

One Step at a Time: The Origins of Sequential Simulation and Beyond

J. Jaime Gómez-Hernández · R. Mohan
Srivastava

Received: date / Accepted: date

Abstract In the mid-1980s, still in his young 40s, André Journel was already recognized as one of the giants of geostatistics. Many of the contributions from his new research program at Stanford University had centered around the indicator methods that he developed: indicator kriging (IK) and multiple indicator kriging (MIK). But when his second crop of graduate students arrived at Stanford, indicator methods still lacked an approach to conditional simulation that was not tainted by what André called the ‘Gaussian disease’; early indicator simulations went through the tortuous path of converting all indicators to Gaussian variables, running a turning bands simulation, and truncating the resulting multi-Gaussian realizations. When he conceived of sequential indicator simulation (SIS), even André likely did not recognize the generality of an approach to simulation that tackled the simulation task one step at a time. The early enthusiasm for SIS was its ability, in its multiple-indicator form, to cure the Gaussian disease and to build realizations in which spatial continuity did not deteriorate in the extreme values. Much of Stanford’s work in the 1980s focused on petroleum geostatistics, where extreme values (the high-permeability fracture zones and the low-permeability shale barriers) have much stronger anisotropy, and much longer ranges of correlation in the maximum continuity direction, than mid-range values. With multi-Gaussian simulations necessarily imparting weaker continuity to the extremes, SIS was an important breakthrough. The generality of the sequential approach was soon recognized, first through its analogy with multi-variate unconditional simulation achieved

J. Jaime Gómez-Hernández
Institute of Water and Environmental Engineering
Universitat Politècnica de València
Camino de Vera s/n
46022 Valencia (Spain)
E-mail: jgomez@upv.es

R. Mohan Srivastava
Tristar Gold Inc., Canada

using the lower triangular matrix of an LU decomposition of the covariance matrix as the multiplier of random normal deviates. Modifying LU simulation so that it became conditional gave rise to sequential Gaussian simulation (SGS), an algorithm that shared much in common with SIS. With nagging implementation details like the sequential path and the search neighborhood being common to both methods, improvements in either SIS or SGS often became improvements to the other. Almost half of the contributors to this Special Issue became students of André in the classes of 1984-1988, and several are the pioneers of SIS and SGS. Others who studied later with André, explored and developed the first multipoint statistics (MPS) simulation procedures, which are based on the same concept that underlies sequential simulation. Among his many significant intellectual accomplishments, one of the cornerstones of André Journel's legacy was sequential simulation, built one step at a time.

Keywords random functions · large grids · stochastic processes

1 Introduction

This paper presents a review of the evolution of conditional stochastic simulation of random fields, from the early works by Journel (1974) until the most recent multiple-point and pattern-based simulations, with sequential simulation acting as the conductor. After its introduction in the late 80s of the past century, sequential simulation has become the cornerstone of many or most of the simulation algorithms of today. This paper, as the rest of the papers in this issue, is a tribute to André Journel and his ingenuity and it is largely biased towards geostatistics. It happens that the two co-authors were classmates and wrote the first code of sequential indicator simulation that was made publicly available under the clear understanding by André of the importance of public domain code for its widespread use and the advancement of research and development.

2 Before

Before simulation, there was estimation, and kriging is the paradigmatic estimation method in geostatistics.

Consider a random function $\{Z(\mathbf{u}), \mathbf{u} \in \mathcal{D}\}$, characterized by all the n -variate distribution functions for any n -tuple of points within \mathcal{D}

$$F(Z(\mathbf{u}_1), Z(\mathbf{u}_2), \dots, Z(\mathbf{u}_n)), \forall n | (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) \in \mathcal{D}. \quad (1)$$

This definition is the classical one, but a random function can also be defined as the rule that assigns a realization $z(\mathbf{u}, \theta)$ to the outcome θ of an experiment

$$Z(\mathbf{u}) \sim \{z(\mathbf{u}, \theta)\}, \forall \theta \in \mathcal{S}, \mathbf{u} \in \mathcal{D}. \quad (2)$$

The interest of this definition is that the random function can be seen as a collection of realizations, and the n -variate distribution functions for any n -tuple of points within \mathcal{D} could be obtained from these realizations.

