Multivariate Geostatistical Simulation of Proportions and Nonadditive Geometallurgical Variables

J.B. Boisvert, Mario E. Rossi and Clayton V. Deutsch

Recovery and plant performance outcomes are influenced by a large number of variables, including head assays, mineralogy and mineral associations. Models that utilize all these variables outperform models based on head assays alone. This paper discusses the geostatistical modeling of the required input variables. Due to the nature of the head assays and mineral association data, unique challenges arise when attempting to model such variables. Specifically, the compositional nature of the variables must be accounted for. Moreover, many of the variables are correlated and require methodologies that are simple, yet effective. In the proposed methodology, data transformations are used to maintain the compositional nature of the variables and PCA analysis is used to consider complex relationships between the variables.

Introduction

In this paper, modeling methodologies are developed for a total of 135 variables, separated into three groups: head grade assay values; grain size measurements; and mineral associations. Significantly more samples exist for the head grade variables, therefore they are modeled first. The grain size and association variables are modeled using the head grade realizations as secondary information.

The head grade and mineral association data are considered compositional, that is, they sum to 100%. A logarithmic transform is used to deal with this constant sum constraint. Normally, these variables would then be cosimulated with sequential Gaussian simulation (SGS); however, the large number of variables available and the large grid size makes this procedure too computationally intensive. An alternative is to perform a principal component (PCA) transform on the logarithmic data to generate uncorrelated variables. SGS is then preformed on the uncorrelated PCA values. The values are back-transformed into original units to generate the realizations. This procedure is used to model the head grade and mineral association data. The grain size data, which are not compositional, are modeled using sequential Gaussian cosimulation for the p_{20}, p_{50} and p_{80} values of each mineral.

Modeling 23 head grade variables

Recall that the plant performance modeling in paper 302 required a total of 23 head grade variables for input into the linear regression models: Cu, U3O8, Ag, Au, Co, Mo, Pb, Zn, Ba, Fe, Al, Si, K, Ca, S, Co2, La, Mg, Mn, Na, P, Ti, Ce. These 23 variables are simulated on a grid with the following dimensions: xmin=56105; ymin=30515; zmin=-1932.5; xsiz=10; ysiz=10; zsiz=15; nx=360; ny=624; nz=119. There are a total of 111,572 head assay samples used in the modeling. Note that the K:AL ratio and $B_{adj}S$ are also required, but are simply calculated from the realizations of K, Al, Ba and S.

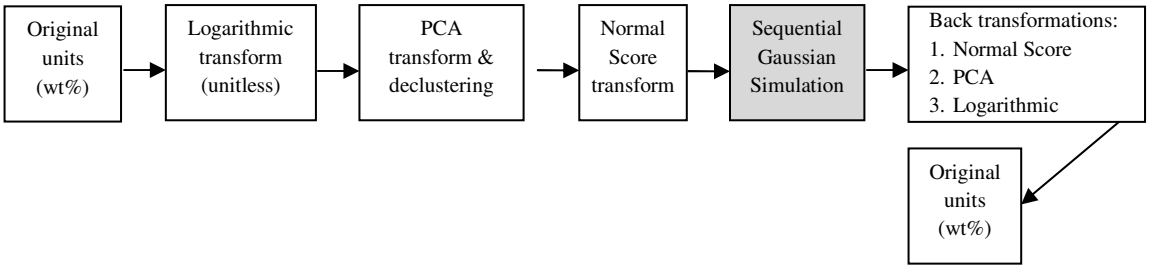
The head grade variables are considered compositional because all chemical and mineral rock components must sum to 100%. Because not all elements in a sample are assayed, the sum of the head grades is always less than 100%. However, in geostatistical modeling if this constraint is not explicitly imposed it can be violated in some areas of the model. For this reason a logarithmic transform of the 24 head grade variables is considered. There are 24 variables because the remaining proportion of the sample is included to impose the 100% constant sum (i.e. the 23 variables listed above + 1 filler variable). The logarithmic transform is:

