

# Recovering Spatial Correlation of Multivariate Block LU Realizations

John G. Manchuk and Clayton V. Deutsch

*A block LU methodology to generate realizations of random variables at a block scale, such as a selective mining unit or panel scale, was developed at CCG; however, the results are strictly local. Results cannot be averaged to a larger block scale or visualized. Doing so is incorrect because the spatial correlation between blocks or panels is not preserved. This paper describes a workflow to reorganize realizations so that the correct spatial correlation is recovered. An advantage of the process is the value for any block and realization does not change in the process; therefore estimates of recoverable reserves are unchanged. Two tools are developed for this research: a program to upscale the LMC required to generate the initial block scale realizations and a program that reorganizes the realizations so that spatial correlation is recovered. Upscaling the variogram is necessary to generate unconditional realizations with the correct correlation that are then used to control the reorganization step. Once reorganized, it is possible to visualize realizations and average them to other scales.*

## 1. Introduction

Predicting local recoverable reserves is a critical step in mining geostatistics. An approach to assess local reserves for a single variable was developed by Ortiz and Deutsch (2007). The program, `blusim`, was designed to predict reserves for blocks such as panels without having to store and process multiple realizations. Panels are discretized into selective mining units (SMU) and the LU decomposition approach to simulation is used to generate realizations that are averaged and discarded. Often there are several variables involved, for example multiple ore grades or oxide concentrations and the problem is multivariate. In the same year (Boisvert and Deutsch, 2007; Boisvert et al, 2008), the approach was enhanced to consider multiple variables using LU decomposition and a linear model of coregionalization (LMC).

A drawback of this approach is that the reserves are local. It is only possible to map the expected value of a set of panels, which gives results similar to mapping kriged estimates. It is possible to look at individual realizations of a single panel independently of all other panels; however, a map of a realization of all panels together would be meaningless and is not advocated. Creating such a map and drawing any conclusions from it is poor practice (Boisvert and Deutsch, 2007).

In this work, a method is developed to reorganize the set of individual realizations of all panels so that the correlation at the panel scale is reproduced. The result is a set of panel scale realizations that can be mapped with the correct spatial structure. To achieve this, a tool to upscale the variogram is required. Generating panel scale realizations uses the variogram from conditioning data that may be sample (core, grab, etc) scale or SMU scale. However, reorganized realizations should reflect the variogram at the panel scale. This paper first goes through a brief review of `blusim` and then introduces the methodology and program to upscale a linear model of coregionalization. Secondly, the reorganization approach is developed. A case study is presented to demonstrate the usefulness of the reorganization approach.

## 2. Background

Panel scale realizations generated from `blusim` consider each panel in a grid of panels independently. Each panel is discretized into a set of smaller scale blocks (SMU's for example); however, this discretization is never observed by the user. For each panel, LU simulation is used to quickly generate a set of realizations at the SMU scale conditioned to any available data. The SMU scale realizations are averaged to the panel scale and discarded, resulting in only a mean and variance for each panel, or other summary statistics such as the grade above a cutoff.

For univariate cases, the covariance matrix for LU simulation is as large as the number of conditioning data (called data from here on) plus the number of SMU scale blocks (called nodes from here on) within a panel. For multivariate data, this size is multiplied by the number of variables involved; therefore, it can become a rather large system. A nice quality of the LU simulation approach is the data to data portion of the covariance matrix and the node to node portion only need to be computed once. Only the data to node portions must be updated for new panels. Another beneficial quality is that realizations are generated by an upper triangular vector product, which is a fast operation for the typical number of data available within a panel. This makes it possible to generate hundreds of realizations for a panel in seconds with a computer.

One option in the `blusim` parameters is to output the panel scale realizations. The resulting file contains all realizations used to evaluate the statistics of the grid of panels. Unfortunately, the realizations cannot be mapped as the spatial variation at the panel scale is not honored. This is especially true when sample data are sparse with a spacing beyond the range of the variogram. However, it is possible to reorganize the output realizations from `blusim` to recover the true spatial continuity. This first requires upscaling the LMC used to generate the realizations to obtain the spatial structure to reproduce.