With the concept of a random function as a collection of realizations, it is simple to distinguish an unconditional random function (the one given by Eq. (1) or (2)) from a random function conditional to a set of n_o observations $\{z(\mathbf{u}_1), z(\mathbf{u}_2), \dots, z(\mathbf{u}_{n_o}); (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{n_o}) \in \mathcal{D}\}$. The latter is made up by the subset of outcomes of the experiment $\theta \in \mathcal{S}$, that is, by the subset of realizations in Eq. (2) that match observed values at observation locations, and it will be represented by $\{\mathbf{Z}(\mathbf{u})|(n_o), \mathbf{u} \in \mathcal{D}\}$.

A conditional estimate, $z^*(\mathbf{u})$, can be obtained, for instance, as the expected value of all the conditional realizations

$$z^*(\mathbf{u}) = E\{\mathbf{Z}(\mathbf{u})|(n_o)\}. \quad (3)$$

This is what kriging does, after imposing certain conditions to the random function such as stationarity of second order. But kriging, which has proven quite valuable in mining, is a locally optimal estimate (Journel, 1989); each estimated value is the best estimate (in a least-squares sense) of the variable at a location given the conditioning data, but when considered pairwise, or in groups of three or more, the optimality is lost, and kriging maps are characterized by a variance much smaller than that of the underlying random function, and also a smooth spatial variability that cannot be observed in any of the realizations. For these reasons, an estimated map should never be used to evaluate the state of a system for which the short scale variability is of paramount importance, such as is the case for the transport of a solute in an aquifer, where the short scale variability of hydraulic conductivity will determine the movement of the solute plume in the aquifer and will never match the movement of the plume in a smooth estimated map (Gómez-Hernández and Wen, 1994). In these cases, there is a need to evaluate the system performance on the conditional realizations, and then, from the outcomes computed on these realizations, retrieve an estimate (as the expected value, for instance). The travel time for a solute to go from A to B estimated in an aquifer with conductivities obtained by kriging will be a very poor estimate of the actual travel time, which would be better estimated by taking the average of the travel times computed on each one of the conditional realizations.

The question is then how to generate realizations from the random function.

2.1 Unconditional realizations

The generation of unconditional realizations was addressed as early as Matérn (1960); but, until the spread of numerical computation, it went unnoticed. Three main approaches could be mentioned for the generation of unconditional realizations before the advent of sequential simulation: spectral methods, matrix factorization and turning bands.

The spectral methods, lead by the work by Shinozuka and Jan (1972), are based on the spectral decomposition of the covariance matrix. Once the spectrum is known, finite ranges of frequencies along each dimension are determined containing as close as possible to 100% of the energy, then a realization of the random field is generated according to the expression

$$z(\mathbf{u}) = \sqrt{2} \sum_{k_1=1}^{N_1} \cdots \sum_{k_d=1}^{N_d} A(\omega_{\mathbf{k}}) \cos(\omega'_{\mathbf{k}} \cdot \mathbf{u} + \phi_{\mathbf{k}}), \quad (4)$$

where, d is the space dimension, N_1, \dots, N_d are the number of frequency increments in which each axis is discretized, $\mathbf{k} = \{k_1, \dots, k_d\}$, $\phi_{\mathbf{k}}$ is an independent phase randomly distributed between 0 and 2π , $\omega'_{\mathbf{k}} = \omega_{\mathbf{k}} + \delta\omega$, with $\delta\omega$ being a small perturbation to avoid periodicity, and

$$A(\omega_{\mathbf{k}}) = (S_0(\omega_{\mathbf{k}}) \Delta\omega_1 \dots \Delta\omega_d)^{\frac{1}{2}}, \quad (5)$$

where S_0 is the spectrum, and $\Delta\omega_1 \dots \Delta\omega_d$ are the increments in which the frequency axes have been discretized.

Of the matrix factorization methods, the best known is the Cholesky decomposition of the covariance matrix (Davis, 1987), whereby any self-adjoint positive definite matrix can be decomposed into the product of a lower triangular matrix by its transpose, $C = L \cdot L^T$. When the matrix C represents the covariances of any pair of points on a given set, then it can be shown that

$$z(\mathbf{u}) = L \cdot \phi \quad (6)$$

results in a realization of the random function with covariance C , where ϕ is a vector of random uncorrelated values drawn from a distribution with zero mean and variance of one.

