics (CCG)
University of Alberta

Sequential Indicator Simulation (SIS) realizations often exhibit high and unrealistic short scale variability; this is due to the uncontrolled transitions between classes and the randomness inside each class introduced by the Monte-Carlo drawing within classes. Despite these problems, SIS has some useful properties that most of the other simulation techniques have not; this motivates further research to overcome the problems of SIS. As a first step towards the improvement of SIS, the impact of this unwarranted short scale variability in the block scale uncertainty is analyzed and compared to Sequential Gaussian Simulation results in a numerical example, obtaining a reduced block scale uncertainty for SIS results. A path for subsequent research work to improve the algorithm and its results is also delineated.

Introduction

Gaussian based simulation techniques are popular and relatively simple methods to simulate continuous variables; however they have some important limitations: they are restricted by the assumption of multigaussianity of the Random Function model, the maximum entropy property related to this model reduces the correlation of extreme values, and a single variogram model is used, locking the correlation at every threshold. These limiutations make these algorithms not suitable for real data sets where a strong spatial correlation of low or high values is present. In addition, they are not sufficiently flexible to handle mixed populations or to incorporate soft data.

Sequential Indicator Simulation (SIS) for continuous variables can handle variables with any type of distribution and that do not fulfill satisfactorily the multigaussian assumption even after normal scores transformation. It offers a greater flexibility by using different variograms for different thresholds, allowing the spatial modeling of continuous variables with complex patterns of spatial distribution for low, median and high values. Besides this characteristic, SIS has other properties that could make it suitable for its use with continuous variables: it allows a straightforward integration of mixed data types, as well as secondary soft data, (Deutsch and Journel, 1997) and it is also widely used and not difficult to understand.

Nevertheless, the SIS algorithm is affected by several practical and theoretical difficulties and disadvantages (Chilès and Delfiner, 1999; Christakos, 2000; Emery and Ortiz, 2004), the most important ones among them, stated in the context of this work, are two:

- The uncontrolled transitions between classes that are product mainly, but not only, of the lack of information about the inter-class spatial cross-correlation results in the unrealistic overlapping of low and high grade patches of SIS realizations.
- The random and uncontrolled drawing of simulated values on the conditional cumulative distribution function (ccdf), which has been constructed by indicator Kriging, results in pure randomness within class boundaries.

Together, both disadvantageous characteristics produce the excess variability in SIS realizations, that is, the spatial persistence of the unstructured short scale intermixing of low and high values all over the area under study, which translates in a reduced variability in the block scale due to the averaging of low and high values that lead to smoothed results. The first drawback can be partially mitigated by introducing the interclass correlation using the complete set of indicator direct and cross-variograms for constructing the ccdf by indicator co-kriging. However, this means an increased effort in modeling the entire variogram matrix for several thresholds using the Linear Model of Corregionalization which not always fits satisfactorily the entire set of variograms, particularly the cross-variograms between extreme thresholds. This approach can also increase the order relation problems in constructing the ccdf. To overcome the second problem the idea is to introduce spatial ordering within classes by considering the spatial correlation among the simulated values inside each class and between the simulated values and the class boundaries.

The objective of this work is to analyze and evaluate numerically the impact of SIS peculiar short scale variability in the block scale uncertainty assessment using a real data set and assessing the sensitivity of this variability to the variogram uncertainty.

Statistical and Geostatistical Data Analysis

The data used in this study is a subset of a drillhole exploration campaign of a Chilean copper deposit. This data subset comprises all the samples corresponding to a single bench of 12m height, located between elevations 3928 and 3940, and belonging to the Tourmaline Breccias rock type. A sample location map is presented in Figure 1.

A total of 180 samples are included in this data set, with copper grade values ranging between 0.21% and 6.89%. Declustering is performed to obtain representative statistics. The declustered mean is 1.26 %Cu, the median is 1.10 %Cu and the variance,0.763 $(\%Cu)^2$; Figure 2 shows the histogram of declustered values. Nine thresholds are chosen to discretize the conditional distributions (Table 1).

Table 1: Thresholds for Indicator Variograms

CDF	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Cu%	0.53	0.66	0.784	0.93	1.1	1.24	1.38	1.62	1.949

Two sets of normal score transformed data are generated. The first one is created without considering the declustering weights to be used in the indicator variography. The second set is generated considering the declustering weights and it is used for the Gaussian simulation. This procedure is necessary because the normal scores transformed data not always have variance equal to one when declustered weights have been used in the transformation.

Spatial Correlation Analysis

From the variogram map of normal scores transformed values, the direction of major continuity in the horizontal plane can be defined with an azimuth between 120º and 135º. Calculation of the experimental variograms confirms the orientation of the major axis.

In order to account for the variogram uncertainty, three models are fitted for the continuous variogram and each indicator variogram; these models correspond to low, medium and maximum continuity scenarios. The variogram model for maximum spatial continuity is defined as the model with the lowest nugget effect and the largest range that can be reasonably fitted to the experimental continuous and indicator variograms. The minimum spatial continuity model has the largest nugget effect and shortest range that can be fitted to the experimental variograms. And the medium spatial continuity model is the best fit model, the one that better describes the experimental variogram. The ratios of ranges and nugget effect of long continuity and short continuity models with respect to the medium continuity model were kept constant when possible for the all indicator variograms. An exponential variogram model was chosen to fit the continuous and all the indicator variograms, the parameters of the resultant fitted models are presented in Table 2 (practical range is shown).