### 3. LMC Upscaling

To reorganize panel scale realizations so that the panel scale LMC is reproduced, the point scale or SMU scale LMC input to `blusim` must be upscaled. A discussion on variogram scaling laws is found in Frykman and Deutsch (1999) and in Journel and Huijbregts (1978). Consider  $V$  to be the scale of the panels, then the averaged variogram can be derived by geostatistical scaling laws. The experimental variogram of core samples is calculated and modeled and the result is scaled to agree with  $V$ . The nugget effect,  $C_0$ , range,  $a$ , and sill,  $C_i, i = 1, \dots, n$  with  $n$  the number of nested structures, are scaled according to Equations (1), (2), and (3) (Kupfersberger et al, 1998). In Equation (2),  $|V|$  and  $|v|$  describe the extent of the volumes and are direction dependent. These results are based on variables that average arithmetically.

$$C_0^v = C_0^v \cdot \frac{|v|}{|V|} \quad (1)$$

$$a^v = a^v + (|V| - |v|) \quad (2)$$

$$C_i^v = C_i^v \frac{1 - \bar{\gamma}(V, V)}{1 - \bar{\gamma}(v, v)} \quad (3)$$

A numerical approach to acquire the variogram for the panel scale is developed to work in conjunction with an automatic variogram fitting tool called `varfit_lmc` (Jewbali, 2009). The program is called `gamupscale` (very similar to `gammabar`) and input parameters are specified in Table 1. On line 1, the size of the block is set equal to the panel size; discretization parameters on line 2 should be set identical to those used in `blusim` for consistency; on line 3, the number of directions to upscale the variogram should be set to the three principle direction of the input LMC that starts on line 10; lines 4 to 6 specify the directions using block offsets much like `gam` from GSLIB (Deutsch and Journel, 1998); the resulting upscaled variograms can be standardized by setting the parameter on line 7 to 1; The measure of dissemination is only possible when one variable is considered (this is a residual parameter from `gammabar`); and from lines 10 to the end of the parameter file, there must be  $N(N + 1)/2$  variograms input where  $N$  is the number of variables specified on line 9.

**Table 1: Parameters for `gamupscale`.**

Line	Parameter	
1	10.0 10.0 1.0	-X,Y,Z size of block
2	4 4 1	-X,Y,Z discretization
3	3 50	-number of direction to upscale, number of lags
4	4.8 0.0 0.0	-direction 1: dx, dy, dz
5	0.0 4.8 0.0	-direction 2: dx, dy, dz
6	0.0 0.0 1.2	-direction 3: dx, dy, dz
7	0	-standardize the variograms
8	0	-calculate measure of dissemination, 1=yes
9	4	-number of variables (LMC if > 1)
10	1 0.001 1 1	-nested structured, nugget, head, tail
11	1 0.999 0.0 0.0 0.0	-struct, cc, angl, ang2, ang3
12	200.0 200.0 50.0	-rangeh1, rangeh2, rangev

The program evaluates the average covariance and average variogram between panels in the directions specified for all variograms in the LMC. A nugget effect and sill of the upscaled variogram is predicted using a separation distance of zero between panels. The nugget effect is approximated using Equation (4), where  $\bar{\gamma}(\mathbf{0})$  is

the average variogram computed with zero separation and  $b_0$  is the nugget effect of the input LMC. The sill is approximated from the average covariance and nugget effect by Equation 5, where  $\bar{C}(\mathbf{0})$  is the average covariance computed with zero separation. These approximations are done for all direct and cross variograms in the LMC. An advantage of the approximation of the nugget effect is that no a priori knowledge of the representative scale of the input LMC is needed. For all other separations greater than zero, the average variogram is calculated using Equation (6).

$$\bar{b}_0 = \bar{\gamma}(\mathbf{0}) \cdot |b_0| \quad (4)$$

$$\bar{\sigma}^2 = \bar{C}(\mathbf{0}) + \bar{b}_0 \quad (5)$$

$$\bar{\gamma}(\mathbf{h}) = \bar{C}(\mathbf{0}) - \bar{C}(\mathbf{h}) + \bar{\gamma}(\mathbf{0}) \quad (6)$$

Output from `gamupscale` is identical to the output from `gamv` and `gam` from GSLIB. One file for each variable pair is output and contains the average variogram for each direction and each lag specified. These files are labelled `gamu_vars_jj.out` with  $i = 1, \dots, N, j = i, \dots, N$ . An additional file labelled `varfitpars.out` that contains the estimates of the nugget effects and sills that can be used to constrain `varfit_lmc`. The upscaled LMC is obtained by running `varfit_lmc` using the output variogram files and parameters. An initial run may be required to assess which variogram structures are appropriate, since a large enough change in scale can lead to a change in structure. One recommendation is to use a Gaussian variogram for one of the structures as upscaling does lead to increased continuity at the origin in many cases. The process of upscaling an LMC is demonstrated later in an example.