The turning bands method, introduced by Matheron (1973) and described in Journel and Huijbregts (1978) uses a clever idea that made it, at the time, the only one capable of generating realizations over very large domains in two and three dimensions. In the turning bands method, the multidimensional simulation is replaced by a (small) number of one-dimensional simulations along lines oriented in different directions. The orientation of the lines is chosen trying to keep a homogeneous density in the circle or in the sphere. The simulated values are obtained by projecting each point \mathbf{u} onto the lines and linearly combining the values from the lines. A realization is obtained as

$$z(\mathbf{u}) = \frac{1}{\sqrt{N}} \sum_{k=1}^N z_{1,k}(\langle \mathbf{u}, \theta_k \rangle), \quad (7)$$

where N is the number of lines, $z_{1,k}$ is a one-dimensional realization along line k that has been generated with a covariance that is derived from the covariance of the random function, and $\langle \mathbf{u}, \theta_k \rangle$ represents the projection of vector \mathbf{u} onto line k .

2.2 Conditional realizations

Unconditional realizations are attractive for theoretical studies, but the practitioner needs realizations that honor the data at data locations. The generation of conditional realizations was introduced by Journel (1974) using an indirect approach that combines unconditional simulations and kriging. It can be shown that the deviations of a conditional realization from the map obtained by kriging are independent of the values at data locations. This property permits the computation of these deviations from an arbitrary unconditional realization simply by kriging the values from the same locations at which the data are sampled. Then, the deviations on the unconditional realization are added to the kriging map obtained with the observed data. The procedure to generate a conditional realization can be summarized with the following algorithm:

1. Generate an unconditional realization $z_{unc}(\mathbf{u})$,
2. Sample the unconditional realization at data locations and, with the sample values as data, perform kriging $z_{unc}^*(\mathbf{u})$,
3. Subtract both fields to obtain the residual field $r(\mathbf{u}) = z_{unc}(\mathbf{u}) - z_{unc}^*(\mathbf{u})$,
4. Perform kriging with the observed values $z^*(\mathbf{u})$,
5. Add the residual field to the kriging map, $z(\mathbf{u}) = z^*(\mathbf{u}) + r(\mathbf{u})$.

As an alternative approach to conditioning, the matrix factorization method was extended by Davis (1987) and Alabert (1987) for the generation of conditional realizations using a smart partitioning of the covariance matrix. Consider a point set split into two subsets, subset 1 contains all points where data are located, and subset 2 all points where the conditional realization has to be generated. Davis (1987) proves that

$$z(\mathbf{u}_2) = L_{21}L_{11}^{-1}z(\mathbf{u}_1) + L_{22}\phi \quad (8)$$

results in a conditional realization, where \mathbf{u}_1 are the points where data were observed, \mathbf{u}_2 are the points where the realization is generated, ϕ is a vector of random uncorrelated values drawn from a distribution with zero mean and variance one, and L_{11} , L_{21} and L_{22} are the submatrices partitioning the lower triangular matrix of the covariance factorization

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} L_{11}^T & L_{21}^T \\ 0 & L_{22}^T \end{bmatrix} \quad (9)$$

where the first rows and columns are associated to point subset 1 and the last rows and columns to the point subset 2.

3 Sequential simulation

In the 1980s, while the new framework of indicator geostatistics (Journel, 1983) was taking shape, the interest in generating realizations from a random function built on the idea of transforming a continuous variable into a collection of indicator variables developed. The indicator formalism is quite simple, any

random function $\{\mathbf{Z}(\mathbf{u}), \mathbf{u} \in \mathcal{D}\}$ can be transformed into a set of K indicator random functions, $\{\mathbf{I}_k(\mathbf{u}), k = 1, \dots, K\}$ corresponding to K thresholds z_k according to the following expression

$$\mathbf{I}_k(\mathbf{u}) = \begin{cases} 1, & \text{if } \mathbf{Z}(\mathbf{u}) \leq z_k \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

Drawing a realization $z(\mathbf{u})$ would be equivalent to drawing K realizations $\{i_k(\mathbf{u}), k = 1, \dots, K\}$ and transforming them into $z(\mathbf{u})$.

The issue of how to generate a realization from an indicator random function was solved by Journel and Isaaks (1984), who use the truncation of a Gaussian realization to generate the indicator realization. In their paper, they explain how to choose the truncation value z_k , and how to determine the covariance function of the Gaussian random function that will yield the desired covariance for the indicator realization. They also describe the conditioning process, which is based on the technique described by Journel (1974) and recapped at the beginning of subsection 2.2.