$$y_i = \ln\left(\frac{x_i}{x_{filler}}\right)$$

where y_i is the new variable to be modeled and x_i are each of the 23 variables to be modeled. This transformation requires that there are no zero values for any variable as ln(0) is undefined. The back transformation is:

$$\frac{x_i = e^{y_i}}{\sum_{i=1}^{24} e^{y_i} + 1}$$

There are now 23 logarithmic transformed variables. There is a complex relationship between these 23 variables, consider their correlations (Figure 1). It would be difficult to reproduce all these relationships with traditional SGS, therefore, an additional transformation is considered. The PCA transform generates 23 new *uncorrelated* variables. These variables are linear combinations of the 23 logarithmic variables but are uncorrelated. An assumption of independence between the 23 variables is then made and all 23 PCA variables are modeled independently with SGS. This ensures good reproduction of the correlation between the 23 variables in the final realizations (Figure 1). An overall summary of the transformations used is shown below:



This methodology assumes that the normal score values of the principal components are independent. The PCA transform ensures that the correlation between the components is zero; however, the components may not be independent. Poor histogram reproduction is seen in original units due to this lack of independence. There are a large number of head assay samples which makes the input histogram reliable; they should be reproduced in the simulation. To obtain reasonable histogram reproduction the final simulations are post processed with TRANS to better match the declustered input histogram (Figure 2). This has little effect on the correlations between variables and individual variable variograms, but improves histogram reproduction.

Details of SGS

Implementation of SGS requires the use of variograms for each PCA variable as well as a number of other important parameters. For all variables considered in this paper, simulation was preformed with 50 nearby data (25 data and 25 previously simulated nodes). Parameters for each variogram can be found in Table 1. Because of the large number of variables, variogram fitting software was used with a visual assessment to locate any major inconsistencies with data.

Declustering was used on the 23 PCA variables to obtain global histograms. A locally varying mean was used in the simulation to consider the nonstationary present throughout the deposit. The mean for each PCA was determined using a moving window average with a radius of 400m in the horizontal direction and a 50% anisotropy in the vertical direction.

Modeling 9 grain size variables

There are three minerals of interest: Brannerite, Coffinite and Uraninite. The p_{20}, p_{50} and p_{80} grain size for each mineral has been measured at 497 locations. There is very little correlation between the minerals (Figure 3) so each mineral is modeled independently. The correlation between the percentiles of each grain size (Figure 3) is reproduced by cosimulating the three percentiles.

The densely sampled 23 head grade values is used to supplement the lack of information for the grain size variables by considering a super secondary variable which is the amalgamation of the 23 PCA head grade variables. This super secondary variable is created differently for each mineral because the correlations between the mineral grain sizes and the PCA head grade variables differ. To generate this super secondary variable, a linear combination of the PCA head grades is determined from the following equations (similar to paper 302 and a technique that is most often used in Bayesian Updating):

$$\begin{bmatrix} \rho_{1,1} & \rho_{2,1} & \cdots & \rho_{n,1} \\ \rho_{1,2} & \rho_{2,2} & \cdots & \rho_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1,n} & \rho_{2,n} & \cdots & \rho_{n,n} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix} = \begin{bmatrix} \rho_{0,1} \\ \rho_{0,2} \\ \vdots \\ \rho_{0,n} \end{bmatrix}$$

Table 1: Variograms for the normal score of the PCA head grade variables. A nugget (C0) and two spherical structures (C1 and C2) were used with no plunge angle.