Table 4: Ranges for the sensitivity analysis

Spatial Continuity:			**Minimum**	**Middle**	**Maximum**
Variogram	Axis	**Az**	**Variogram ranges & N.E.**		
NS-cont	**Major**	120	260	280	320
	Minor	30	130	140	200
	Nugget		0.35	0.25	0.15
I10	**Major**	120	153	170	196
	Minor	30	18	20	28
	Nugget		0.35	0.25	0.15
I20	**Major**	120	198	220	253
	Minor	30	63	70	98
	Nugget		0.14	0.1	0.06
I30	**Major**	120	180	200	230
	Minor	30	126	140	196
	Nugget		0.42	0.3	0.18
I40	**Major**	140	162	180	207
	Minor	60	99	110	154
	Nugget		0.21	0.15	0.09
I50	**Major**	120	162	180	207
	Minor	30	99	110	154
	Nugget		0.14	0.1	0.06
I60	**Major**	130	144	160	184
	Minor	40	99	110	154
	Nugget		0.28	0.2	0.12
I70	**Major**	130	117	130	150
	Minor	40	36	40	56
	Nugget		0.28	0.2	0.12
I80	**Major**	120	90	100	115
	Minor	30	36	40	56
	Nugget		0.56	0.4	0.24
I90	**Major**	120	90	100	115
	Minor	30	36	40	56
	Nugget		0.28	0.2	0.12

Uncertainty Analysis

Sequential Gaussian Simulation and Sequential Indicator Simulation are performed over a 2m x 2m x 12m cell size grid covering 300m in the east-west direction, 600m in the north-west direction and coincides with the bench 3928. A search ellipse defined by a major axis of 300 m with an azimuth of 120º and a minor axis of 150m was used in both SGS and SIS, with a maximum number of original samples of 10, and maximum number of simulated nodes of 6. In order to improve the histogram reproduction in SIS, 200 tabulated quantile values were used to control the cumulative distribution interpolation between the nine thresholds.

Three groups of 100 SGS and three groups of 100 SIS realizations are generated using the corresponding minimum, middle and maximum continuity variogram models. Realizations are clipped by the boundary of the Tourmaline Breccias rock type. The histograms of all the simulated values in every realization are shown in Figures 3 and 4, next to sample realization maps. These point support realizations are averaged to a SMU size of 12m x 12m x 12m. The histograms and sample maps of the averaged values are presented in Figures 5 and 6. While the histogram reproduction is comparable for both simulation techniques, the patchiness of SIS realizations is clearly evident in Figure 4 where abrupt and disordered transitions between classes can be observed. However, when averaged to 12m x 12m blocks, SGS realizations preserve the distinction of high and low grade zones, but SIS realizations tend to smooth this distribution making this delineation very fuzzy and yielding lower means and variances than SGS.

Sensitivity of SGS and SIS to the Spatial Correlation

In order to assess and compare the sensitivity of both SGS and SIS techniques to the uncertainty in the parameters of spatial continuity, a group of simple transfer functions is defined:

- Recoverable metal content;
- Recoverable proportion of total tonnage; and
- Average grade above cut-off.

The cut-off is fixed as 1.5% Cu, which is compatible with an open pit copper mine (this is very high, but used for illustration purposes only). These transfer functions are calculated for all realizations at the SMU scale. The standard deviation of the transfer function is used as a measure of uncertainty. The comparative results of the variogram continuity impact on transfer functions are presented as histograms in Figures 7, 8, and 9.

In figure 9, it can be observed that the tonnage proportion above cut-off yields to a lower dispersion in the histograms corresponding to SIS averaged realizations, with a variance around 2.1 for all variograms models, which is below the variance of the same transfer function applied to SGS averaged realizations. For SIS, the direct relationship between the mean of the recoverable average copper grade above cut-off and the spatial continuity of the variogram model is clear, while SGS does not show any particular relation between both. This relationship can be explained by the decreasing influence of the relatively few high values as the nugget effect increases and the variogram range decreases, lowering in this way the proportion of simulated values over the upper thresholds and increasing the smoothing of SIS in the SMU scale averages.

Interpretation of Results

As seen in Figures 3 and 4, the declustered sample histogram is slightly better reproduced by the Sequential Gaussian Simulation than by the Sequential Indicator Simulation even when quantile interpolation between thresholds is used. The realization plots of SGS look better structured and more geologically appealing, while realizations generated with SIS show the characteristic patchy appearance, with the presence of adjacent low and high grade zones, and randomness inside classes.

When averaged to a larger SMU size, SGS realizations clearly preserve the structure of high and low grade areas in concordance with the original sample values, but this differentiation is less apparent in the SIS realizations, becoming almost indistinguishable for the indicator simulations when a low continuity variogram model is used. The smoothing in the SMU scale is caused by the increasing of simulated values heterogeneity all over the study area as the nugget effect increases and the range decreases. This smoothing effect of decreasing variogram continuity is reflected as a smaller variance at the SMU scale in SIS realizations, but not in SGS realizations.

Overall uncertainty has been quantified by the standard deviation of the response: the average grade above cut-off, the proportion of recoverable tonnage and the recoverable metal content. The response of overall uncertainty to changes in the variogram continuity is summarized in Figure 10:

- The uncertainty in the recoverable tonnage (Figure 10a) is larger for SGS results but decreases as the spatial correlation in the variogram model increases. SIS results present a lower recoverable tonnage uncertainty and change only slightly as the modeled continuity of grades changes.
- For SIS results, the standard deviation of the average cooper grade above cut-off (figure 10b) increases with the continuity of the variogram model. This positive relationship among the average ore grade uncertainty and the spatial continuity can be explained by the increased smoothing of the SMU grades related to the loss of structure in SIS realizations when the spatial continuity is reduced. Thus, not only the variance of all simulated values reflect this but also the standard deviation of the average ore grade transfer function, which can even be comparable to the standard deviation corresponding to SGS results when a high continuity variogram model is used. This function applied to SGS results does not show a clear tendency related to the level of spatial continuity.
- The recoverable metal content (Figure 10c), being a product of the previous transfer function, reflects the same features, this is, a relatively low standard deviation of the recoverable tonnage, and positive relationship between the spatial continuity and the uncertainty in the average grade above the cut-off when these transfer functions are applied to SIS results.

Conclusions and Future Work

As expected, the increased and unrealistic short scale variability in the SIS realizations not only produces a misleading low variability at the SMU scale; it can also be observed in the SIS realizations maps and in the sensitivity analysis for different transfer functions when imposing different continuity in the variogram models. It is important to notice that the uncertainty in SGS results is generally larger than when using SIS, for all the transfer functions considered. The homogeneous mixture of low and high values of SIS results yields to an “artificially” reduced uncertainty at block support or after a transfer function is applied.

The loss of spatial structure in the short scale, translates in the increasing of smoothness at the SMU scale, lowering the uncertainty on the average ore grade for SIS results. Intuitively, one would expect the opposite: The higher the spatial continuity, the lower the uncertainty in the transfer function. SGS results suit better this expected behaviour, uncertainty in the recoverable tonnage and recoverable metal content decreases slightly as variogram continuity increases.