But the extension of Journel and Isaaks (1984) approach to generate multiple indicator realizations and then transform them into a realization $z(\mathbf{u})$ never took place. There was a fundamental need to devise a new conditional simulation algorithm that could overcome the problems of the existing ones, namely,

1. The difficulty of generating very large conditional realizations,
2. The difficulty of generating non-Gaussian realizations.

Of all the unconditional methods described above, only the turning bands method coupled with the conditioning approach by Journel (1974) could generate very large conditional realizations; the problem with this method is that it needed many bands to produce realizations that did not display artificial continuities in the directions of the lines on which the one-dimensional random functions are generated, and that the linear combination of random values from many lines necessarily results in realizations of a multivariate Gaussian random function.

It was in the Fall of 1988 during one of the Topics in Advance Geostatistics classes at Stanford University when André presented the seminal idea of indicator sequential simulation, a revolutionary concept in stochastic simulation that would allow the generation of very large conditional realizations of non-Gaussian random functions. We were a handful of students in the class, one of them was Mohan Srivastava, who, over the next weekend, wrote a prototype in FORTRAN of the simulation algorithm. The following week, he made a presentation of the problems he found for the implementation and showed some of the first successful realizations; then, Jaime Gómez-Hernández picked it up from there to write an ANSI-C code, ISIM3D, who was eventually published (Gómez-Hernández and Srivastava, 1990) and became the first publicly available sequential simulation code. The sequential simulation algorithm was born to address the problem of generating conditional realizations of indicator-based random functions, but it will quickly be evident that it could be used

for the generation of realizations of many other random functions, including the multiGaussian one.

The idea behind sequential simulation is quite simple: any multidimensional random function distribution can be rewritten as the product of univariate distributions by recursively applying the definition of conditional probability. Indeed, Eq. (1) can be rewritten as

$$\begin{aligned} F(Z(\mathbf{u}_1), Z(\mathbf{u}_2), \dots, Z(\mathbf{u}_n)) &= F(Z(\mathbf{u}_1)) \cdot \\ &F(Z(\mathbf{u}_2)|Z(\mathbf{u}_1)) \cdot \\ &F(Z(\mathbf{u}_3)|Z(\mathbf{u}_1), Z(\mathbf{u}_2)) \cdots \\ &F(Z(\mathbf{u}_n)|Z(\mathbf{u}_1), \dots, Z(\mathbf{u}_{n-1})), \end{aligned} \quad (11)$$

where each component is the conditional distribution of a random variable given all the random variables that appear before in the product. The problem of drawing a realization from the multivariate distribution (1) is replaced by the problem of drawing n values from univariate conditional distributions. This transformation was already proposed by Rosenblatt (1952) and mentioned by Johnson (1987). The extension to generate realizations conditional to a set of (n_o) data is trivial

$$\begin{aligned} F(Z(\mathbf{u}_1), Z(\mathbf{u}_2), \dots, Z(\mathbf{u}_n)|(n_o)) &= F(Z(\mathbf{u}_1)|(n_o)) \cdot \\ &F(Z(\mathbf{u}_2)|(n_o), Z(\mathbf{u}_1)) \cdot \\ &F(Z(\mathbf{u}_3)|(n_o), Z(\mathbf{u}_1), Z(\mathbf{u}_2)) \cdots \\ &F(Z(\mathbf{u}_n)|(n_o), Z(\mathbf{u}_1), \dots, Z(\mathbf{u}_{n-1})), \end{aligned} \quad (12)$$

there is only a need to replace the previous conditional distributions by distributions that are also conditional to the data.

The sequential simulation algorithm for the generation of a realization of a random function over a point set of size N conditional to n_o observation data would be as follows

1. Define a permutation of the numbers 1 to N that will identify the sequence in which the conditional univariate distributions will be built,
2. Sequentially visit all nodes according to the previous ordering and compute, at each node i , the conditional distribution of variable $Z(\mathbf{u}_i)$ given the (n_o) data and all previously simulated random variables $\{Z(\mathbf{u}_1) = z_1, Z(\mathbf{u}_2) = z_2, \dots, Z(\mathbf{u}_{i-1}) = z_{i-1}\}$

$$F(Z(\mathbf{u}_i)|(n_o), Z(\mathbf{u}_1) = z_1, Z(\mathbf{u}_2) = z_2, \dots, Z(\mathbf{u}_{i-1}) = z_{i-1}),$$

3. Draw a value z_i from this distribution and incorporate it to the conditioning data set for the simulation of the next node,
4. Go back to step 2 until all nodes have been simulated.