#### 4. Reorganizing Realizations

Panel scale realizations from `blusim` only provide local distributions of uncertainty that can be utilized on a panel by panel basis. Any form of uncertainty for a subset of panels cannot be evaluated; however, the local distributions can be used to generate realizations that do possess the correct spatial correlation. These can be visualized and used to assess more than just local uncertainty. For each panel, the set of realizations represent an empirical CDF (ECDF). The set of ECDF's for all panels are used in the following workflow to create a new set of realizations that have the correct LMC:

1. Generate initial `blusim` realizations (this includes declustering, normal score transformation and LMC modeling).
2. For each panel, sort the estimates in ascending order to be used as ECDF's. These are denoted by  $\mathbf{z}_k, k = 1, \dots, N$ , where  $N$  is the number of panels.
3. Upscale the LMC to the panel scale by executing `gamupscale` and `varfit_lmc`.
4. Generate unconditional correlated realizations using the upscaled LMC and the same panel scale grid used for the `blusim` realizations.
5. Convert the unconditional realizations into lookup tables. This is done for each panel by determining the sorted index, which is done as follows with  $\mathbf{y}$  representing the set of unconditional realizations for a given panel:
  - a. Sort  $\mathbf{y}$  and retain indexes  $I$  so that  $\mathbf{y}(I)$  is sorted in ascending order.
  - b. Sort  $I$  and retain indexes  $J$  so that  $\mathbf{y}(I(J)) = \mathbf{y}$ .
  - c. The set,  $J$ , are the lookup indexes.
6. Use the lookup indexes to re-organize the ECDF's from step 2:  $\mathbf{z}'_k = \mathbf{z}_k(J), k = 1, \dots, N$ .

An advantage of this workflow is the new set of realizations consists of the same estimates as the original set of realizations; therefore, local volumes and averages are unchanged. It is also straightforward to implement. For step 4, `sgsim` can be used to generate a set of unconditional, uncorrelated realizations that are correlated using a utility called `correlate_realizations`. A similar utility called `correlate` already exists; however, it is intended for a Bayesian updating workflow. `Correlate_realizations` takes a set of `sgsim` realizations and a correlation matrix as input parameters (Table 2). The Cholesky decomposition of the correlation matrix,  $\mathbf{R}$ ,

in Equation (7), where  $\mathbf{L}$  is lower triangular, is used in a matrix multiplication with the realizations in Equation (8), where  $\mathbf{Y}$  is the set of realizations. Correlated realizations are output to a single file.

**Table 2: Parameters for correlate\_realizations.**

Line	Parameter	
1	2	-number of variables to correlate
2	sgsim1.out	-file 1
3	1	-column number
4	sgsim2.out	-file 2
5	1	-column number
6	50 50 1	-nx, ny, nz
7	-998.0 999.0	-trimming limits
8	1.0 0.5	-correlation matrix: 11 12
9	1.0	- 22
10	sgsim.out	-output file

$$\mathbf{R} = \mathbf{L}\mathbf{L}^T \quad (7)$$

$$\mathbf{Y}' = \mathbf{L} \cdot \mathbf{Y} \quad (8)$$

Steps 5 and 6 are built into a utility called `reorder_realizations`. Parameters are shown in Table 3. The program takes realizations from `blusim`, builds the ECDF for each panel (Step 2), and samples from it using the indexes from step 5, which are determined from the unconditional realizations input on line 3. For the file on line 3, it is always assumed to have the same number of variables as on line 2. An index table that was used to define the new order is output to the file on line 7 (one for each variable) and the reordered realizations to the file on line 8. A temporary directory is input on line 9 because there is a substantial amount of disk I/O operations required. To speed up the process, input files are first converted to direct access files and stored in the specified directory. If the directory does not exist, it is created. These files can be deleted afterwards.

