Geostatistics - Geostatistical simulation: basic concepts and sequential Gaussian simulation

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

September 21, 2024

Summary

MultiGaussian and indicator kriging allow determining the local uncertainty at every location. However, these methods do not permit quantifying uncertainty over a larger volume, or characterizing the expected variability in a sequence of points that will be transferred to another process. In this chapter, we introduce the concept of geostatistical simulation and explain why it is necessary and how it is different from estimation. We also introduce sequential Gaussian simulation, which is one of the most used algorithms for multiGaussian simulation, and is a natural extension from multiGaussian kriging.

1 Introduction

In many applications, the major objective has been, and still is, to obtain the "best" estimate of the variable studied. To achieve this objective, estimation methods have progressed from geometric triangulation and polygonal approaches to a variety of kriging algorithms. All of these estimation approaches produce a map of locally averaged values and in the case of kriging, a map of the estimation variance at each estimated location. The resultant models produced by these estimation methods have several limitations. These limitations include:

- The spatial variability and histogram of the estimates is "smoothed" compared to that known from sample data. This is known as the smoothing effect of kriging.
- The assessment of uncertainty (by means of the kriging variance) is strongly controlled by the sampling configuration without reference to the magnitude of the sample grades that inform the estimated value. This means that the kriging estimate is homoscedastic, that is, the variance does not depend on the value of the estimate. In practise, variables such as grades following a lognormal distribution show proportional effect, that is, the variability is proportional to the estimated grade. In mathematical terms this is called heteroscedasticity.
- It is very difficult to obtain a quantification of the joint uncertainty of a collection of blocks because the uncertainty in each block is not independent of the uncertainty in the adjoining block.

In contrast to the estimation methods, geostatistical simulation provides maps of the variable that:

- Honor the sample data values.
- Reproduce the histogram.
- Reproduce the spatial variability of the variable of interest.

In addition, simulation is a probabilistic procedure that results in **many different realizations of the same attribute**. Each realization honors sample values, geological interpretation, data statistics and spatial continuity.

Unlike estimation, where a best map can be generated under some definition of quality (for kriging it is the minimization of the mean square error), realizations are accepted or rejected based on their capacity to honor the data, geology, histogram, variogram, and any secondary information. Therefore, there is no single best realization, since they are all considered equally likely to occur. The set of equally probable realizations allows us obtain a distribution of the possible grade at that location. Similar procedures exist for categories such as the geological unit. This distribution of outcomes is interpreted as the uncertainty in the variable. This allows to report an uncertainty value that is conditioned by the surrounding sample values and geological attributes. The set of simulated values on a point scale also allows us to consider the uncertainty of a collection of related points or joint uncertainty. This means a change of support can be performed, and simulations at block support are obtained. The set of block support realizations represents the uncertainty of the variable at that support. For example, the net result of averaging the simulated

point grades within an arbitrary shape for each simulation is a distribution of the uncertainty of the expected grade for that volume (e.g. stope or production period volume). This cannot be obtained by kriging methods without assuming independence between each volume.

In summary, conditional simulation provides a quantification of the uncertainty surrounding an estimate. These values may account for the support and information effects, hence allowing quantification of uncertainty in any response variable. Response variables that depend upon several input variables can be correctly assessed with simulation and their uncertainty, quantified. This is the case of the response in mine planning, in geometallurgical studies or in geotechnical classification.

2 An introductory mining example

This introductory example is based on a satellite zone of mineralization at a gold mine in eastern Canada, that has potential for development as an underground operation to contribute additional tonnes and grade to the mill of the operation. Uncertainty about the continuity of the gold mineralization leads to uncertainty in predictions of cash flows.

Several hundred surface and underground diamond drill holes through the zone exist, confirming significant gold mineralization (**Figure 1**). High grade assays mostly fall along a few roughly planar structures that have been grouped into a series of lenses within a broader altered and mineralized shear system (**Figure 2**). All of these lenses follow the major foliation direction with an east-west strike and a

northerly dip of 65 to 70°.

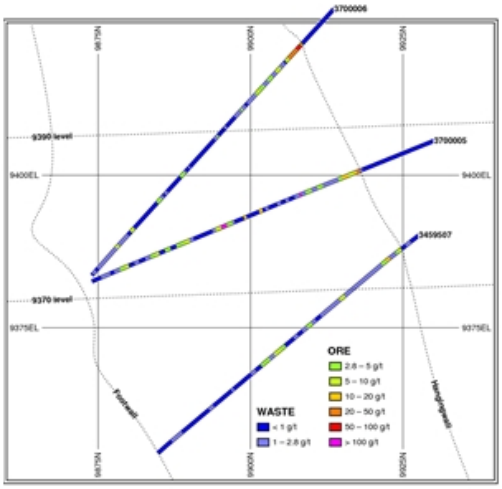

Figure 1: Drillhole intersections over a representative cross section of the deposit.