The algorithm is ready to implement if there is a way to compute the conditional distribution of any of the random variables $Z(\mathbf{u}_i)$ given any number of conditioning random variables.

The reason why sequential simulation was first developed for the simulation of realizations from indicator-based random functions is because the introduction of non-parametric geostatistics (Journel, 1982, 1983) was aimed, precisely, at the computation of the conditional distribution of a random variable given a number of data using a non-parametric indicator-based approach. Later, it was realized that sequential simulation could be extended to the simulation of multiGaussian random functions, since the univariate conditional distributions, in this case, are Gaussian with mean and variance given by the solution of a set of normal equations (Anderson, 1984).

3.1 Implementation issues

While the theory was simple and the core of the algorithm was straightforward, some implementation issues had to be addressed.

3.1.1 Size of conditioning data set

Whether an indicator-based random function or a multiGaussian was chosen, the calculation of the conditional distributions required the solution of a set of kriging equations: indicator kriging for the non-Gaussian random function and simple kriging for the Gaussian one. The kriging system has as many equations as number of conditioning data. Therefore, it seems impossible to strictly apply the decomposition (12) when n is large. The size of the kriging systems must be limited to a maximum value for the algorithm to be practical for the generation of realizations over hundreds of thousands or millions of points.

Following a criterion similar to the standard implementation of kriging, the system of equations is limited to the closest conditioning values within a search neighborhood up to a predetermined maximum number.

All conditional distributions in (12) are approximated as

$$\begin{aligned} F(Z(\mathbf{u}_i)|(n_o), Z(\mathbf{u}_1) = z_1, Z(\mathbf{u}_2) = z_2, \dots, Z(\mathbf{u}_{i-1}) = z_{i-1}) \approx \\ F(Z(\mathbf{u}_i)|Z(\mathbf{v}_1) = z_1, Z(\mathbf{v}_2) = z_2, \dots, Z(\mathbf{v}_{n_{max}}) = z_{n_{max}})), \end{aligned} \quad (13)$$

where $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{n_{max}}\}$ is a location subset selected out of the set of locations made up by the n_o conditioning data and the $i - 1$ previously simulated points that are within the search neighborhood centered at the location for which the conditional distribution is being estimated.

3.1.2 Random path

Once the decision of limiting the number of conditioning points is taken, the decomposition of the multivariate random function is no longer exact. This approximation raises the concern of whether the order in which the nodes are visited has an adverse effect on important spatial statistical properties of

the realizations. In the first implementations, a random path was used, and this is the common approach used in most codes. The way to generate such a random path is not trivial, the first approach proposed by Gómez-Hernández and Srivastava (1990) was to use a congruential generator (Bratley et al., 1983) to produce all numbers between 1 and a sufficiently large power of 2 to cover all nodes. The problem with this approach is that subsequent realizations always use exactly the same path changing only the starting point. A better approach is to generate a random permutation of all integers between 1 and N , what can be achieved with N random swaps in an ordered sequence (Durstensfeld, 1964).

The decision to implement a random path was motivated largely by the observation of artificial stripes in the realizations when the sequence moved from one node to an adjacent neighbour, along the rows or columns of the grid. While immediately obvious, and usually undesirable, this was an artifact that cannot be identified or corrected solely through two-point statistics; it is a higher-order artifact in a simulation procedure governed by two-point statistics.

Liu and Journel (2004) studied the possibility of using a structured path to facilitate the reproduction of very large correlation ranges, and recently, Nussbaumer et al. (2018) have analyzed exhaustively the impact of path choice concluding that the traditionally-used random path generates realizations with small biases regarding the underlying covariance functions.

3.1.3 Computational speed

With the need to solve N kriging equations, it was clear that the covariance function had to be evaluated numerous times. A way to speed the computation time was to precalculate, and store in a lookup table, all of the covariance values that would repeatedly be required during the simulation process. These covariance functions can be precomputed easily when the realization is performed on a regular grid and all the available conditioning data are assigned to the nearest grid node.