In despite of the unwanted consequences of uncontrolled transitions between different classes and the randomness inside classes, SIS has some desirable properties that Gaussian simulation techniques do not have. One of these properties observed and utilized in this work is its flexibility to handle different spatial continuity models by defining a different variograms model for each threshold. The unique properties of SIS are a strong motivation to research techniques and methods to overcome the severe limitations of this simulation technique.

Future work in this direction may be focused on two objectives: (1) Accounting for cross correlation between classes in order improve the control in the transitions between them. (2) Reducing the randomness inside each class by introducing spatial correlation. In order to accomplish the first objective the main idea is the implementation of the full indicator Cokriging in the SIS algorithm. This requires the simultaneous modeling of $K(K+1)/2$ indicator direct and cross-variograms for K thresholds using the Linear Model of Coregionalization (LMC). Nevertheless, this model often gives a coarse approximation to experimental indicator cross-variograms of thresholds widely separated, particularly when the continuous variable exhibits a low nugget effect. In response to this inconvenient the adjacent thresholds indicator approach (Goovaerts, 1994) is suggested; this is, to use just the cross variograms between the closest upper and lower classes to the corresponding threshold and the direct variograms and discarding the indicator cross variograms of farther thresholds to permit a good LMC fitting. To accomplish the second objective it has been suggested to introduce the distance of the node to the closest class edge as a secondary correlated variable (Ortiz, Neufeld, and Deutsch, 2005). Further work is needed to determine the best approach to correct the randomness inside classes.

References

Babak, O. and Deutsch, C.V., 2005. A *Short Note on How Response Uncertainty is Affected by Local Data and Parametric Uncertainty*. Centre for Computational Geostatistics, Report 7.

Chilès, J.P. and Delfiner, P. *Geostatistics, Modeling Spatial Uncertainty*. Wiley Series in Probability and Statistics. New York, 1999.

Christakos, G. *Modern Spatiotemporal Geostatistics*, Oxford University Press, New York, 2000.

Deutsch, C.V., 2005. *A Sequential Indicator Simulation Program for Categorical Variables with Point Data: BlockSIS.* Centre for Computational Geostatistics, Report 7.

Deutsch, C.V., Ortiz, J.M., and Neufeld, C., *New Software for Fitting Indicator Covariances and for Indicator Kriging/Simulation*. Centre for Computational Geostatistics, Report 7.

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

Emery, X. and Ortiz, J. M., *Shortcomings of multiple indicator kriging for assessing local distributions,* Applied Earth Sciences (Trans. Inst. Min. Metall. B), Vol. 113, B249-B259, December 2004.

Goovaerts, P. *Comparative Performance of Indicator Algorithms for Modeling Conditional Probability Distribution Functions*, Mathematical Geology, Vol. 26. No. 3, 1994.

Ortiz, J.M., Neufeld, C. and Deutsch, C.V., A Short Note on Within Class Modeling in Indicator Simulation. Centre for Computational Geostatistics, Report 7.