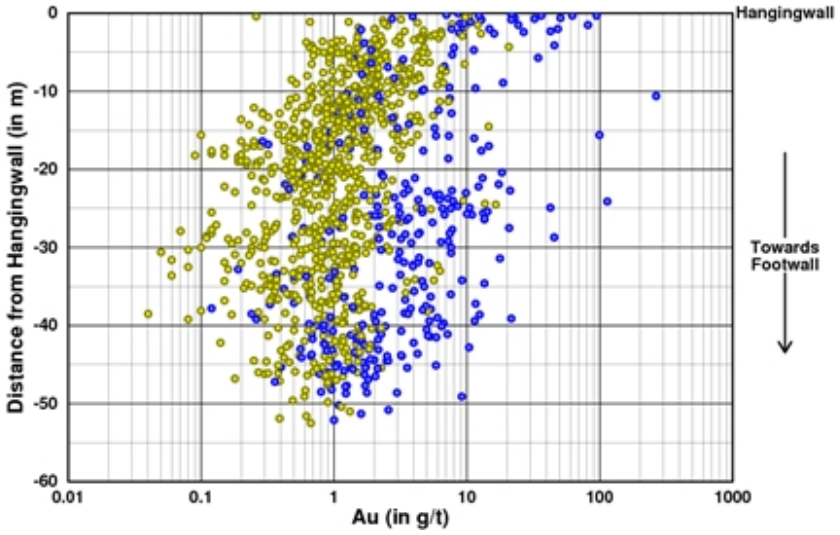

Figure 2: Grade trend as a function of the distance to the hanging wall.

Three realizations, out of a hundred created with a geostatistical simulation method, are shown in **Figure 3**. It can be seen that these realizations follow the general direction of continuity of the grades, honor the data, and show realistic variability of the grades at neighboring locations. Although they have the same general features, they are different. At every location, different values are simulated over the realizations. The spread of these values is a reflection on the uncertainty of the grade at that location. Closer to conditioning data, simulated values tend to have less dispersion, reflecting higher confidence.

For comparison, an inverse distance weighting estimate is depicted in **Figure 4**, which shows very different features. The continuity is honored (imposed by the search plan), but the map does not reflect any short range variability, and provides very hard transitions between high and low grade areas, due to the conditioning. A single map is obtained with this method, without access to the expected variability.

Considering the expected variability provided by the realizations from the geostatistical simulation method, the reserve grade within the production stopes is computed for each realization and fed into a cash flow analysis. This reflects the chance of realizing the plan. The expected cash flow from the inverse distance estimate is shown in **Figure 5**. A negative cash flow in the first period is compensated with positive cash flows in subsequent years. Now, if we look at the cash flows obtained for each realization, we see that these may fluctuate and there is some chance of not obtaining the expected value of the project (**Figure 6**). The net present value can be computed for every realization and compared with the expected NPV. Furthermore, if uncertainty is deemed too large, corrective actions can

