Geostatistics - Geostatistical simulation: examples

Julián M. Ortiz - https://julianmortiz.com/

October 5, 2024

Summary

Sequential Gaussian and sequential indicator simulation are commonly used to build models of continuous and categorical variables. In this chapter, we present examples of implementation of both methods and show how these are integrated to account for joint geological and grade uncertainty, in the context of a resource modeling example in mining.

1 Example

1.1 Categorical simulation

In this example, we start by considering the rock types as a categorical variable. For simplicity, we model the presence and absence of unit 20. In order to prepare the data, we code the rock type to discriminate between samples with rock type 20 vs all the rest. The new codes are:

- Rock type 20: 200
- All other rock types: 100

Sequential indicator simulation (for categorical variables) requires the definition of the number of categories (2 in this case), the categories codes (200 and 100 in this case), the global proportions, the indicator variograms and a set of parameters for the search.

Global proportions

Categorical variables, as continuous variables, may be sampled in a non-representative way. The high grade domain may be preferentially sampled, thus, when computing the global proportions, we need to correct for this spatial bias via declustering.

The same parameters used for declustering the grades are used to decluster the proportions. A cell of size 35 by 35 by 12 m^3 is used to determine the weights assigned to each sample. Statistics before and after declustering are presented in **Figure 1**. It can be seen that the declustered proportions are as follows:

- Code 100 (all other rock types): 65.108 %
- Code 200 (rock type 20): 34.892%

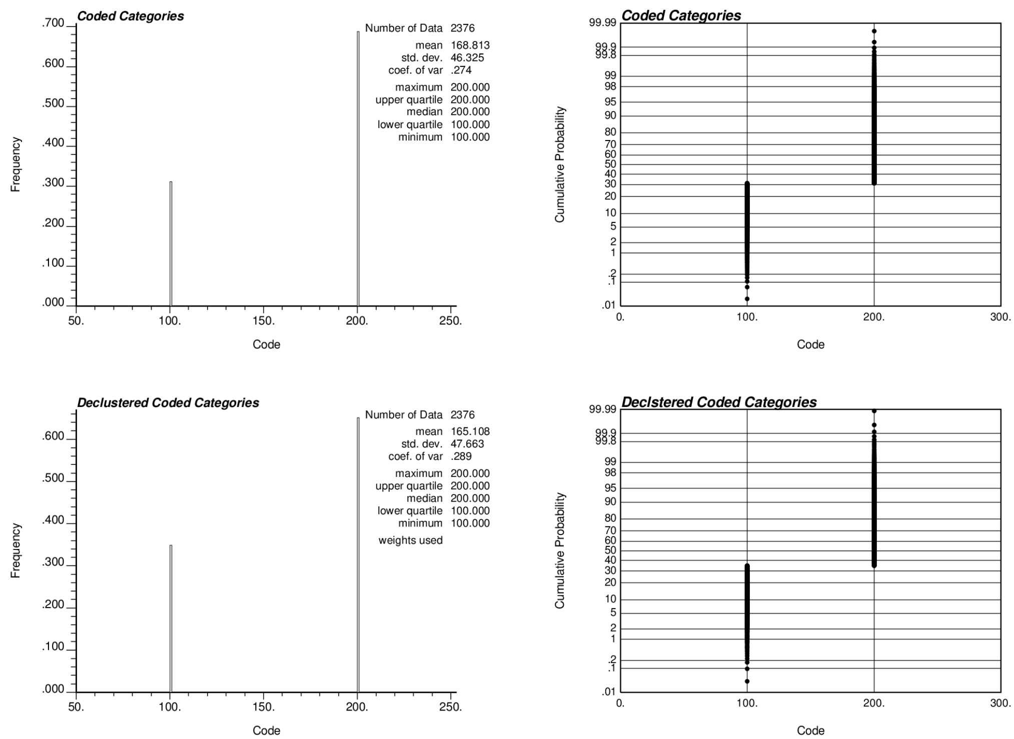

Figure 1: Raw and declustered statistics of the proportions of the coded rock types.

Indicator variograms

Since the sample locations are the same as for the grades, the search parameters for inferring indicator variograms are kept as before. They are summarized in **Table 1**.