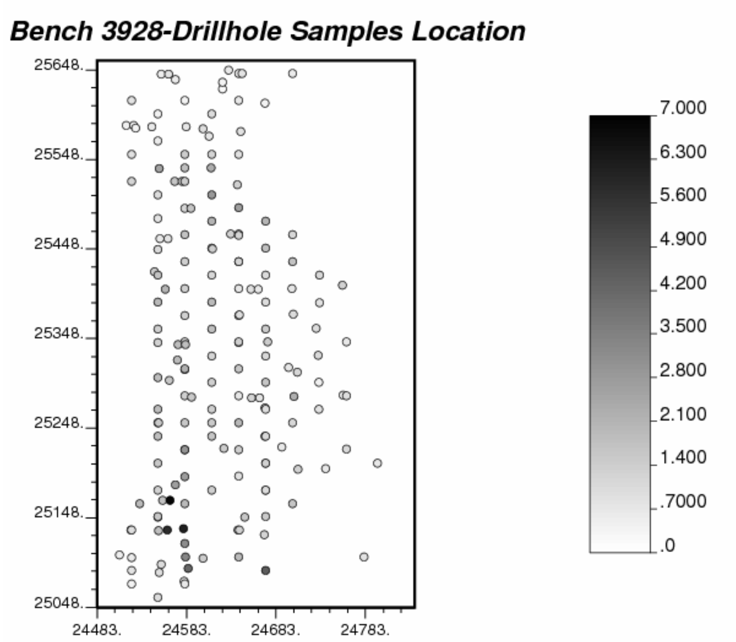


Figure 1: Drillhole sample locations in bench 3928

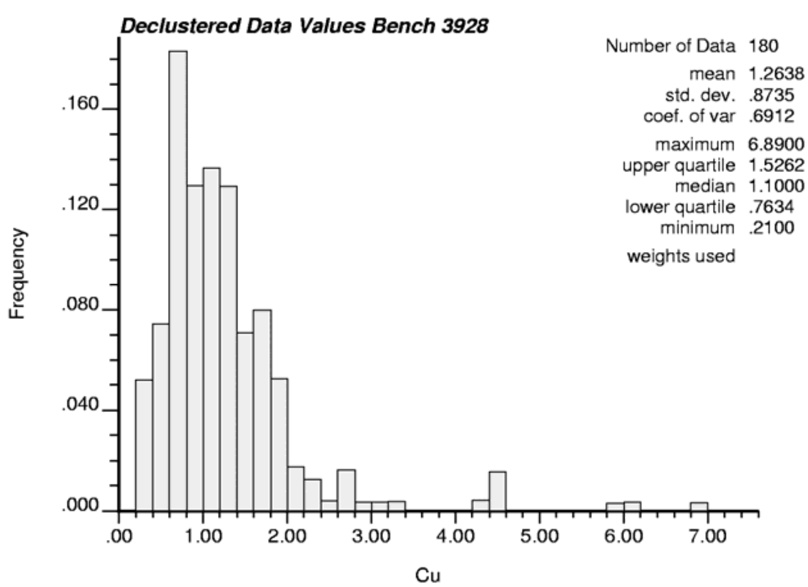


Figure 2: Histogram of declustered grades

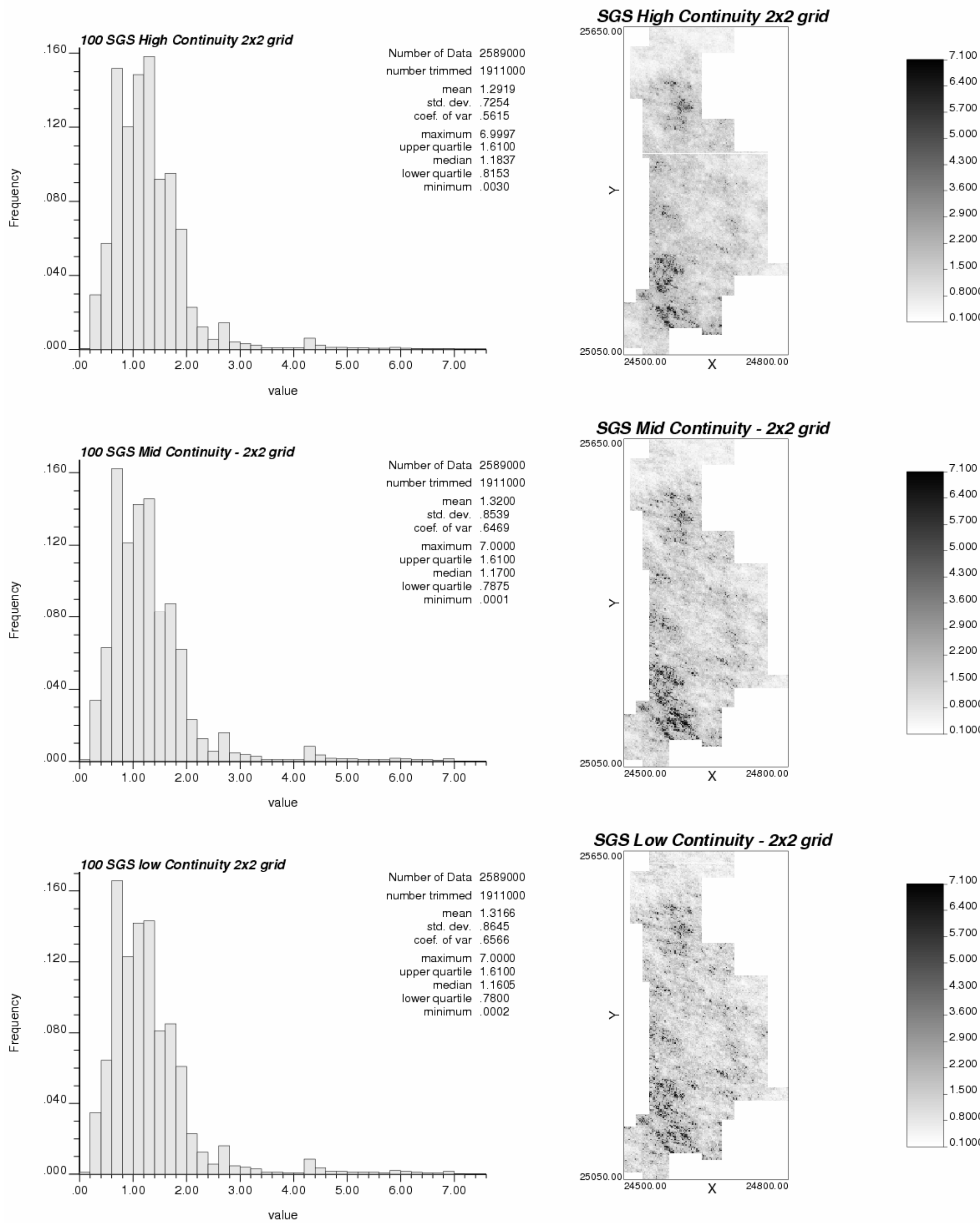


Figure 3. Left: Histograms of all simulated values by SGS with high (top), middle and low (bottom) spatial continuity variogram models. Right: sample realizations.