**Table 3: Parameters for reorder\_realizations.**

Line	Parameter	
1	blusim.out	-file with realizations to reorder
2	2 1 2	-number of variables and columns
3	sgsim.out	-file with unconditional realizations
4	1 2	-columns, same number as first file
5	100 100 1	-nx, ny, nz
6	-998.0 -999.0	-trimming limits
7	order.out	-output file for index table
8	blusim_ord.out	-output file for reordered realizations
9	./tempfiles/	-directory for intermediate files

## 5. Example

A two dimensional data set with four variables (Ni, Fe, MgO, and Co) is used to demonstrate the presented approach. Figure 1 shows the sample locations shaded by nickel grade. The experimental LMC for input to `blusim` was computed for the normal scores of the four variables using `gamv` and modeled using `varfit_lmc` (Figure 2). Next, `blusim` was executed to generate 50 panel scale realizations. Panels were 20 by 25 meters and the SMU size was 2.85 by 3.58 meters, which discretizes the panels into 7 by 7 blocks. The whole grid consisted of 50 by 44 panels covering an area of 1000 by 1100 meters. Normally, `blusim` would be executed using the four variables in their original form. Then `blusim` handles the normal score transform, possibly with declustering weights (not used for this example) and the resulting realizations are in the associated grade units. For purposes of checking that the realization reordering methodology results in realizations that reproduce the correct LMC, an

additional run of `blusim` was done using the normal score variables as well. Resulting realizations are Gaussian and experimental variograms should coincide with the upscaled LMC (Figure 3).

50 unconditional realizations for each of the variables were generated using `sgsim` and the upscaled variograms. The four sets of realizations were correlated using `correlate_realizations` and the correlation matrix computed from the sills of the upscaled LMC (Equation (9)). After reorganizing the `blusim` realizations using `reorder_realizations`, reproduction of the panel scale LMC was improved, especially for the shorter lags with higher covariance (Figure 4). Reordering was applied to the realizations in grade units as well. Figure 5 shows a realization of each variable.

$$R_{ij} = \frac{C_{ij}}{C_{ii}^{1/2} \cdot C_{jj}^{1/2}} \quad i, j = 1, \dots, 4 \quad (9)$$

## 6. Conclusions

A methodology and set of programs to reorganize multivariate block realizations so that each realization has the correct spatial structure was presented. The values of each block are not changed in the process; therefore, statistics for individual blocks are preserved. Realizations were shown to reproduce the variogram after being reorganized. Having corrected the spatial structure, it is possible to map individual realizations and use them to assess global uncertainty. Initial, unorganized realizations from `blusim` can only be used for local uncertainty assessment.

## References

- Boisvert, J.B. and Deutsch, C.V., 2007, Multivariate block LU simulation. Center for Computational Geostatistics, 9, 305
- Boisvert, J., Ortiz, J., and Deutsch, C.V., 2008, Local recoverable reserves prediction with block LU simulation. International Journal of Mining and Mineral Engineering 1(1):3-21
- Deutsch, C.V. and Journel, A.G., 1998, GSLIB: geostatistical software library and user's guide. Oxford University Press, 384 pp.
- Frykman, P. and Deutsch, C.V., 1999, Geostatistical scaling laws applied to core and log data. SPE 56822
- Jewbali, A., 2009, Finding the nearest positive definite matrix for input to semi-automatic variogram fitting (`varfit_lmc`). Center for Computational Geostatistics, 11, 402
- Journel, A.G. and Huijbregts, Ch.J., 1978, Mining geostatistics. Academic Press Limited, 600 pp.
- Ortiz, J. and Deutsch, C.V., 2007, Local recoverable reserves prediction with block LU simulation. Center for Computational Geostatistics, 9, 304