Notice that data are coded as indicators for the two cases considered: code 100 and code 200. In each case, the codes are transformed to a binary indicator. It is obvious from looking at the formula of the experimental indicator variogram that in the binary case, the output is the same if the indicators are interchanged, as only transitions from one category

Direction	Lag [m]	Lag Tolerance [m]	Azimuth [°]	Azimuth Tolerance [°]	Dip [°]	Dip Tolerance [°]	Bandwidth [m]
N30°E	35.0	17.5	30	22.5	0	15.0	30.0
N120°E	35.0	17.5	120	22.5	0	15.0	30.0
Vertical	12.0	6.0	0	180.0	-90	15.0	20.0

Table 1: Parameters for indicator variogram calculation

to any other matters for the computation of the indicator variograms.

The resulting experimental indicator variograms are then modeled with a nugget effect and two nested structures. Notice the non-stationarity evident in the experimental results (**Figure 2**). Despite this, we model the indicator variograms as stationary random functions (**Table 2**), with a sill (that depends on the proportions). The plots are standardized to a unit sill to ease the modeling.

> Notice that in most cases, the categories are not properly modelled with a stationary random function, since they tend to present very clear structure (trends) in space, rather than being spread all over the domain in a regular fashion. Therefore, categorical simulation methods always struggle to represent properly the spatial distribution of categories.

Type	Sill	Angle 1	Angle 2	Angle 3	Range Y”	Range X”	Range Z”
Nugget	0.01						
Spherical	0.29	30	0	0	170	140	85
Spherical	0.70	30	0	0	170	140	∞

Table 2: Parameters of the indicator variogram model

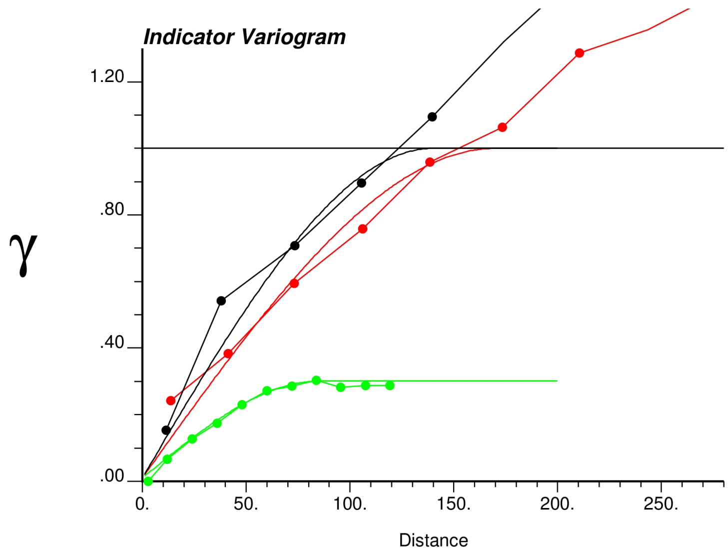

Figure 2: Experimental indicator variogram and fitted model in the three principal directions.

Search plan

A point support grid (consistent with the point support that is later used for simulating the grades) is considered (**Table 3**). The search plan consists of 24 samples and previously simulated nodes within a 70*m* search radius. Since the grid is quite dense, each sample is assigned to the closest node, requiring a single search for conditioning information and making the simulation faster. Multiple grids are used with three levels.

1.2 Results of categorical simulation

The results of sequential indicator simulation are 20 dense realizations conditioned by the sample data. The E-type es-

Coordinate	Nodes	Min Coordinate (m)	Spacing (m)
East (X)	400	24450.5	1.0
North (Y)	600	25050.5	1.0
Elevation (Z)	12	3826.0	12.0
Total	2880000		

Table 3: Definition of the point grid for simulation.

timate provides a map of the likelihood of finding each rock type. The map of conditional variances, reflects the areas where the category is uncertain. Four realizations and the e-type and conditional variance maps are shown in **Figure 3**.

From the maps of the realizations it is clear the non-stationary nature of the variable. The Tourmaline breccia unit (rock type 20, coded as 200) is mainly concentrated at the center of the domain. Some occurrences appear in the simulations at the edges of the domain, but these happen because we use simple indicator kriging, so where conditioning is weak, simulated categories are drawn from the global (prior) proportions.

The e-type map clearly shows where each category should be found. All the green areas are mixed cases, where some simulations output a value of 100 and others a value of 200. These are the areas close to the unknown boundaries between the main unit and the others. This is again seen in the conditional variance map, where a clear countour of the boundary is evident in red. Further to this, areas with large uncertainty are seen where conditioning is scarce.