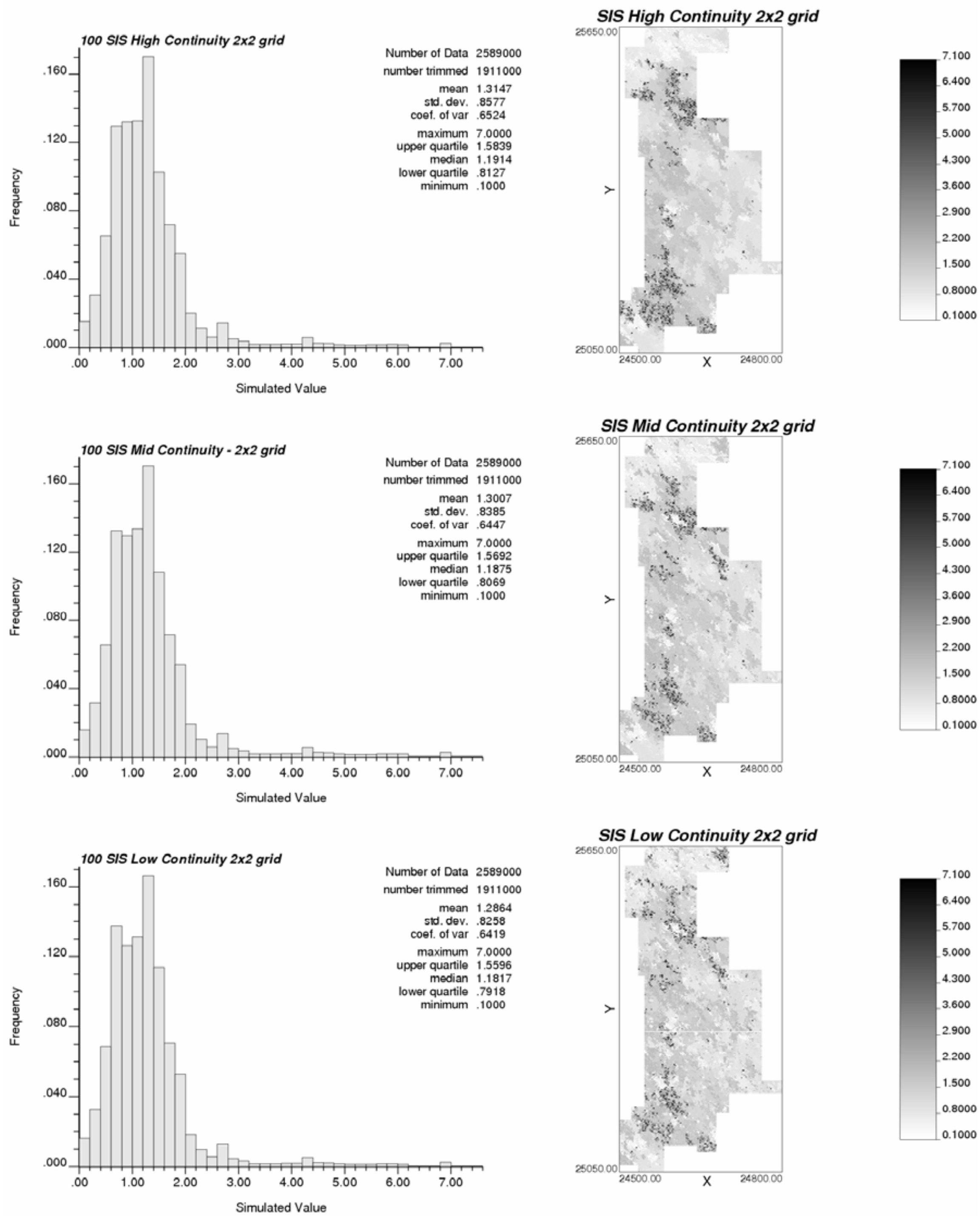


Figure 4. Left: Histograms of all short scale simulated values by SGS with high (top), middle and low (bottom) spatial continuity variogram models. Right: sample realizations.

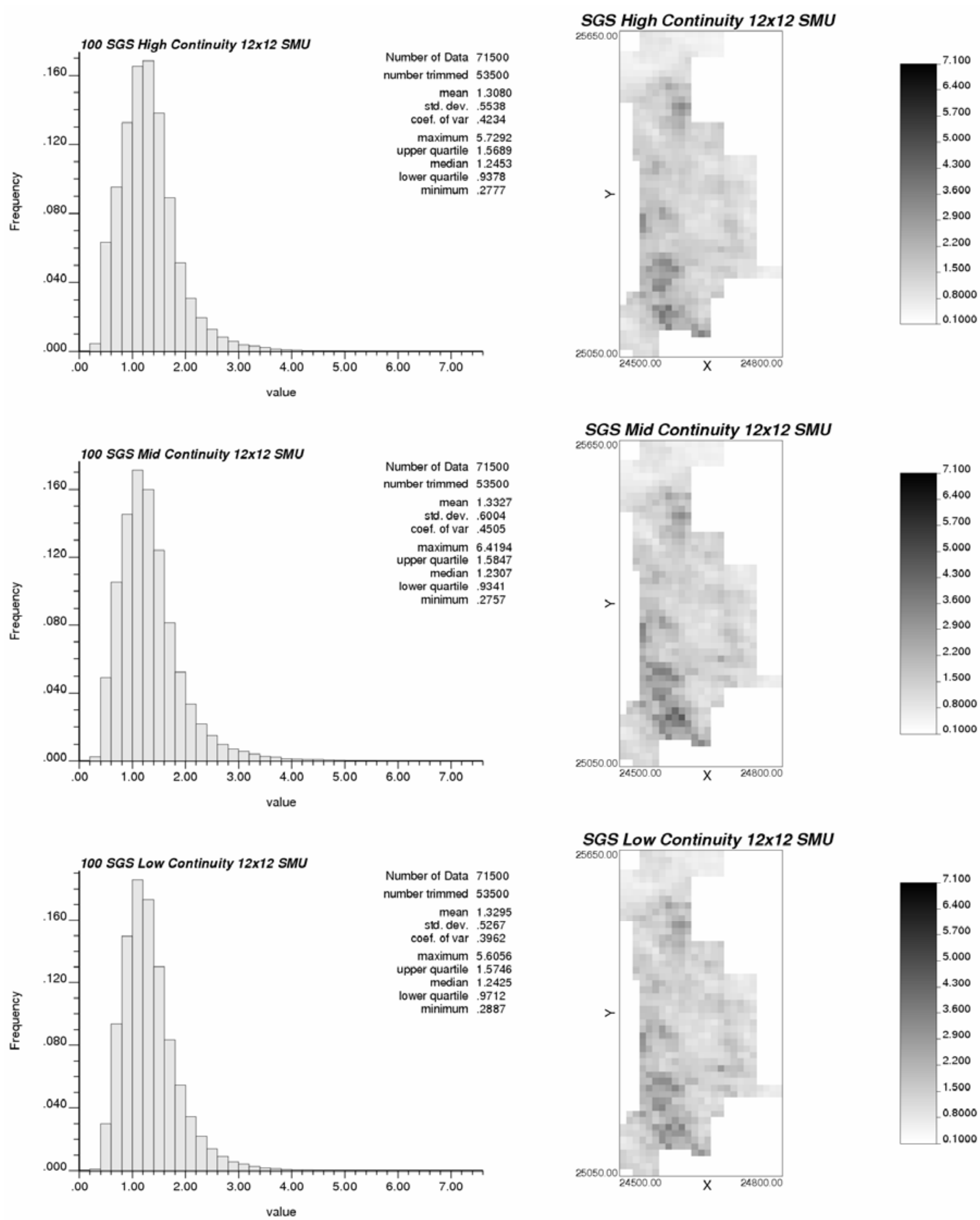


Figure 5. Left: Histograms of SGS results averaged in a 12m x 12m SMU scale corresponding to high (top), middle and low (bottom) spatial continuity variogram models. Right: Sample Realizations.