These and a few other implementation issues along with verification tests can be found in Gómez-Hernández and Cassiraga (1994); Gómez-Hernández and Journel (1993); Journel (1989) or Deutsch and Journel (1992).

4 After

The concept of sequential simulation was so simple and so potent that it quickly became the workhorse for the generation of realizations from many different types of random functions.

4.1 Sequential Gaussian simulation

After the first applications of indicator conditional simulations (Gómez-Hernández, 1989; Journel and Gómez-Hernández, 1993), it became evident that the algorithm could be implemented for the multivariate simulation of multiGaussian random functions (Deutsch and Journel, 1992; Gómez-Hernández and Journel, 1993; Verly, 1993). As already mentioned, the univariate conditional distributions associated to a multiGaussian distribution are always Gaussian, the mean and the variance of which are obtained by the solution of a set of simple kriging equations (Journel, 1989) built with the conditioning variables. In the case of a multivariate simulation, the conditional distribution remains Gaussian and its mean and variance result from the solution of a set of cokriging equations (Gómez-Hernández and Journel, 1993).

4.2 Transition probabilities

An alternative implementation to the indicator sequential simulation was presented by Carle and Fogg (1996) and developed into a computer code by Carle (1999). They work with transition probabilities instead of indicator variograms, aware that they contain the same statistical information but are easier to interpret when looking at the spatial distribution of facies. In addition, the implementation of the full cokriging equations whereby, each indicator variable is estimated not only using the indicator data for the given threshold, but also for the indicator variables for all other thresholds, was easier to perform. The concept of sequential simulation remains the same, it only changes the way the conditional distributions in Eq. (12) are obtained.

4.3 Direct sequential simulation

Sequential Gaussian simulation would generate realizations with a given mean and a given covariance. In theory, both the mean and the covariance could be non-stationary, but they must be known and seldom this possibility is considered. The realizations generated with this algorithm will have all the Gaussian features associated to a multiGaussian distribution, in particular, all univariate distributions will be normal. When the starting data set has clearly a non-Gaussian histogram, it does not seem appropriate to use a Gaussian simulation algorithm since the histogram of the final realizations will be normal rather than non-Gaussian. To circumvent this problem, there exists the possibility of transforming the original data into their normal scores (that is, each value becomes the score of a standard Gaussian distribution with the same cumulative probability), computing the covariance of the normal scores, performing a Gaussian simulation in normal-score space, and then transforming back the results to the original space. The final realizations would have univariate distributions conforming with the non-Gaussian histogram of the

data. But, the covariance deduced from the realizations would not necessarily coincide with the one deduced from the data; since the covariance that was used for the simulations was that of the normal scores.

The question that arose was, could sequential simulation still be used to generate realizations with a given covariance and a given histogram different from a Gaussian bell? The response came with the finding by Journel (1994) that in order to reproduce a given covariance, the conditional distributions in Eq. (12) need not be Gaussian, any univariate distribution with mean and variance given by the solution of the simple kriging equations would yield realizations with the desired covariance. This was a surprising claim to the many who believed, incorrectly, that simple kriging requires normality; it was Journel who pointed out that the second-order properties of the distribution defined by the simple kriging mean and variance are correct regardless of the univariate distribution of the data values. All that the assumption of normality achieves is an easy and convenient specification of the entire distribution from its first and second order moments. Journel suggested that by smartly selecting the shape of the conditional distribution, the desired histogram could be obtained without sacrificing the reproduction of the target covariance. Later, Nowak and Srivastava (1997) demonstrated its application in a mining context, Soares (2001) presented direct sequential co-simulation, and Oz et al. (2003) wrote a software code in FORTRAN.

With direct sequential simulation it is possible to generate realizations with a given covariance and a given histogram, without the need of any forward and back transforms to and from normal space.