Figures

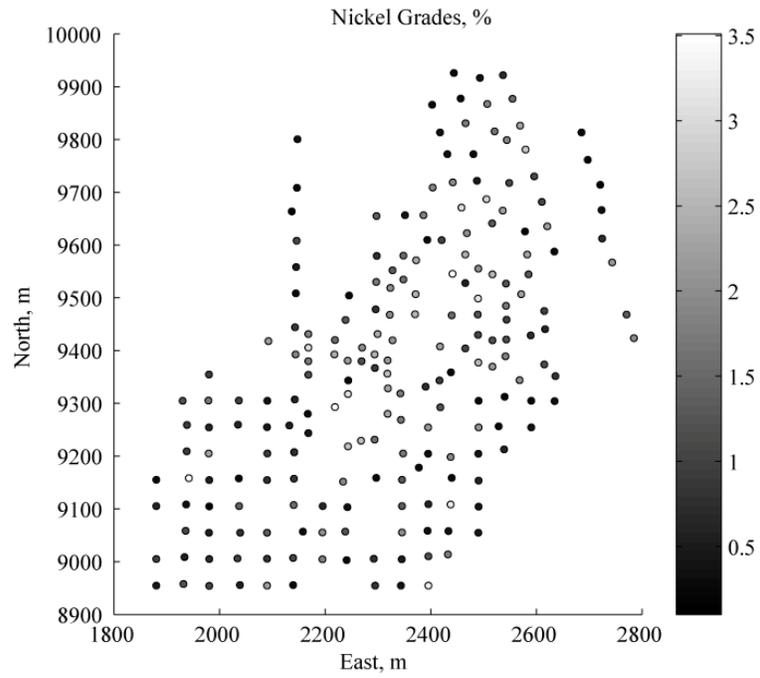


Figure 1: Locations of samples.