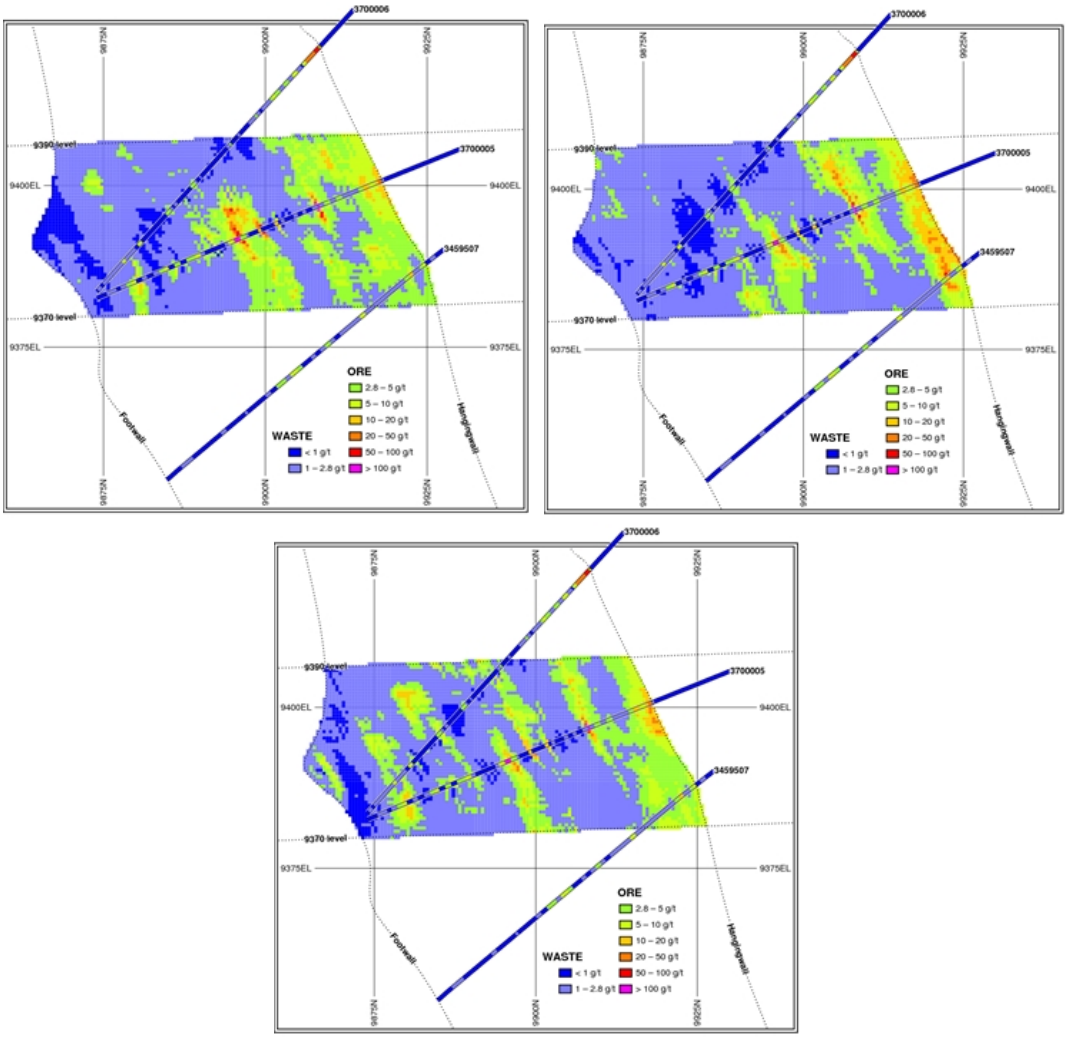

Figure 3: Three realizations obtained with a geostatistical simulation method.

be taken, such as additional drilling, or changes in the plan to minimize risk by starting with areas where confidence is higher and the fluctuation in the cash flow is lower. In this case, the NPV may be as low as $4.8 million and as high as $8.9 million (**Figure 7**). The decision maker should be

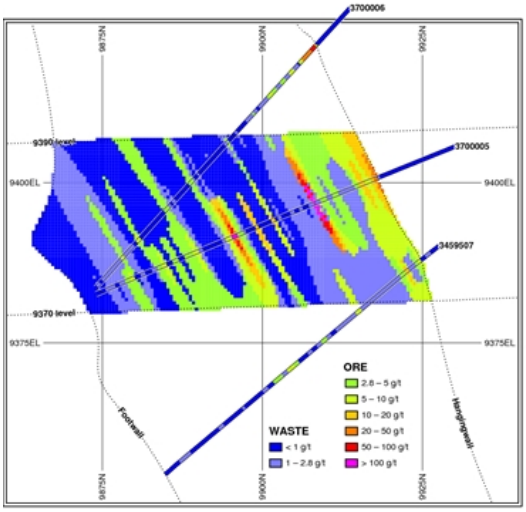

Figure 4: A map of inverse distance weighting estimates.

aware of this uncertainty to proceed with the plan.

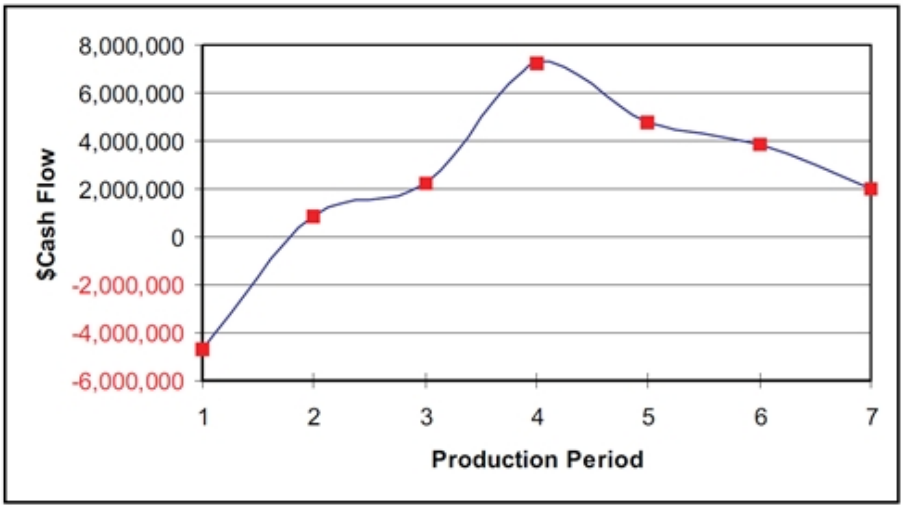

Figure 5: Expected cash flow obtained from the inverse distance estimates. This is usually the basis for decision making and planning.