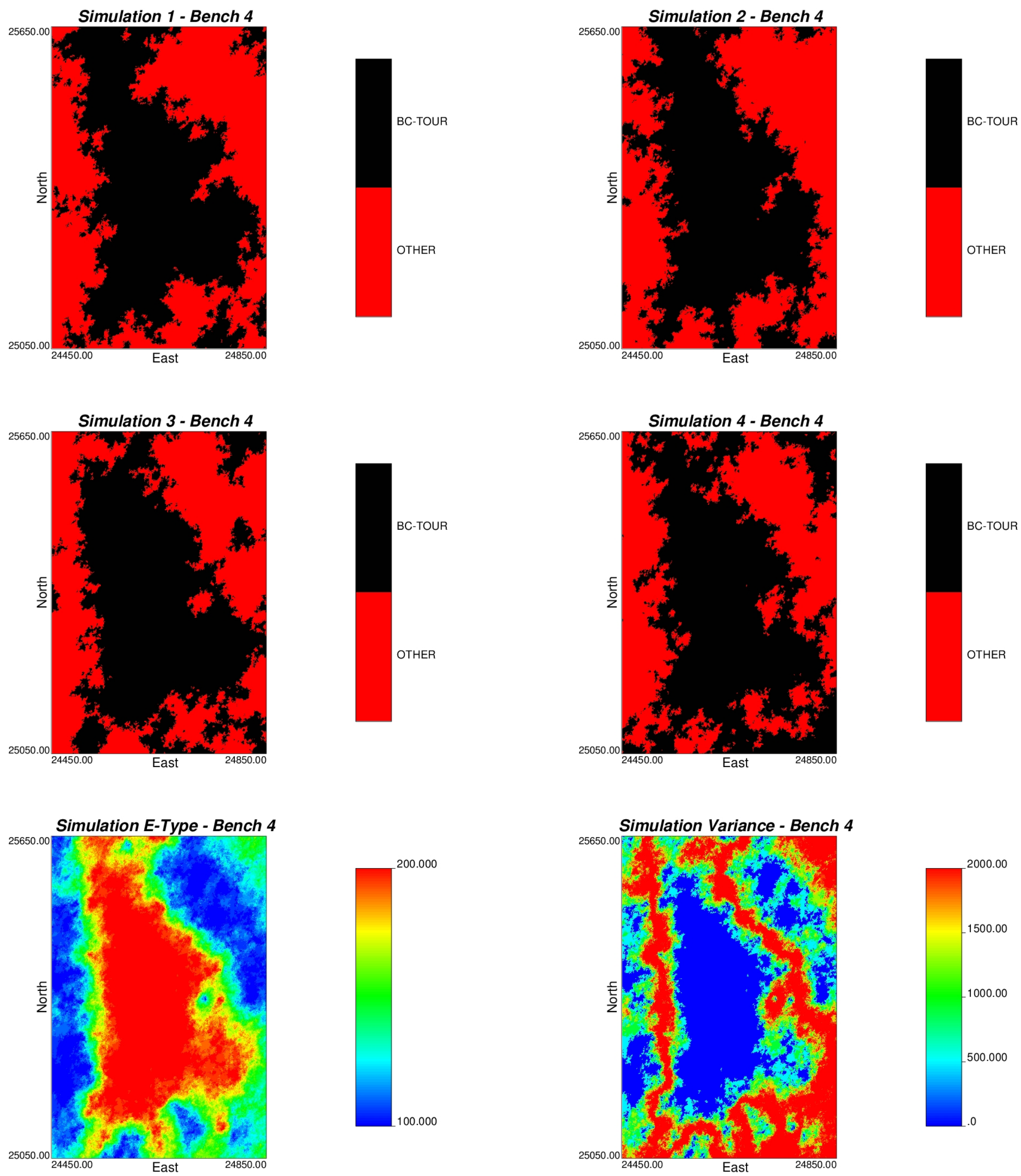

Figure 3: Four realizations and the e-type and conditional variance maps for a representative bench.

1.3 Continuous simulation

The next step is to simulate the spatial distribution of grades, but this has to be done for each domain separately. Since we have two categories, grades are simulated independently in each domain (we will assume hard boundaries between the units).