4.4 Faster sequential simulation

By approximating the conditional distribution of Eq. (12) by Eq. (13), sequential simulation turned out to be a fast algorithm capable of generating multiple realizations over very large domains. Dimitrakopoulos and Luo (2004) report the number of floating-point operations needed to generate a realization over N points when using an approximation to the conditional distribution using only the closest n_{max} conditioning points to be in the order of $O(Nn_{max}^3)$, a very substantial reduction from the $O(N^4)$ needed if the exact equation would be used. Yet, Dimitrakopoulos and Luo (2004) claimed that further speed improvements could be achieved by embedding the matrix factorization approach by Davis (1987) into the simulation process. Their main idea is to generate the realizations, not one point at a time, but by groups of nearby points that would be conditioned to the same subset of conditioning data. Each simulation of a set of points would be performed using matrix factorization. There is an optimal number of points to be simulated at once to achieve the maximum reduction in simulation time. This value is quantified by the authors at 80% of the maximum number of conditioning points retained. The combined use of matrix factorization and sequential simulation is termed generalized sequential simulation.

4.5 Multipoint geostatistics

The International Geostatistics Conference of Tróia was memorable. Not only some of the previous works were presented there, but, most remarkably, the Tróia proceedings also included the seminal paper that gave birth to multipoint geostatistics. The work by Guardiano and Srivastava (1993) established the premises on how to generate realizations drawn from random functions imposing higher-order statistics than a covariance. Their approach required a dramatic shift in the computation of the univariate conditional distributions involved in Eq. (12) from a theoretical result (i.e., the solution of a kriging system) to a training-image-based approach.

Equation (12), or better, its approximation (13), is also the foundation of multipoint sequential simulation. But now, computing the probability that a random variable at a given location is below a certain threshold conditional to a set of nearby values is not derived from the expression of a multivariate random function, but rather experimentally calculated by searching a training image for all places where the same pattern of conditioning data repeats. The analysis of the values that in the training image fall at the same relative location with respect to the conditioning pattern serves to build the local conditional distribution.

Again, another brilliant idea coming from Journel's lab at Stanford, easy to describe, but not so easy to implement. It is worth pointing out that André always refused to co-sign a paper unless he was instrumental in the implementation of the ideas presented. Multipoint geostatistics was one of those ideas, and credit should be given to him, but André always said "ideas are worth nothing unless they are implemented" and this is the reason why he does not appear as co-author of any of the early papers: he did not work hands-on in writing the code and running the tests, he was "only" a supervisor. It was not until much later that a paper by him was published explaining why and how multipoint geostatistics were born (Journel, 2005).

4.5.1 Training image

The concept of a training image as a substitute for a random function model was new. It was received with skepticism, and it is still today subject of criticism by some researchers. A training image is a point set, generally depicted as a raster image over a regular grid, large enough to be able to find multiple replicates of n -tuples of given values from which to construct a frequency distribution at the nearby locations that will approximate the conditional distributions at those locations. If the size of the n -tuple is small, it will be easy to find several replicates of it in the training image and to build a conditional frequency distribution, but if n is large, for instance, above six or seven, the training image must be very large to find a sufficient number of replicates to allow a meaningful approximation of a probability distribution.

The size and shape of the set made up by the n -tuple plus the point being simulated will determine the order of the statistics that are being extracted

from the training image and infused into the realization, as well as the ranges over which these multipoint statistics are under control.

The construction or the selection of a training image is another issue subject to much debate, which will not be discussed here (see, for instance, Mariethoz and Caers, 2014).

4.5.2 Lookup tables

Scanning the training image for each point being simulated to find the replicates of the conditioning data was very computationally intensive. Strebelle (2000, 2002) devised a lookup table approach coupled with a tree search to first build the conditional distributions of all possible conditioning data configurations and, second, retrieve quickly a given distribution from the table.

Building the lookup table was time consuming and limited the high-order statistics that would be used to generate the realization. This problem was solved by Mariethoz et al. (2010) with their direct sampling approach. They realized that exhaustively searching the training image for all replicates of a given conditioning set and then drawing a random value from this set was the same as doing a random search and retaining the first replicate, thus speeding up considerably the drawing from the conditional distribution.

4.5.3 Continuous variables

Most of the first implementations of the multipoint sequential simulation was done for a binary variable (Guardiano and Srivastava, 1993; Strebelle, 2000, 2002). It is easier to find replicates of a conditioning data set when the data values can only be ones or zeroes. The extension to categorical variables with more than two categories was obvious in theory, but the construct of the lookup tables and the search tree becomes cumbersome and inefficient. Attempting to perform the simulation of continuous variables was out of the question, even considering a binning approach that would transform the continuous variable into a categorical one.