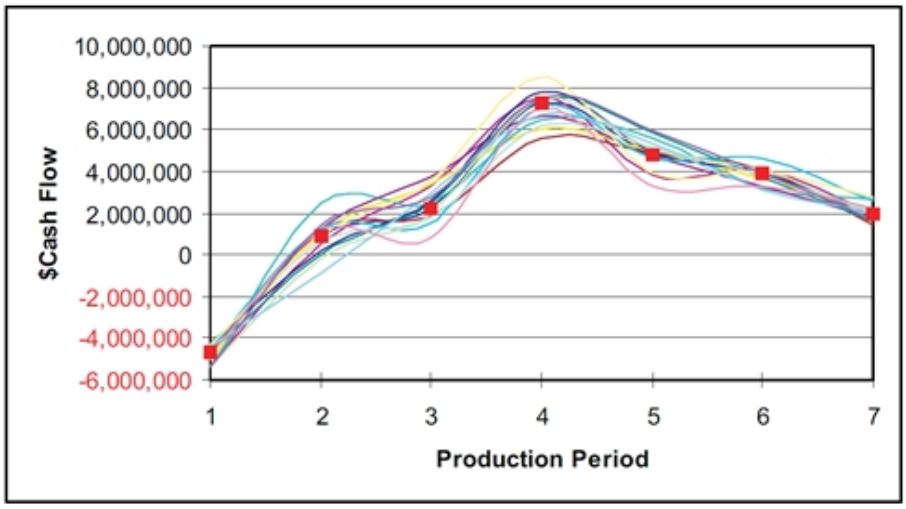

Figure 6: Cash flow uncertainty - only 10 realizations are shown. The expected cash flow from the inverse distance model is shown with red squares.

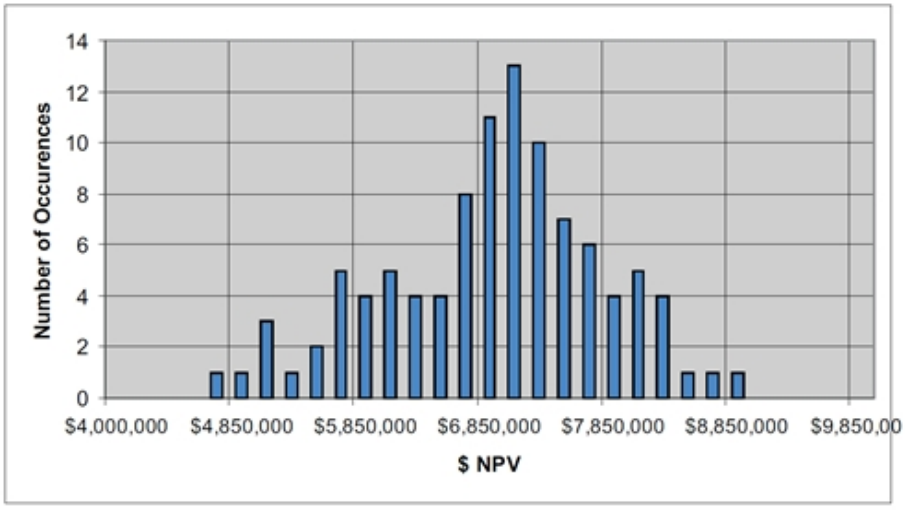

Figure 7: Net present value (NPV) uncertainty for the project.

The previous example is very straightforward, but illustrates that understanding the uncertainty in the estimate

may change the decisions made.

There are many difficulties that appear as a consequence of knowing the uncertainty. Communicating uncertainty and risk is not straightforward. Making decisions accounting for risk is also a very difficult task. In mining, designing the outline of an open pit or a stope accounting for the uncertainty on the grades is an extremely complex optimization problem. The same happens in other fields: imagine trying to decide what area needs to be decontaminated when you have a limited budget and 100 models displaying different possible spatial distributions of the pollutant.

> Notice that it is not the same: "to decide accounting for uncertainty" than "to assess the uncertainty given a fixed mine plan"

3 Geostatistical simulation

3.1 Principles

The main idea of geostatistical simulation is to **draw multiple realizations of a random function**. All the realizations will share the spatial structure, which is often imposed by the variogram, under the assumption that relationships between pairs of points (so called two-point statistics) are sufficient to capture the main features of the model.