The procedure is:

- Decluster the data: this can be done jointly or by category
- Transform to normal scores: the grades within each category are transformed independently to a standard Gaussian distribution.
- Calculate and model variogram of normal scores of grades: this is done independently for each category.
- Perform simulation: grades are simulated with sequential Gaussian simulation independently for each category. Notice that the simulation is not yet constrained to the extent of the domain (which is determined in a stochastic manner with sequential indicator simulation). Therefore, the entire domain is simulated.

Once simulations are completed, the models can be merged with the indicator simulations such that:

- For realization 1 to L:
  - Read the value of each node in the indicator simulation realization.
  - Assign the simulated grade for the corresponding category to that location.

Declustering and transformation

The declustered distribution and corresponding normal score transform is shown as an example, for unit 100 in **Figure 4**.

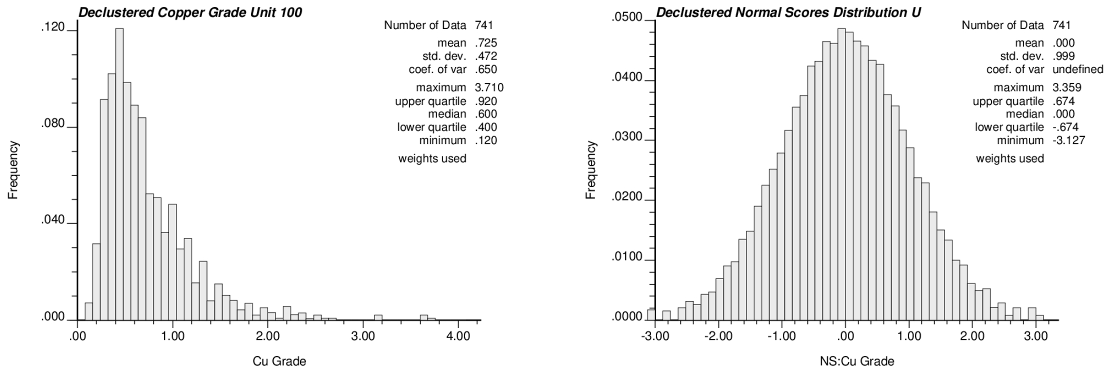

Figure 4: Declustered and transformed histogram for unit 100.

Variograms of normal scores

The variograms for unit 200 were shown when presenting multiGaussian kriging. Here, we show the resulting variograms of normal scores for unit 100 (**Figure 5**).

Search plan

The parameters for Gaussian simulation are the same as for indicator simulation in terms of the grid definition and search plan.

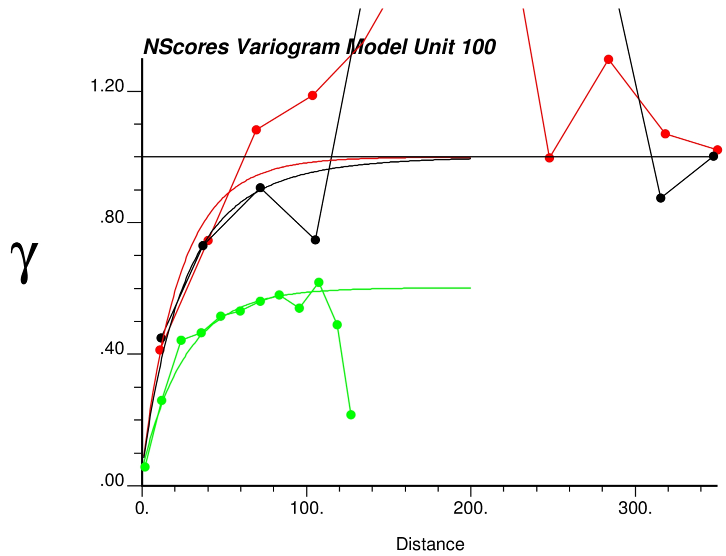

Figure 5: Variogram of normal scores for unit 100.

1.4 Results of continuous simulation

The results of the simulation of grades are presented in **Figures 6** and **7**.

The realizations capture the spatial correlation and anisotropy and show the true variability between points. The e-type follows closely what was seen before with kriging. The conditional variance highlights the proportional effect and the uncertainty in areas of scarce sampling.

Notice for unit 100 the smooth behavior at the center of the deposit, where scarce conditioning exists (since this is where unit 200 is). The estimate tends to the global mean and the variance is high. Also, it can be seen that for the simulation of grades in unit 200, in the borders, conditioning is weak generating a high variance.