Variable Name	C0	C1	C2	Azimuth 1	Dip 1	Range 1			Azimuth 2	Dip 2	Range 2		
						Major	Minor	Vertical			Major	Minor	Vertical
NS:PCA 1	0.11	0.345	0.544	104	-75	118	79	65	100	-86	1141	1556	548
NS:PCA 2	0.035	0.608	0.357	186	83	67	54	63	158	-56	1417	606	482
NS:PCA 3	0.219	0.348	0.432	360	-80	282	110	197	360	-80	294	1193	945
NS:PCA 4	0.212	0.283	0.505	38	-76	314	79	108	349	-82	530	1627	1488
NS:PCA 5	0.292	0.378	0.33	290	-40	166	166	209	290	-40	670	1449	1303
NS:PCA 6	0.081	0.716	0.202	106	-89	59	54	48	113	-68	535	350	192
NS:PCA 7	0.107	0.302	0.59	50	-76	85	44	55	38	-61	716	1571	947
NS:PCA 8	0.168	0.415	0.417	88	-89	101	60	53	106	-79	471	606	247
NS:PCA 9	0.19	0.455	0.356	89	90	80	64	54	109	-69	496	454	237
NS:PCA 10	0.19	0.545	0.266	311	-12	54	62	73	354	-31	398	210	1020
NS:PCA 11	0.216	0.442	0.342	130	-80	96	68	72	130	-80	550	442	284
NS:PCA 12	0.188	0.426	0.386	281	-16	53	57	81	353	-39	296	247	672
NS:PCA 13	0.239	0.376	0.385	21	83	76	50	55	101	-42	446	713	311
NS:PCA 14	0.201	0.544	0.254	214	-2	49	42	61	224	-45	272	169	290
NS:PCA 15	0.451	0.463	0.085	292	-15	104	141	263	283	24	3791	943	25404
NS:PCA 16	0.234	0.561	0.205	23	-83	68	46	55	44	-58	280	280	784
NS:PCA 17	0.465	0.45	0.085	307	-7	99	122	203	283	-81	43720	1311	35267
NS:PCA 18	0.29	0.424	0.286	198	-5	52	52	67	194	-34	999	374	487
NS:PCA 19	0.211	0.559	0.23	100	-70	55	55	47	145	-73	839	220	148
NS:PCA 20	0.195	0.564	0.241	326	-5	53	57	65	5	-16	684	480	1160
NS:PCA 21	0.332	0.627	0.042	280	-20	51	57	70	280	-20	25464	535	8428
NS:PCA 22	0.305	0.25	0.445	294	-30	81	106	157	281	-61	683	683	365
NS:PCA 23	0.598	0.19	0.212	232	70	142	106	132	231	-53	2037	1022	786

The right hand side of this equation contains the correlation between one of the grain size variables and the 23 input head grade variables to be merged. The left hand side is the correlation between all 23 PCA head grade variables; note that the left hand side contains 1.0 on the diagonal and 0.0 for all off diagonal terms because the PCA values are uncorrelated. This is done for the p_{50} value for each mineral and the same super secondary variable is used for modeling the p_{20}, p_{50} and p_{80}. This single super secondary variable allows for the cosimulation of the three percentiles and only one exhaustive secondary variable. Without merging all secondary variables into a super secondary, the grain size simulations would have to consider 23 separate secondary variables in order to use all the available information from the head grade variables.

The super secondary variable is used as a collocated secondary variable for each of the grain size models. Note that for the grain size variables neither a logarithmic nor a PCA transformation is considered because there are only three variables (p_{20}, p_{50} and p_{80}) for each mineral. Cosimulation of three variables can be accomplished in a reasonable amount of CPU time. This procedure is repeated for Brannerite, Coffinite and Uraninite. This includes building a new super secondary variable for each mineral.

Very few data exist for the grain size variables and the variograms are unstable (Figure 5). The same variograms are used for the p_{20}, p_{50} and p_{80} of each mineral. The spatial structure for the p_{20}, p_{50} and p_{80} are similar (Figure 5); differences are likely due to a lack of data. Parameters for the variograms used are shown in Table 2.

Table 2: Variograms for the grain size data. A nugget (C0) and two spherical structures (C1 and C2) were used with no plunge/dip angle and no horizontal anisotropy.