Now, to make this useful, we create realizations from a random function that shares the spatial continuity features with our data. From the limited sample data available, we can infer a histogram and variogram (and sometimes, we

can go beyond and infer patterns or multiple-point statistics). This modeled three dimensional variogram characterizes the spatial continuity of the random function. In addition to this, we want the random function to honor the data, that is, we expect to recover the sample values at sample locations.

Realizations from a simulation method are called conditional, if they reproduce the sample values at sample locations. This means, that all realizations will share some common values. Since spatial correlation controls the dispersion of values in space, if one node is fixed with a known value, its surroundings will be correlated to that value. Therefore, around each sample location, the spread of values will be constrained, thus representing a lower uncertainty.

> The true distribution of the regionalized variable can be interpreted as one particular realization of the random function. Therefore, the simulated realizations can be interpreted as other possible scenarios.

Since we do not know exhaustively the true distribution, we can use each realization as one possible truth, and expect that it will behave similarly to the true distribution, since it shares its statistical and spatial statistics, and it reproduces the sample values at sample locations.

Therefore, **to assess the performance of the true variable when subject to a process, we can instead, assess the performance of each one of the realizations and use the output as a representation of the uncertainty in the response**.

We will now introduce the different methods and discuss the assumptions underlying each method.

3.2 Methods and assumptions

There are many approaches to simulation. The following is a non-exhaustive list of the most relevant approaches.

- MultiGaussian methods:
  - Sequential Gaussian simulation (continuous)
  - Turning bands simulation (continuous)
  - Matrix decomposition method, known as LU simulation (continuous)
- Indicator methods:
  - Sequential indicator simulation (categorical)
  - Sequential indicator simulation (continuous)
- Mixed models (continuous models used to obtain categorical methods):
  - Truncated Gaussian simulation (categorical)
  - PluriGaussian simulation (categorical)
- Multiple-point simulation:
  - Single Normal Equation (SNESIM) (categorical)
  - FILTERSIM (categorical and continuous)
  - Direct Sampling (categorical and continuous)
- Other methods
  - Simulated annealing (categorical and continuous)
  - Direct sequential simulation (continuous)
  - P-field simulation (continuous)

MultiGaussian methods rely obviously on a multiGaussian assumption. Indicator methods do not rely on a distributional assumption, which generates some problems of consistency. However, they show more flexibility and can be a good solution for some problems.

Truncated Gaussian and pluriGaussian simulation are categorical methods, based on simulating one or more continuous multiGaussian variables and then truncating the continuous result to categorize it into one or more classes. This allows obtaining a large spectrum of continuity models for the categorical variable. The main difficulties with these methods is conditioning a continuous simulation to categorical data and finding the spatial continuity with which the continuous fields must be simulated so that, after truncation, the continuity of the categorical variable is honored.

Multiple-point simulation methods go beyond the use of variograms. Instead, pattern statistics are inferred that account from multiple points at a time. This normally requires the use of a training image or of some kind of abundant and dense training data. These methods are very similar to methods used in texture synthesis in computer vision, therefore a wealth of methods have been researched in recent years and continue showing great potential for practical applications.

Finally, the methods classified under “other methods” are indirectly linked to a multiGaussian assumption (direct simulation and p-field simulation), or use an optimization-based approach to simulate (simulated annealing).

4 Sequential Gaussian simulation

4.1 Theoretical foundation

The sequential Gaussian simulation method is founded in the multiGaussian assumption of the variable. This can be interpreted as follows: the random function is composed of random variables $\{Z(\mathbf{u_i}), \forall \mathbf{u_i} \in D\}$ that form a multivariate Gaussian distribution. Another way of understanding this idea is to think of any set of R variables (each variable represents a random variable at a particular location $\mathbf{u_r}, \forall r = 1, ..., R$). The scatter plots between these variables are ellipsoids of R dimensions and form a joint multivariate Gaussian distribution. Therefore, if any of the locations has a known value (a sample, that simply means the variance of that random variable collapses to 0), then, the conditional multivariate distribution of the remaining $R-1$ locations can be inferred using the multiGaussian hypothesis. That is, the mean and variance can be obtained by regression (or, what is the same, by simple kriging), and the distribution remains an $R-1$-dimensional multiGaussian distribution.

Under this premise, if we have N locations to simulate, $\mathbf{u_1}, ..., \mathbf{u_N}$, we can model them as a standardized multiGaussian distribution. The joint pdf is:

$$f_{Y_1 ... Y_N}(y_1, ..., y_N) = \frac{exp\left(-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\right)}{\sqrt{(2\pi)^N |\boldsymbol{\Sigma}|}} \quad (1)$$

where Y_i is short notation for the random variable $Y(\mathbf{u_i})$.