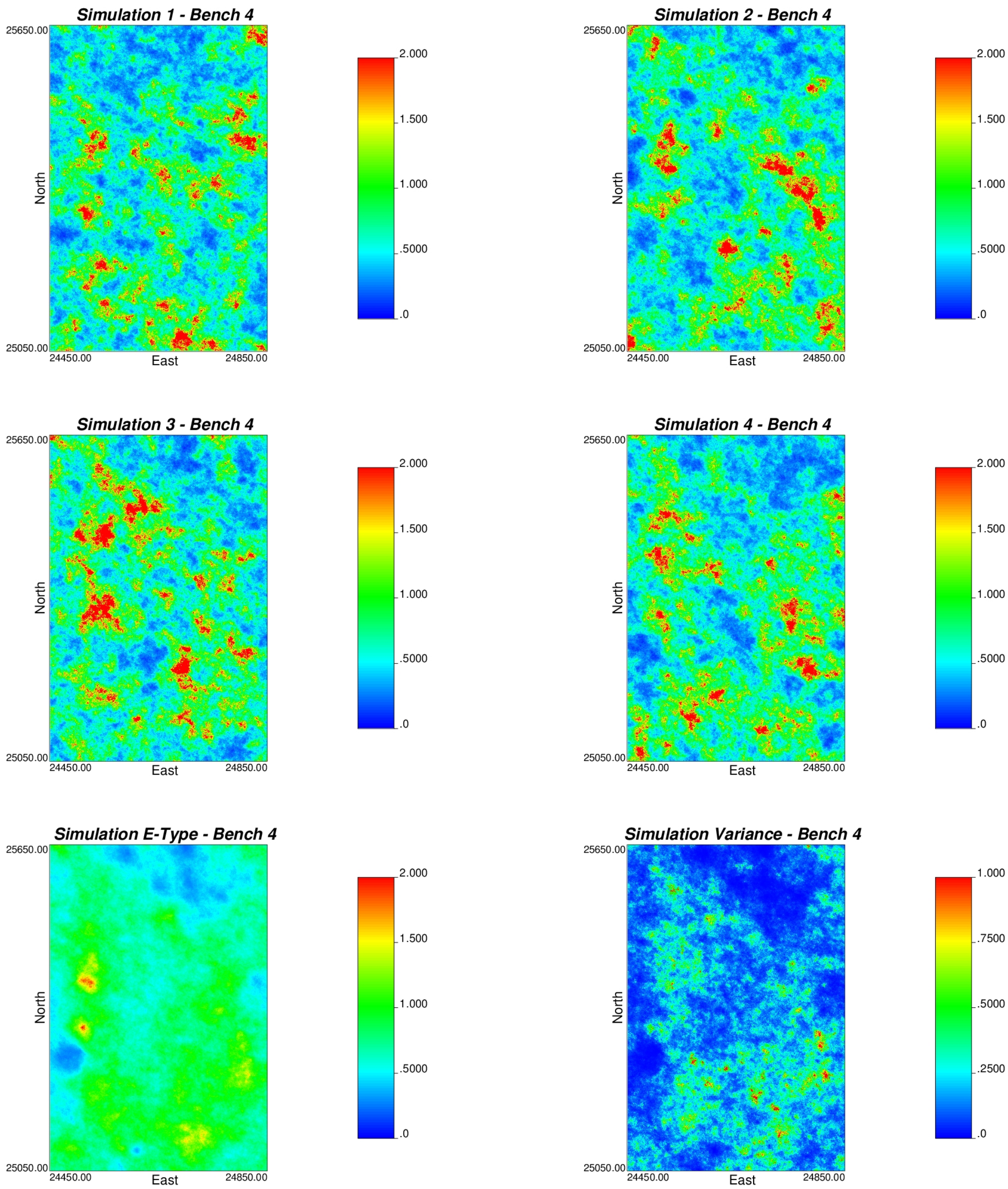

Figure 6: Four realizations, the e-type estimate and the conditional variance for grades in unit 100.