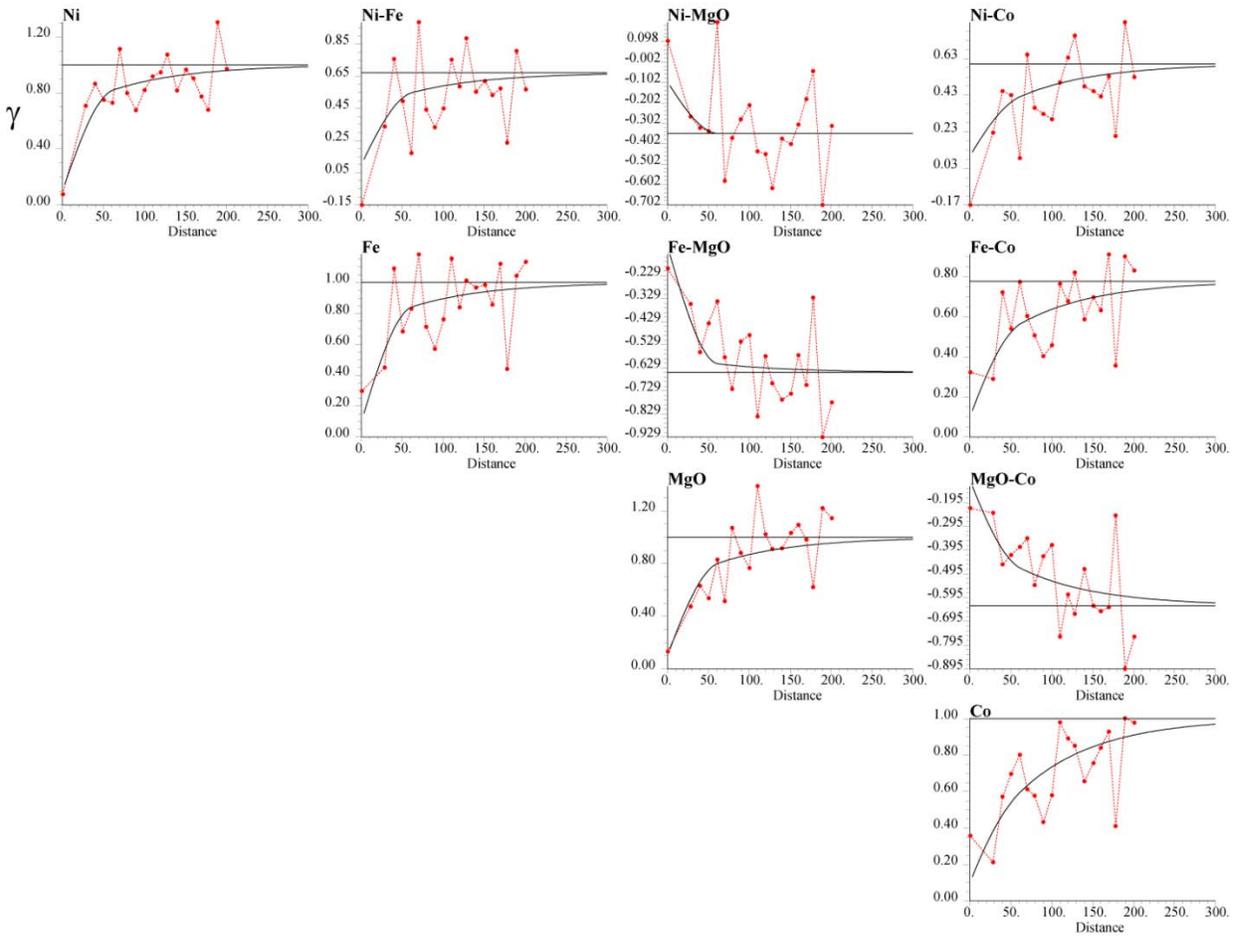


Figure 2: LMC of normal scores in the East direction. Dashed line shows experimental and solid lines show the fit model.