This can be simplified, since we work with standardized variables.

$$f_{Y_1 \dots Y_N}(y_1, \dots, y_N) = \frac{exp\left(-\frac{1}{2}\boldsymbol{y}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{y}\right)}{\sqrt{(2\pi)^N |\boldsymbol{\Sigma}|}} \tag{2}$$

The correlation between two locations will be controlled by the correlogram at their separation distance **h**.

Thus,

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma^2_{Y_1} & C_{Y_1 Y_2} & \cdots & C_{Y_1 Y_N} \\ C_{Y_2 Y_1} & \sigma^2_{y_2} & \cdots & C_{Y_2 Y_N} \\ \vdots & \vdots & \ddots & \vdots \\ C_{Y_N Y_1} & C_{Y_N Y_2} & \cdots & \sigma^2_{Y_N} \end{pmatrix} = \begin{pmatrix} 1 & \rho_{Y_1 Y_2} & \cdots & \rho_{Y_1 Y_N} \\ \rho_{Y_2 Y_1} & 1 & \cdots & \rho_{Y_2 Y_N} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{Y_N Y_1} & \rho_{Y_N Y_2} & \cdots & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \rho(\mathbf{h_{12}}) & \cdots & \rho(\mathbf{h_{1N}}) \\ \rho(\mathbf{h_{21}}) & 1 & \cdots & \rho(\mathbf{h_{2N}}) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(\mathbf{h_{N1}}) & \rho(\mathbf{h_{N2}}) & \cdots & 1 \end{pmatrix}$$

The conditional distribution can be understood in terms of Bayes law (notice that we simplified the notation, to keep it short):

$$f_{Y_1 \dots Y_k | Y_{k+1} \dots Y_N} = \frac{f_{Y_1 \dots Y_N}}{f_{Y_{k+1} \dots Y_N}} \tag{3}$$

Now, if we look at a single conditioning variable,

$$f_{Y_1 \dots Y_{N-1} | Y_N} = \frac{f_{Y_1 \dots Y_N}}{f_{Y_N}} \tag{4}$$

This idea can be applied sequentially to simulate a random function:

- Start by imposing the conditioning data: $Y(\mathbf{u_1}), \dots, Y(\mathbf{u_n})$
- Simulate the first node $Y(\mathbf{u_{n+1}})$ conditioned to the available data
  - The conditional distribution of $Y(\mathbf{u_{n+1}})$ is a Gaussian distribution:

    $$f_{Y_{n+1} | Y_1 \dots Y_n} = \frac{f_{Y_1 \dots Y_{n+1}}}{f_{Y_1 \dots Y_n}}$$

    with parameters

    $$\begin{aligned} \boldsymbol{\mu_{1|2}} &= \boldsymbol{\mu_1} + \boldsymbol{\Sigma_{12}} \boldsymbol{\Sigma_{22}^{-1}} (\boldsymbol{y_2} - \boldsymbol{\mu_2}) \\ \boldsymbol{\Sigma_{11|2}} &= \boldsymbol{\Sigma_{11}} - \boldsymbol{\Sigma_{12}} \boldsymbol{\Sigma_{22}^{-1}} \boldsymbol{\Sigma_{21}} \end{aligned} \tag{5}$$

    where $\boldsymbol{Y_1}$ is the 1-dimensional vector composed of variable $Y(\mathbf{u_{n+1}})$ and $\boldsymbol{Y_2}$ is the n-dimensional vector composed of variables $Y(\mathbf{u_1}), \dots, Y(\mathbf{u_n})$.
  - Simulate a value from the distribution of $Y(\mathbf{u_{n+1}})$ by Monte-Carlo simulation.
  - Use the data to condition the subsequent nodes to be simulated.
- Simulate the subsequent nodes $Y(\mathbf{u_{n+k}})$ conditioned to the available data and previously simulated nodes

- The conditional distribution of $Y(\mathbf{u_{n+k}})$ is a Gaussian distribution:

$$f_{Y_{n+k}|Y_1 \ldots Y_n, Y_{n+1} \ldots Y_{n+k-1}} = \frac{f_{Y_1 \ldots Y_{n+k}}}{f_{Y_1 \ldots Y_n, Y_{n+1} \ldots Y_{n+k-1}}}$$

  with parameters as in **Equation 5**, now with $\boldsymbol{Y_1}$ being the 1-dimensional vector composed of variable $Y(\mathbf{u_{n+k}})$ and $\boldsymbol{Y_2}$ is the $n+k-1$-dimensional vector composed of variables $Y(\mathbf{u_1}), \ldots, Y(\mathbf{u_{n+k-1}})$.
- Simulate a value from the distribution of $Y(\mathbf{u_{n+k}})$ by Monte-Carlo simulation.
- Use the data to condition the subsequent nodes to be simulated until the last node is simulated.