Variable Name	C0	C1	C2	Range 1		Range 2	
				Horizontal	Vertical	Horizontal	Vertical
Brannerite	0.4	0.2	0.4	200	20	200	150
Coffinite	0.4	0.2	0.4	400	20	400	300
Uraninite	0.4	0.2	0.4	200	20	200	350

Modeling 100 association matrix variables

Modeling the association matrix utilizes a combination of the techniques previously discussed. The matrix is a 10x11 matrix where each row sums to 1.0 (or 100%). Consider this sample:

	Brannerite	Coffinite	Uraninite	Pyrite	Chalcopyrite	Bornite	Chalcocite	Other Sulphides	Acid Soluble Gangue	Acid Insoluble Gangue	Free Surface
Brannerite		8.02								88.18	3.80
Coffinite	1.71		1.64			0.25	0.24		3.50	90.67	2.00
Uraninite		23.51								76.49	
Pyrite											
Chalcopyrite						2.83			2.59	88.43	6.15
Bornite		0.18			0.93				15.50	75.89	7.49
Chalcocite		0.30							0.87	97.91	0.92
Other Sulphides										100.00	
Acid Soluble Gangue		0.05			0.02	0.32	0.01			91.16	8.44
Acid Insoluble Gangue	0.04	0.19	0.01		0.08	0.22	0.16	0.02	12.82		86.45

Each element in the matrix represents the % surface area of interaction between minerals determined from mineral liberation analysis. Each row sums to 1.0; however, each column does not sum to a constant value as the values are standardized by the proportions. There are a total of 100 elements in the matrix to be modeled, ignoring the diagonals. An assumption that the rows are independent is made to reduce the problem to simulating 10 independent sets (rows) of 10 dependant variables (columns). To maintain the constant sum constraint the logarithmic transformation is applied to each row resulting in the need to model 9 logarithmic variables, the logarithmic transform reduces the number of variables by 1. The PCA transformation is applied to reproduce the correlation between variables in each row. The principal components of each row are normal score transformed and then simulated with SGS. There are a total of 490 data available for simulation of association variables.

As with the grain size variables, the head grade simulations provide a super secondary variable to use in collocated SGS. There are a total of 23 (normal score PCA) head grade simulations to be combined into a single super secondary variable for each of the 100 elements in the association matrix. The PCA transform is done in such a way that the amount of data explained by each principal component can be measured. Some components 'contain' more information than others. In this case the first 5 components of the head grade realizations contain over 75% of the information in the original head grades. Only the first 5 principal components generated in the head grade modeling are combined into the super secondary variable to reduce the computational requirements of the methodology. Moreover, the super secondary variable is only used for the first 4 of the 9 principal components of the association variables. Because there are 100 association variables to model, CPU time becomes an issue.

A variogram is required for each of the 90 principal components (10 sets/rows with 9 principal components in each). As with the head grade variables these variograms were fit with automatic variogram fitting software and visually inspected for inconsistencies.

Special considerations for the association data

Missing or "null" values always pose a problem in compositional data modeling. In this instance there are some entries that are missing because a particular mineral does not appear in a given sample. For rows that have some missing values but still sum to 1.0, the missing values are reset to 0.0001 or 0.01%. In some cases there are entire rows that are missing. This is because the mineral does not appear at that location; however, in these cases all values cannot be set to a small value as they would not sum to 1.0. The solution undertaken in this study was to remove the samples where the entire row was missing. When performing SGS at this location the values in that particular row are simulated as if the data did not exist (in fact this data does exist and has a value of zero). There is a miss-match between the missing values at this location and the simulated values given the surrounding data. It is intended that if these models are to be used in the future, some type of rock type modeling is preformed. At these mismatch locations the missing minerals have a 0.0 proportion and the mismatched association values are ignored.

Histogram/variogram reproduction

There are 135 variables modeled in total. The histogram/variogram reproduction for the first 3 realizations have been assessed. The following discussion compares the input histograms and variograms to the realization outputs.

Head grade variables

The head grade variables reproduce the histogram quite well (Figure 2) because of post-processing with TRANS. Variogram reproduction is checked in normal score units of the principal components.