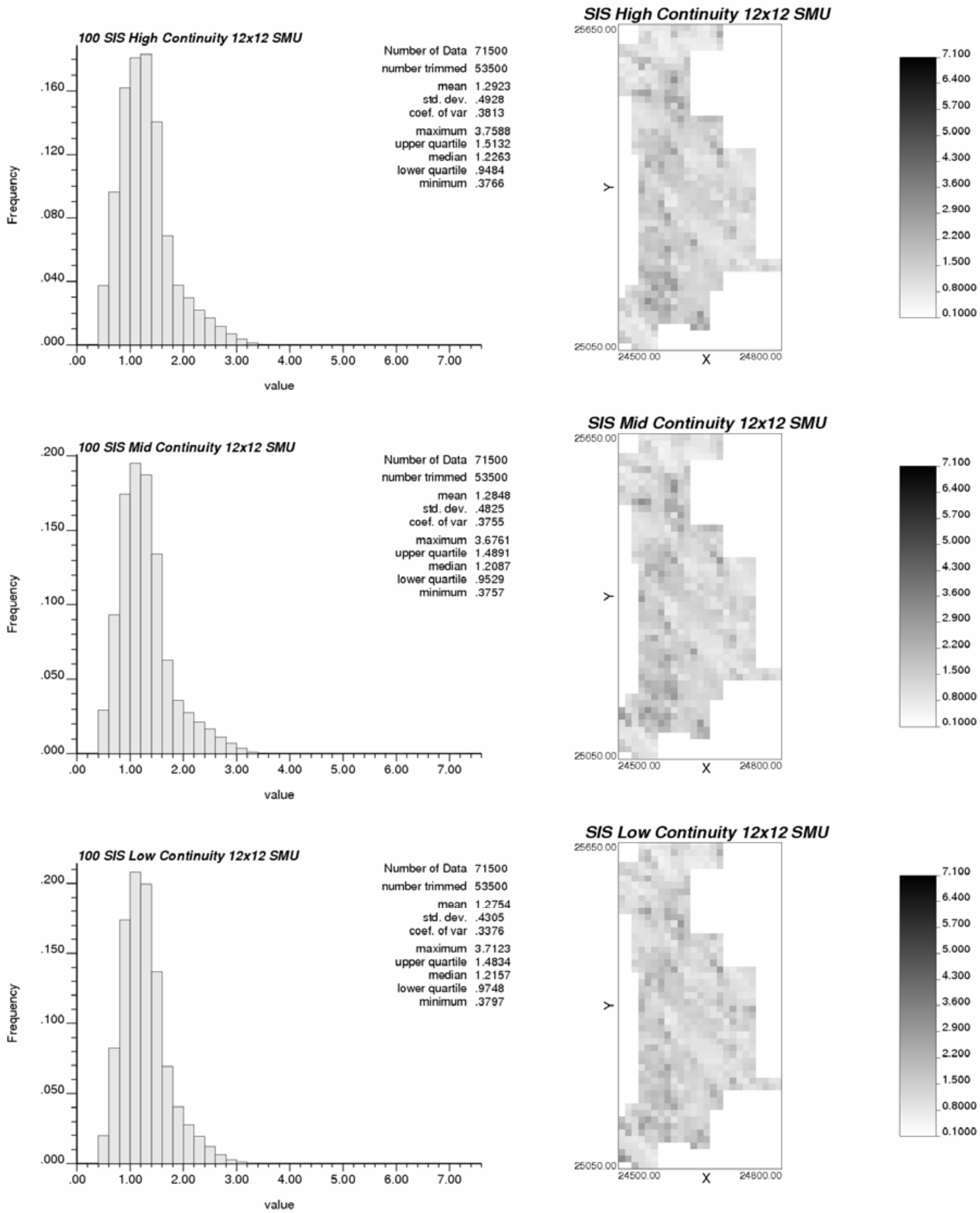


Figure 6. Left: Histograms of SIS results averaged in a 12m x 12m SMU scale corresponding to high (top), middle and low (bottom) spatial continuity variogram models. Right: Sample Realizations.