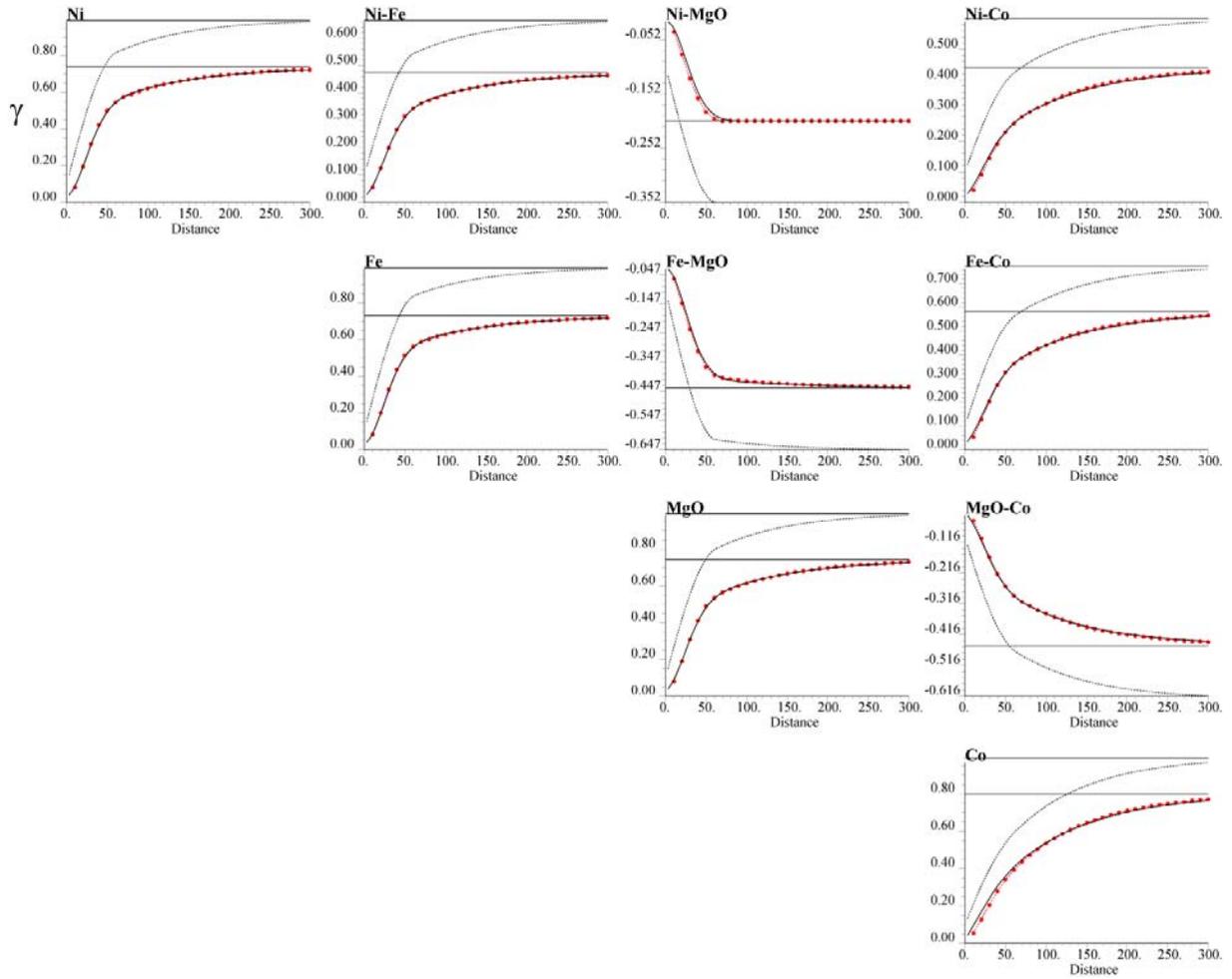
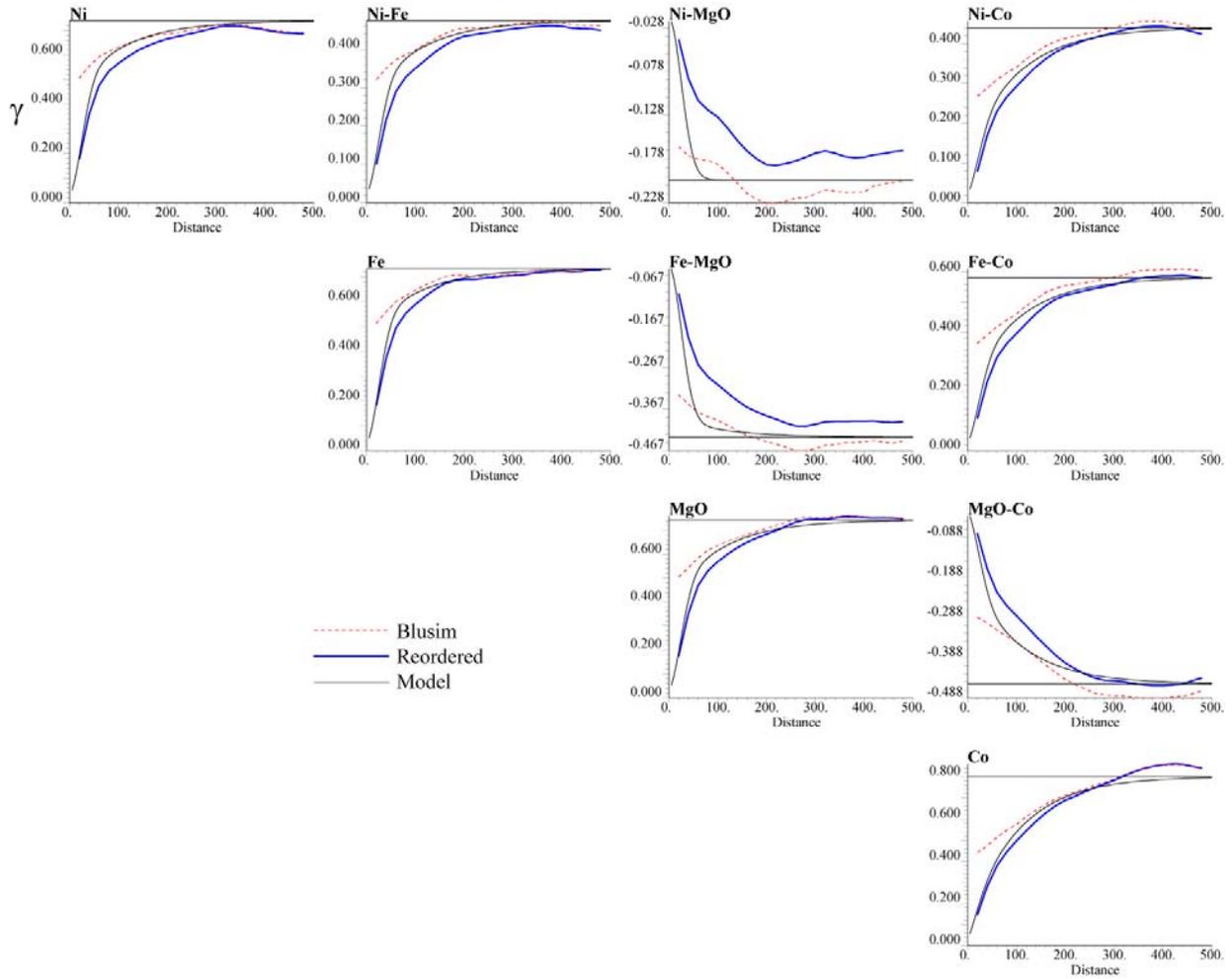


Figure 3: Upscaled LMC. Points output from `gamupscale`, dashed line is point scale, solid line is panel scale.