4.2 Steps

We have already covered most of the basic steps for sequential Gaussian simulation, when presenting multiGaussian kriging. For simulation, a multiGaussian kriging is performed at every location, providing the mean and variance of the conditional distribution, from which a simulated value is drawn. However, one additional consideration is required to impose the spatial correlation between simulated points: **previously simulated values are used to condition subsequent simulated values, to ensure their spatial correlation is honored**.

The implementation proceeds sequentially. At every node, the sample values within a neighborhood as well as the previously simulated nodes in the neighborhood, are used to compute the kriging estimate and variance (in Gaussian transformed units). To avoid artifacts due to the use of a

regular grid of simulation points (and the screening effect of kriging), a random path is created to visit the nodes in a random order in each realization.

From the kriging estimate and variance of normal scores, a simulated value can be computed by Monte Carlo simulation, since under the multiGaussian assumption, we know that the shape of the conditional distribution is Gaussian and its mean and variance are the simple kriging estimate and variance.

The steps to implement sequential Gaussian simulation are:

1. **Representative distribution**: decluster the data to obtain a representative distribution.
2. **Normal score transform**: transform the sample data according to the transformation between the representative (declustered) distribution and a standard normal distribution, to obtain the normal scores of the data.
3. **Variogram of normal scores**: compute the experimental variogram of normal scores and fit a model.
4. **Verify the multiGaussian assumption**: there are many approaches to check if the multiGaussian assumption is reasonable. They include: checking the h-scatter plots, comparing variograms of a truncated multiGaussian variable with the experimental indicator variograms of the continuous variable, computing variograms of different order and checking some ratios, etc.
5. **Perform the simulation**:

    (a) **Visit a node**: following the random path visit a node and check that it is not informed by a sample.

(b) **Search for data and previously simulated points in neighborhood**: find the nearby normally transformed sample data and any previously simulated node in the search neighborhood.

(c) **Perform simple kriging**: compute the simple kriging estimate and simple kriging variance of the normally transformed values.

(d) **Simulate a value from the conditional distribution**: Using a Gaussian distribution with mean equal to the simple kriging estimate and variance equal to the simple kriging variance, simulate a value by Monte Carlo simulation. This entails drawing a uniform random value in the interval $(0, 1)$ and computing the inverse of the cdf of the following distribution: $\mathcal{N}\left(y^*_{SK}(\mathbf{u_0}), \sigma^2_{y,SK}(\mathbf{u_0})\right)$. Notice that the uniform random number drawn represents the cumulative probability, and the dimulated value is the corresponding quantile.

6. **Back transform the simulated values**: using the representative distribution and its relationship with the standard Gaussian distribution, back-transform every simulated node to the original units.

It is interesting (although we are not going to show it here) that the collection of simulated values obtained at the end of step 5, follows a standard Gaussian distribution. Therefore, the back-transformation in step 6, brings back the representative distribution in original units over the domain.

The result of the simulation process is a set of multiple realizations, where each realization reflects the variability

expected in space of the variable, and the set of realizations summarize the uncertainty at every location.

Notice that this is a point-support result. The spatial variability and the uncertainty at every location represents points (or to be exact, the support of the samples used to condition the simulation and to infer the variogram). No change of support has been done until now.

4.3 Post-processing

The simulation results at point support can be treated as equally probable scenarios of the true distribution of the variable over the domain. None of them is a good representation by itself, and each will perform poorly if it is used as a predictor. However, the set will give valuable information.

In most applications, it is not the point uncertainty what is relevant, but the joint uncertainty when considering a larger volume. In fact, if the goal were to obtain the point uncertainty, multiGaussian kriging or multiple indicator kriging would be appropriate approaches. When considering many points forming a volume, however, we need to understand their joint uncertainty, which accounts for the spatial correlation between points.

Every point-support realization can be seen as one possible scenario (that already captures the spatial correlation between points). If we apply an operation or a transfer function over the simulated domain, we can emulate the behavior of the true (but unknown) distribution.

For example, we may be interested in knowing the uncertainty of the grade at a mine, over a production period. The

mine plan will state which volume is extracted. Each realization can be processed to obtain the average grade over that production period. Each one will give a slightly different result. Pooling together the results of a large number of realizations, the distribution of "averages over the production period" can be obtained. The spread reflects the uncertainty and the mean represents the expected grade of the volume. Notice that this expected grade is probably not the best estimation of the true grade, since simple kriging was used. However, the approach let us infer the variability, which was not possible with kriging (at least not accounting for the proportional effect).

This same approach can be applied to any transfer function, not just a change of support.

4.4 Implementation details

In practice, there are several tricks that are convenient or are needed to implement sequential Gaussian simulation.