Grain size variables

Histogram and variogram reproduction (Figure 6) is heavily influenced by the secondary variables as there was very little grain size data. For this reason the histogram and variogram reproduction for the grain size variables does not exactly match the input. Moreover, the grain size variables are sparsely sampled suggesting that the input histogram and variogram may be unreliable. Some deviation from the input parameters due to the secondary information is warranted.

Association matrix variables

There are a total of 100 association variables. Histogram reproduction is not perfect. The resulting histograms and variograms deviate from the input for two reasons:

1) Lack of independence of the principal components.
2) Influence of the super secondary attributes on the models.

The post processing applied to the head grade variables could be applied to the association variables but this would be highly CPU intensive.

Conclusions

Paper 302 presented three linear regression models for the prediction of plant performance from head assay, mineralogy and association variables. This paper presented a methodology for the spatial modeling of these variables. The intention is to use the models developed in paper 302 from the pilot plant trials with the spatial models in this paper to predict plant performance. The cost of obtaining samples of plant performance (i.e. pilot plant runs) is very high. Building models, such as those presented in paper 302 and 303, based on the sparse sampling of mineral recovery, acid consumption and work indexes allows for the mapping of these variables for all locations in the mine. This provides an advanced measure of complex process based variables that rarely have sufficient data density to generate appropriate variograms and prove difficult to effectively model.

Figure 1: Left - correlation between the head grade variables. Right: correlation in one simulation. Correlations calculated in original units.

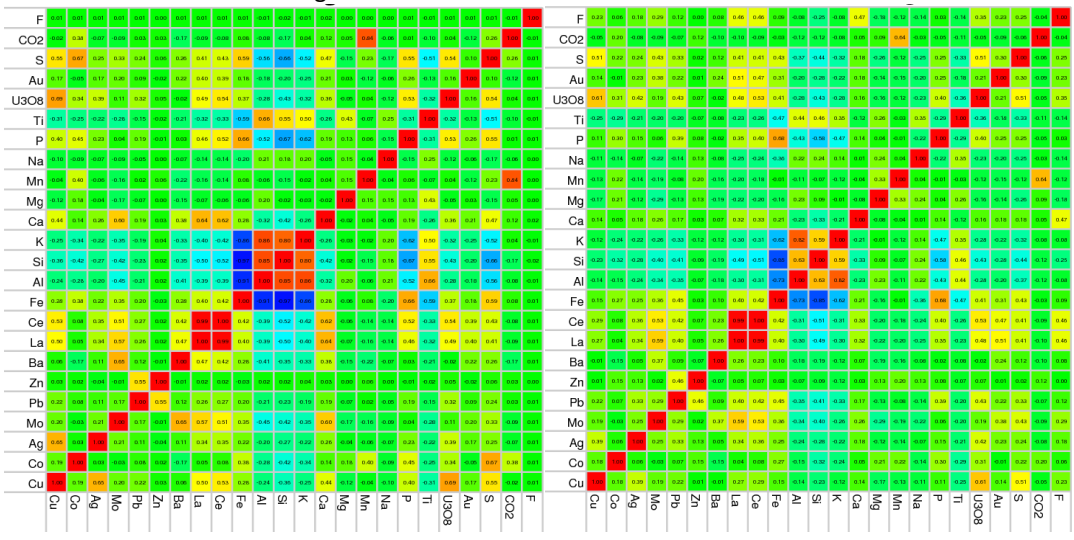


Figure 2: Histogram reproduction for 25 head grade variables after post processing. Black – 3 realizations. Red – Input histogram with 7038 data.