**Figure 4:** Variogram reproduction before and after realization reordering. Blusim and reordered variograms are the average of 50 realizations.

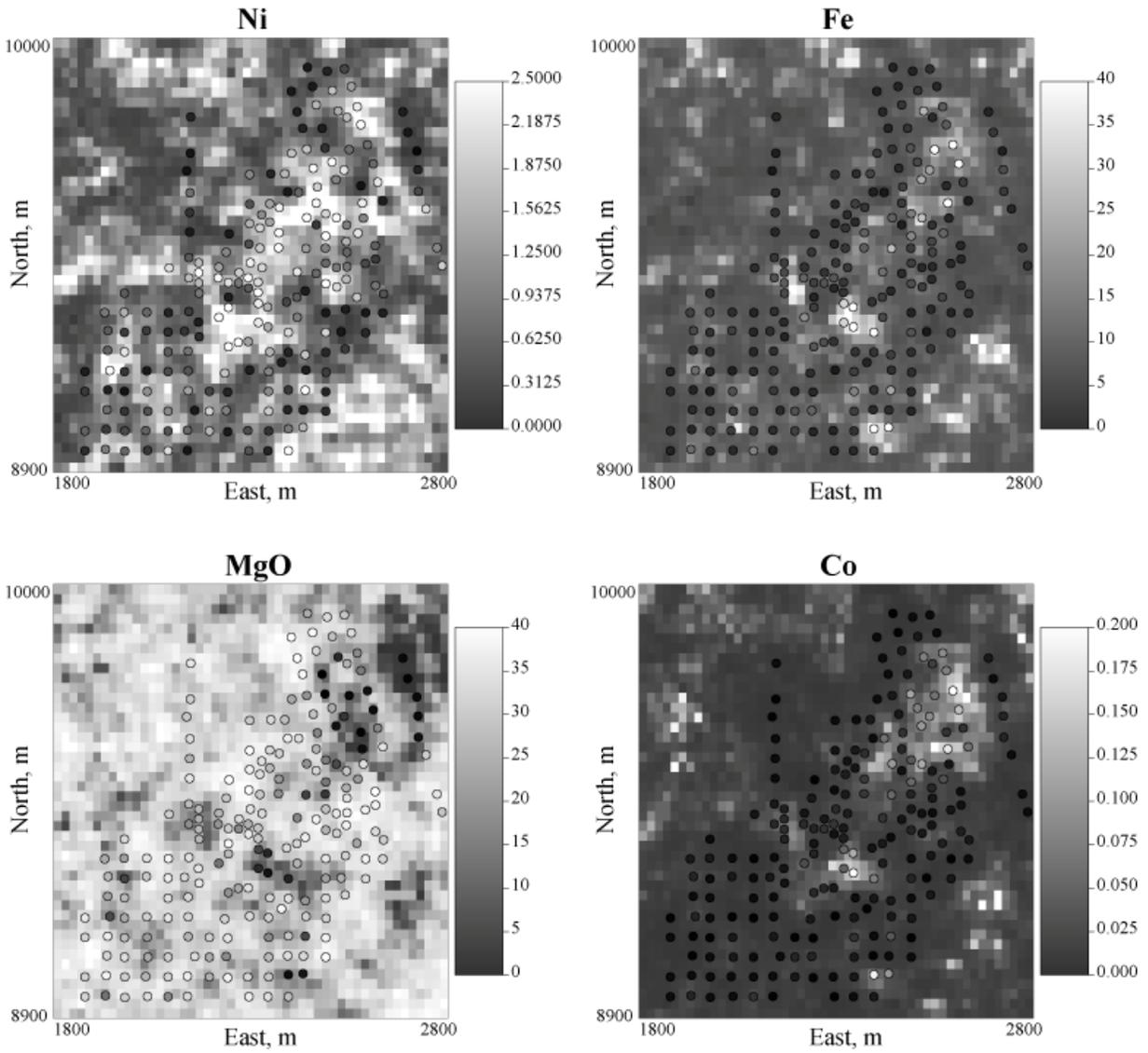


Figure 5: Reordered realizations and conditioning data.