Local neighborhood

First, the sequential approach builds conditional distributions with an increasing amount of conditioning information. This happens because, for the first simulation location, only the n samples (the hard data) will be used to determine the conditional expectation and conditional variance of the distribution at the node location (through simple kriging of the normal scores). However, the simulated value drawn from the Gaussian distribution with the parameters inferred by kriging, is used as conditioning for all subsequent nodes.

So, for the second node to be simulated the conditioning information will consider the n hard data and the 1 previously simulated node. Similarly, for the node k in the sequence, all $k-1$ previous nodes already have a value that will be used to condition this location. Plus, we have the n hard conditioning samples. The kriging system will have $n+k-1$ equations.

It is clear that, as the sequential simulation progresses, we accumulate a potentially very large number of previously simulated nodes (a simulation grid may easily have millions of nodes). Therefore, it becomes impractical to use *all* the conditioning information. We therefore use a local neighborhood, and impose some constraints in the number of samples and previously simulated nodes that are used. Since simulation relies more heavily on stationarity, it is reasonable to use as much conditioning information as possible, as inference of the conditional mean and variance of the distribution improves as more information is used. On the other hand, since solving the kriging system is computationally costly and this cost increases exponentially with the number of equations, we would prefer to keep this number low, or what is the same, the number of conditioning data small.

Notice that conditioning to previously simulated nodes imposes the correlation between simulated locations, to capture the spatial variability (from the variogram).

Two searches: for data and for previously simulated nodes

Simulation is done, in most cases, over a regular grid. This simplifies the search for previously simulated nodes, since only a limited number of nodes around the simulation node can be checked to find the neighboring information. The sample data, on the other hand, are not regularly spaced and in most cases the sample locations will not coincide with any node in the grid. Therefore, to search for sample data, the search needs to compute distances between the samples and the simulation node, sort and determine which ones are within the local neighborhood.

There are many search algorithms to improve the efficiency of the search, but they are all quite heavy computationally, especially considering that this needs to be done millions of times (it is done once at every node and for every realization). A super-block search or a kD-tree can be used to improve the search, but it is still a burden for the computation of kriging.

If the simulation grid is dense enough, sometimes the sample data are assigned to the nodes on the grid. The assumption is that by displacing the samples to the closest node, there is little loss of accuracy in the final model. This is reasonable if the nodes are closely spaced with respect to sparse samples. In this case, samples are easier to find within the neighborhood, as only the nodes around the simulated node are to be searched. Samples are, in this case, treated as previously simulated nodes.

Having to do two searches is computationally costly, therefore assigning data to nodes is a good approach to save simulation time.

Random path and multiple grids

Nowhere in the theoretical formulation of sequential simulation is a requirement for the nodes to be randomly visited. In theory, one could visit the nodes of the grid in order following a regular path and condition to the samples and previously simulated nodes (within the search neighborhood). However, kriging may generate some artifacts due to the screening effect. Therefore, a random path is used to avoid these artifacts.

The random path is, however, not completely random, since another trick is used to ensure long range continuity is reproduced. It is easy to imagine that as simulation progresses, more and more informed nodes become available to condition subsequent simulation nodes. Therefore, in the search neighborhood, and considering that a maximum number of previously simulated nodes is used to limit the size of the kriging system, only the closest previously simulated nodes will be used for kriging. This implies that the continuity will be imposed for short distances, but the long range continuity may not be properly captured. The multiple grids approach considers simulating first a coarse grid of points and then refining the grid as simulation progresses, to ensure that at early stages of the simulation, the long range continuity is locked by the simulated points over the coarse grid. Within each grid (coarser to finer), the nodes are visited randomly. In summary, by using multiple grids, the path is structured from a coarse grid to a fine grid, but nodes are visited randomly within each level.

Simple or ordinary kriging?

We know that simple kriging relies heavily on the assumption of stationarity of the mean. This makes it inflexible to local changes in the mean of the variable. Could we use ordinary kriging to make it more robust to local changes? Unfortunately, the answer is: no. Simple kriging identifies the conditional expectation and conditional variance in the case of a multiGaussian variable. If we were to use ordinary kriging, we would not be estimating properly these two parameters (conditional mean and conditional variance). In fact, the kriging variance in ordinary kriging is always higher than that of simple kriging. Therefore, the conditional Gaussian distributions obtained will show more variability than their "true" variability. When drawing values from that distribution by Monte Carlo simulation, we will be overestimating the variability. This translates in a variance inflation in the global result of the simulation and, as a consequence, the global histogram is not reproduced and the model is biased.

Computational issues

Simulations are heavy in computational requirements. Depending on the implementation, the full array of simulated values needs to be kept in RAM memory, so when simulation grids are large, this becomes a problem. In terms of CPU, they are also very demanding and therefore, they may take some time to be completed. At each node a search, a kriging and a simulation is done. This is repeated for all the nodes and repeated again for multiple simulations. We will not dwell about these details in these notes.

Index