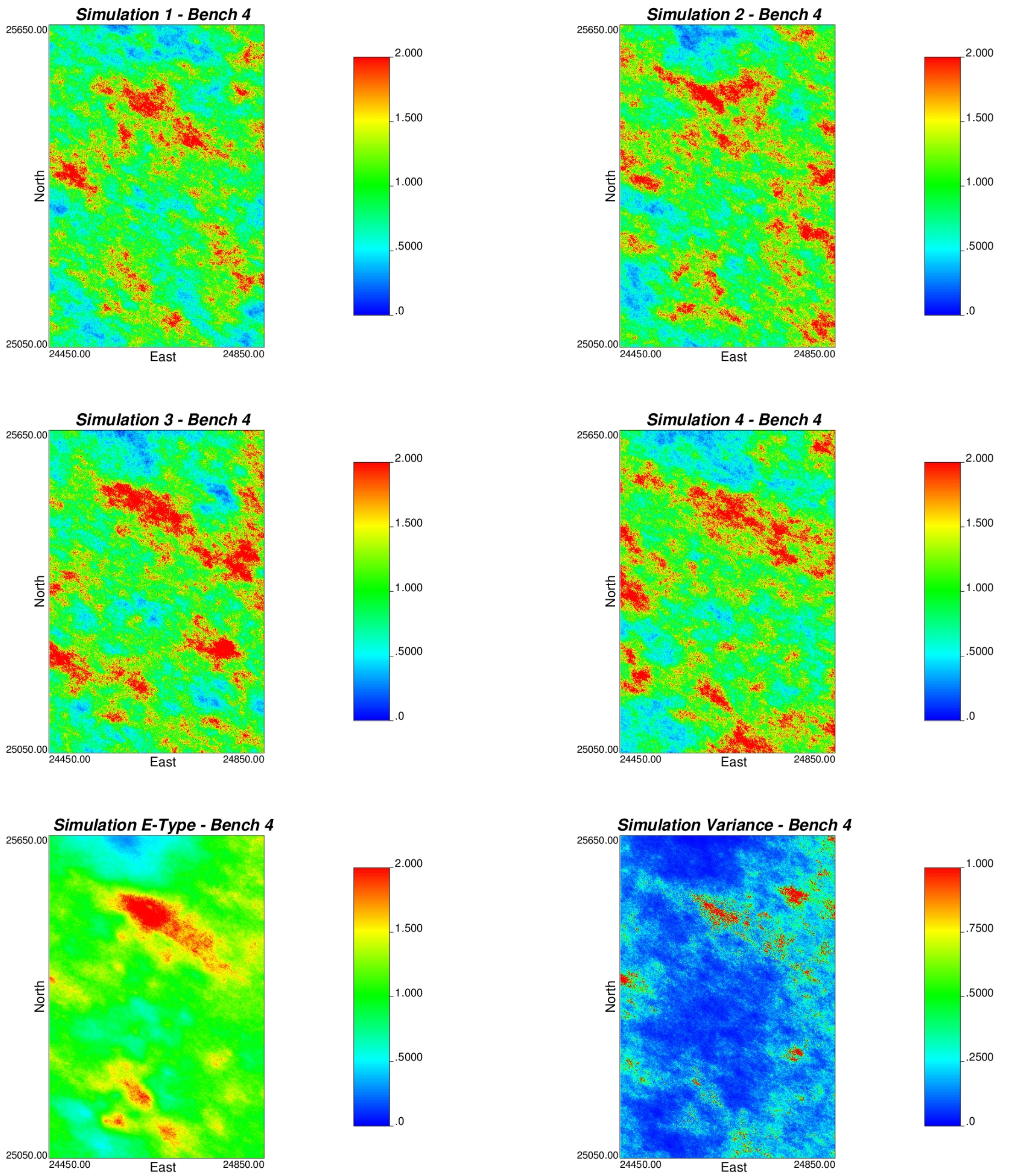

Figure 7: Four realizations, the e-type estimate and the conditional variance for grades in unit 200.

1.5 Merging the models

The final step is to merge the categorical simulations representing the extent of the domains, and the continuous simulations representing the uncertainty in the grade distribution.

This is done by taking a very simple approach: the simulated grade for realization l at location $\mathbf{u}$ is assigned based on the simulated category $s^l_{sim}(\mathbf{u})$ at that location in realization l of the set of categorical simulations. This means if category $s^l_{sim}(\mathbf{u}) = s_k$ then, the grade assigned at location $\mathbf{u}$ is that of the simulations for category s_k at that location, for that same realization.

The results of the assembled model are presented in **Figure 8**.

This set of realizations captures the joint uncertainty in the extent of the geological units and in the grades. It can be processed to represent the expected uncertainty of the true deposit, considering the samples available.

1.6 Final comments

In this example, we showed how to combine the categorical and continuous simulation models to account for the uncertainty in both the distribution of geological domains and the grade distribution within domains.