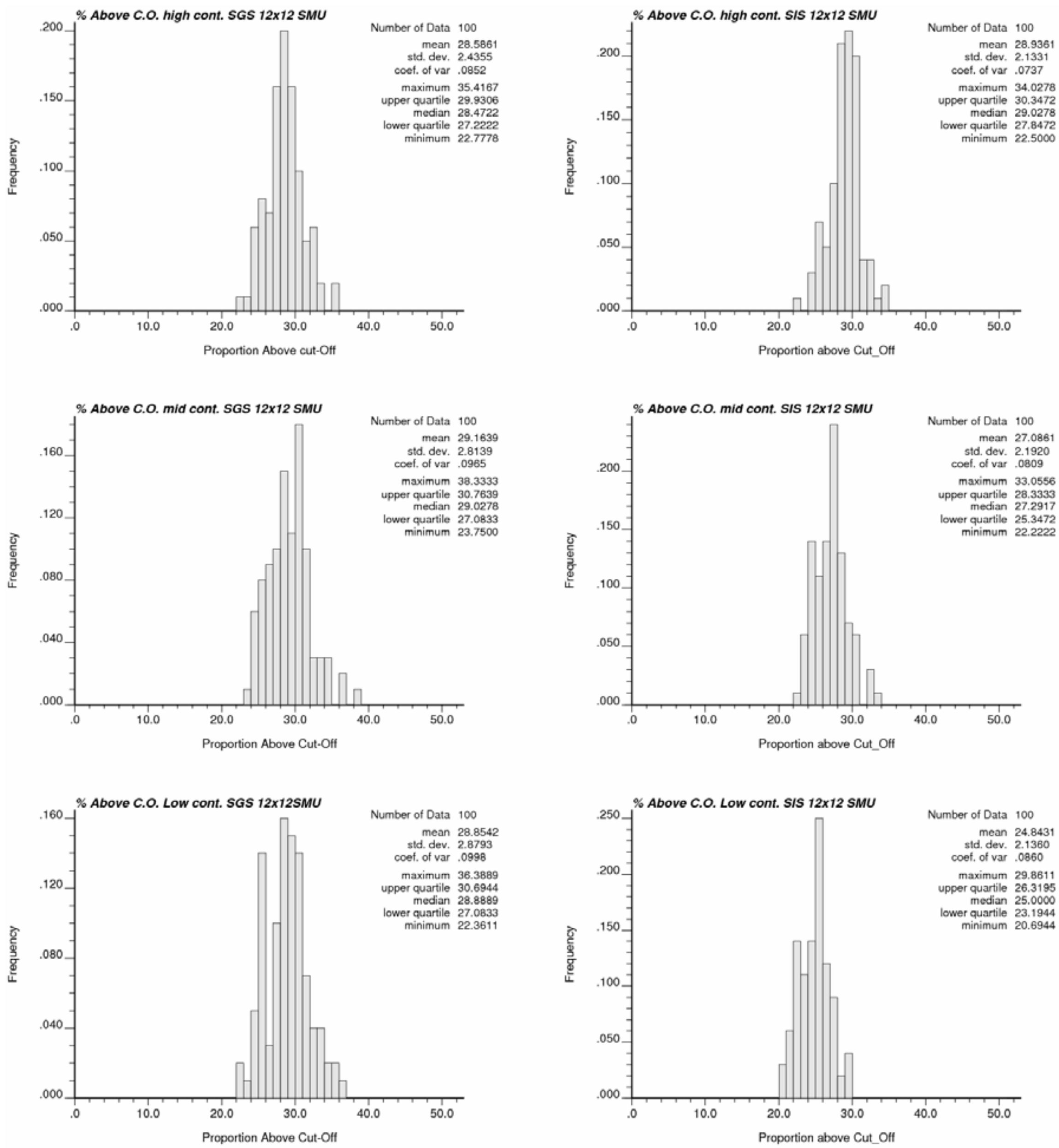


Figure 7: Histograms of the proportion of recoverable Tonnage for 100 SGS (left) and 100 SIS (right) realization using different spatial continuity models.

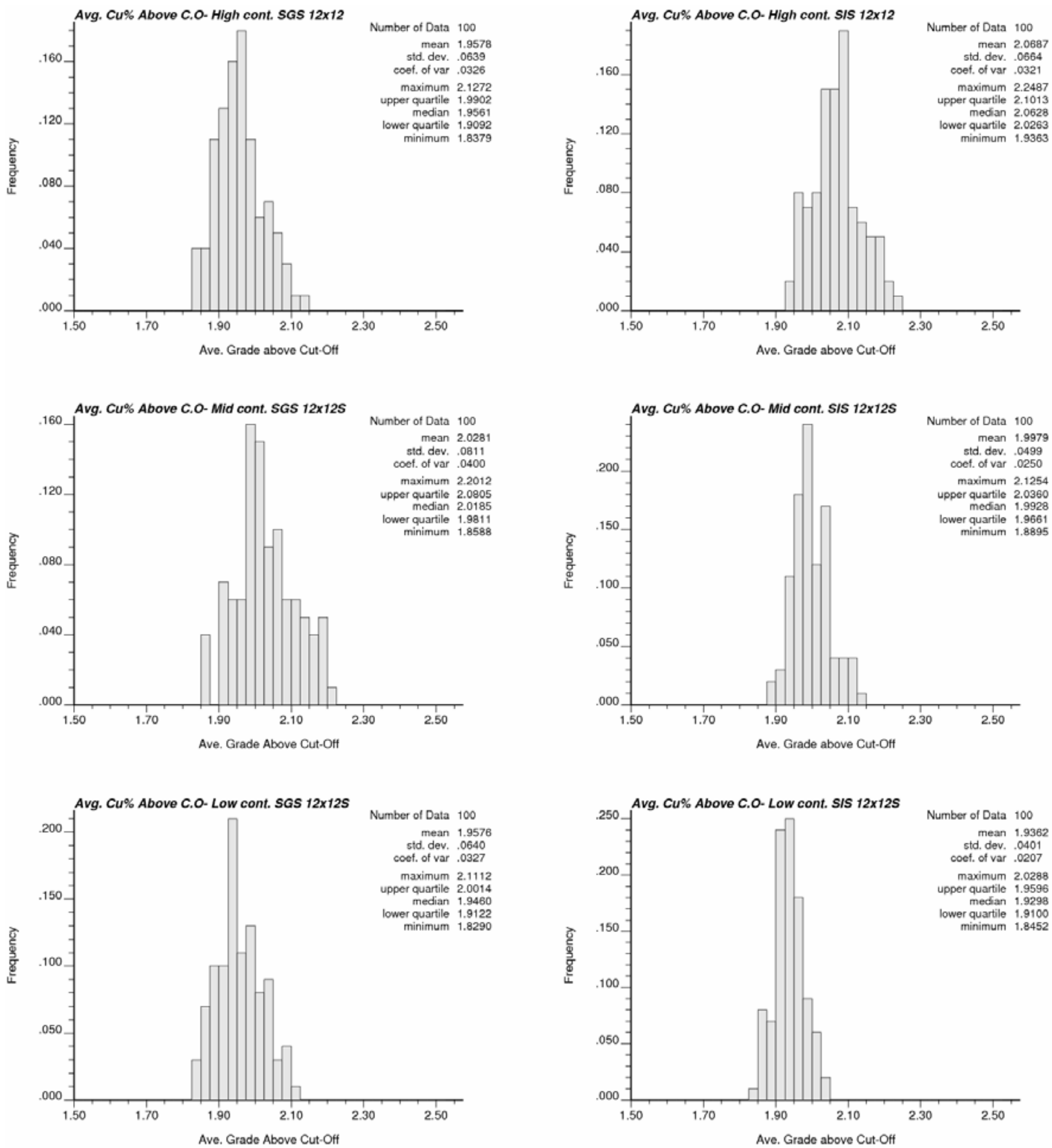


Figure 8: Histograms of the average grade above cut-off for 100 SGS (left) and 100 SIS (right) realization using different spatial continuity models.