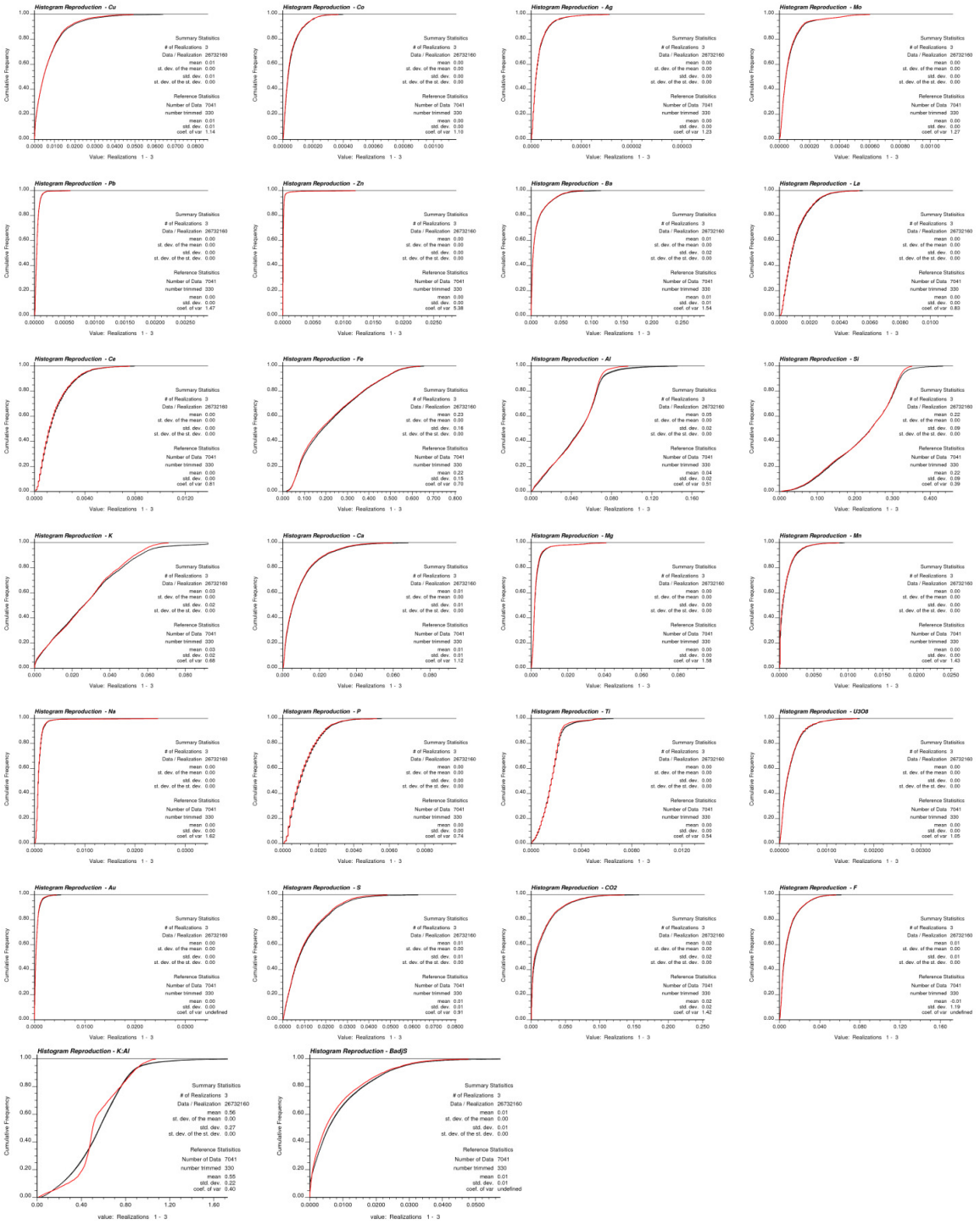


Figure 3: Correlation between the grain size variables. Minerals are simulated independently because of the small correlation between minerals.