The resulting set of realizations can be used to determine the response of the deposit to any transfer function, by processing each realization through the transfer function and obtaining the statistical distribution of the response.

It should be noticed that by integrating the categorical and continuous models, the hard boundaries between categories, end up “diluted” in the set of models. This reflects the fact that we are uncertain about the location of the boundary between categories (although we may be sure the the boundary is hard).

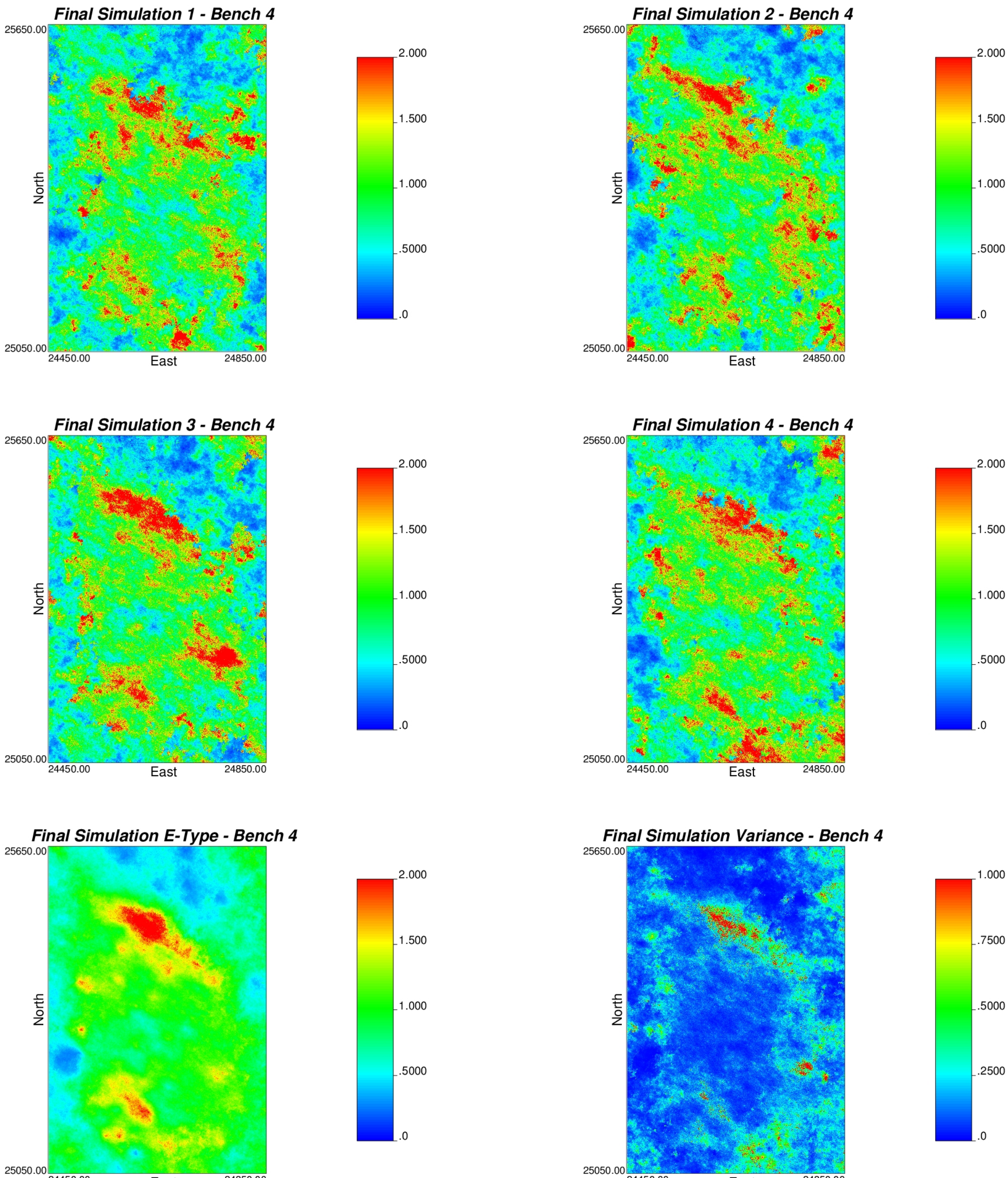

Figure 8: Four realizations, the e-type estimate and the conditional variance for the grades in the final assembled model.

Index