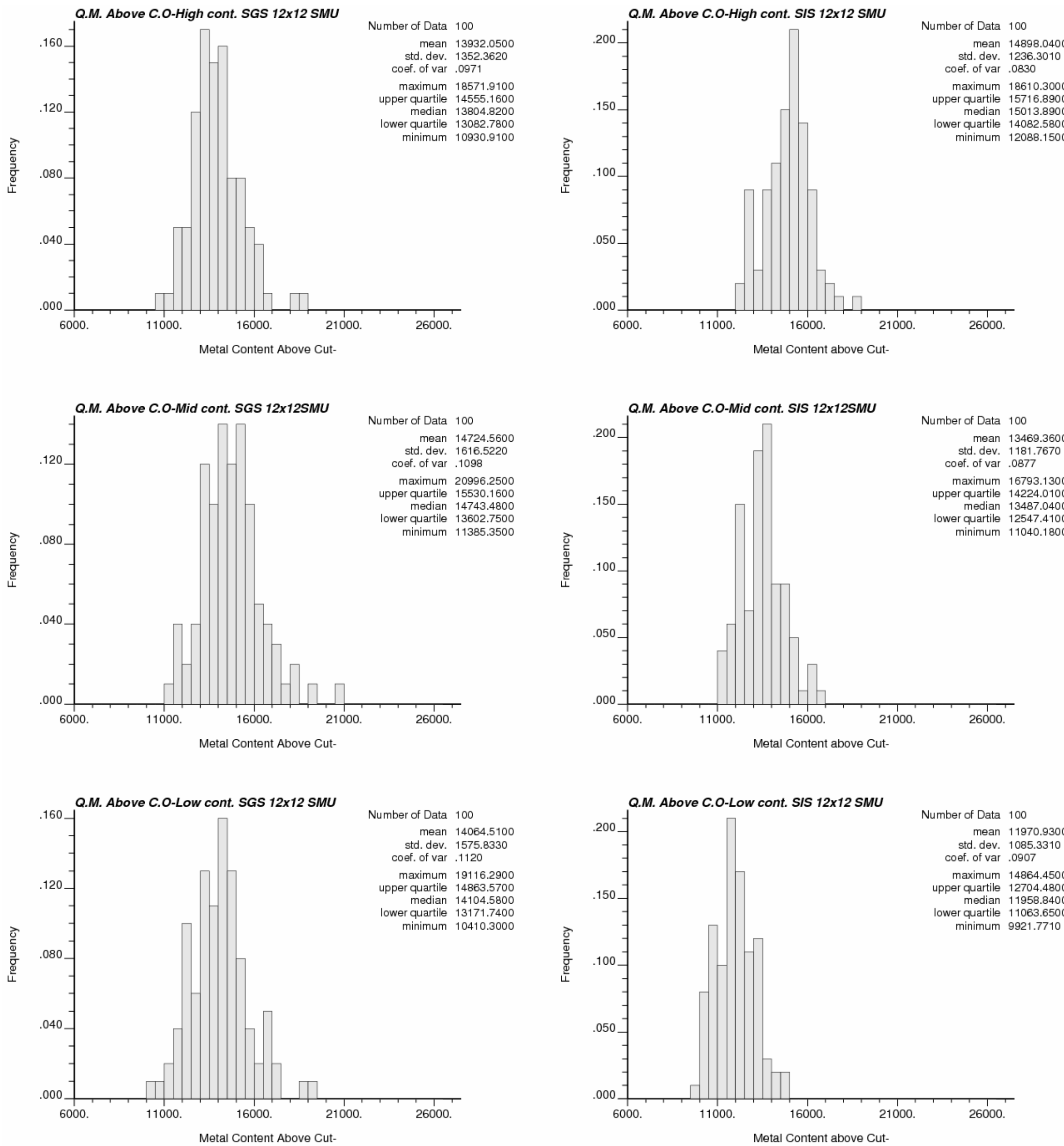
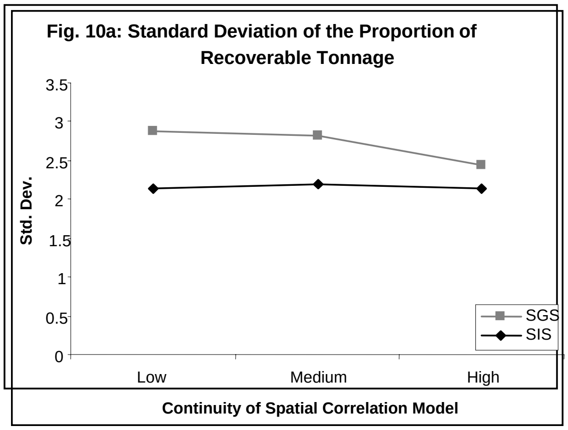
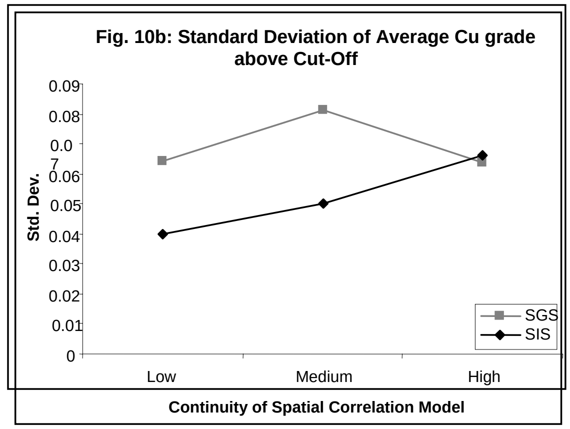
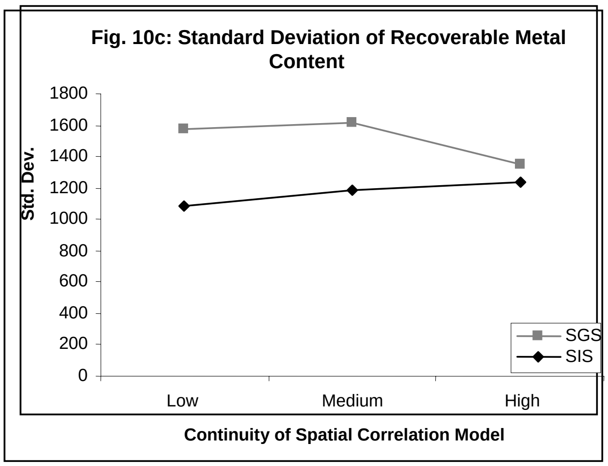


Figure 9: Histograms of the average grade above cut-off for 100 SGS (left) and 100 SIS (right) realization using different spatial continuity models.







Figures 10a to 10c: Comparison of the uncertainty measure (standard deviation) of different transfer functions applied to SGS and SIS results in the SMU scale.