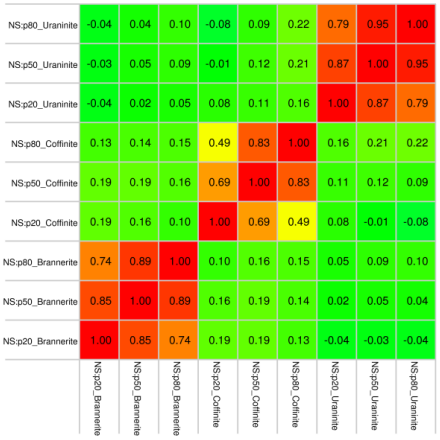
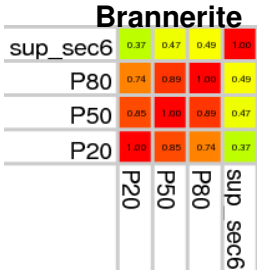
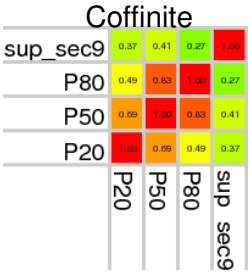
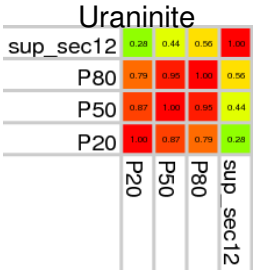


Figure 4: Correlation between the grain size variables. Above – correlations from 497 data to the super secondary variables. Below – correlations from one grain size simulation







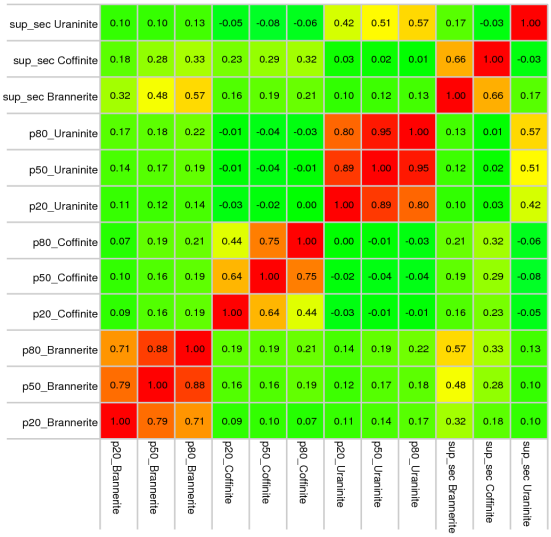


Figure 5: Modeled variograms for the 9 grain size variables. The same variogram was used for the percentiles of each mineral.

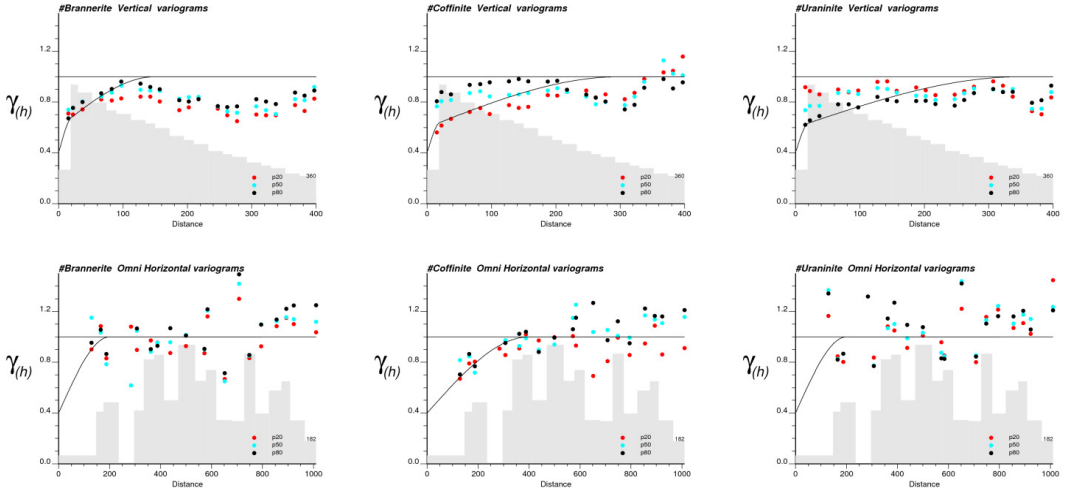


Figure 6: Histogram reproduction for grain size variables. Black – 2 realizations. Red – Input histogram.