Again, the direct sampling approach by Mariethoz et al. (2010) solved the problem. Once the conditioning data set is defined, some tolerances are applied to the conditioning data values and the training image is scanned to find replicates in which the values are within the tolerance limits, once a match is found, the value at the position of the point being estimated is retrieved and assigned to the simulating node.

A thorough analysis of all the implementation issues and the sensitivities to some of the tuning parameters in the direct sampling algorithm can be found in the work by Meerschman et al. (2013).

4.6 Sequential simulation with patterns

The concepts of multipoint geostatistics and training images is taken one step further by considering the simulation of blocks of neighboring points all at once

Arpat (2005); Arpat and Caers (2007) . The concept of sequential simulation remains but the decomposition of the multivariate random function (12) is replaced with

$$\begin{aligned}
 F(Z(\mathbf{u}_1), Z(\mathbf{u}_2), \dots, Z(\mathbf{u}_n)|(n_o)) &= F(B(\mathbf{v}_1)|(n_o)) \cdot \\
 &F(B(\mathbf{v}_2)|(n_o), B(\mathbf{v}_1)) \cdot \\
 &F(B(\mathbf{v}_3)|(n_o), B(\mathbf{v}_1), B(\mathbf{u}_2)) \cdots \quad (14) \\
 &F(B(\mathbf{v}_{n_b})|(n_o), B(\mathbf{v}_1), \dots, B(\mathbf{v}_{n_b-1})),
 \end{aligned}$$

where $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{n_b}\}$ are the centroids of the blocks $\{B(\mathbf{v}_1), B(\mathbf{v}_2), \dots, B(\mathbf{v}_{n_b})\}$ that fully tessellate the initial point set $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$, with $n_b \ll n$. The above expression is simplified as in Eq. (13) and only the closest simulated blocks and conditioning data are retained to build the conditioning pattern that is scanned on the training image in search for replicates. By simulating blocks of points, the number of steps in the sequential simulation is reduced and therefore the speed is increased; also, accordingly to the authors, this approach is able to capture curvilinear and complex features from the training image.

The concept of simulation with patterns was further developed by Mahmud et al. (2014) who used conditional image quilting.

4.7 Sequential simulation with high-order spatial cumulants

An interesting alternative to the multipoint sequential simulation approach described above is sequential simulation using high-order spatial cumulants as proposed by Dimitrakopoulos et al. (2010) and Mustapha and Dimitrakopoulos (2010, 2011). Cumulants can be regarded as a generalization of the covariance function to orders higher than two; as a matter of fact, the cumulant of second order is the covariance. Any multivariate probability function can be expressed in terms of its moments or in terms of its cumulants. Dimitrakopoulos et al. (2010) show that the cumulants of order three to five can capture the same complex features as the multipoint approach. Cumulants can also be used to determine the conditional distributions in Eq. (12). Mustapha and Dimitrakopoulos (2010) write these conditional distributions as Legendre polynomial expansions where the coefficients of the expansion are written in terms of the cumulants. The cumulants are derived from the data in a similar way as a covariance or a variogram; however, they contain higher-order moment information and can be used for the generation of realizations from non-Gaussian random functions. Recently, the technique has been extended to the joint simulation of several variables (Minniakhmetov and Dimitrakopoulos, 2017) and an alternative formulation has been proposed in which, instead of using Legendre polynomials, orthogonal splines are employed and cumulants are replaced by alternative coefficients that can be estimated from the data (Minniakhmetov et al., 2018).

A key aspect of this high-order spatial cumulant approach is that it is data-driven as opposed to the multipoint approach that is training-image driven; the latest work in this regard proposes a training-image-free approach to high-order simulation (Yao et al., 2020).

5 Not sequential simulation

Sequential simulation is a widely spread concept for the drawing of realizations from random functions of different kinds, from Gaussian to indicator-based to training-image-based to non-Gaussian of other types. But there are other approaches that are also being used that are not based on the sequential simulation concept. Without trying to be exhaustive, some of those alternative methods are gradual deformation (Hu, 2000), pluriGaussian simulations (Armstrong et al., 2011; Galli et al., 1994), or Markov chain Monte Carlo (Fu and Gómez-Hernández, 2009).

6 Conclusion

The sequential simulation algorithm has dominated the field of stochastic simulation of spatial random functions since the late 1980s. Its implementation for the generation of realizations drawn from random functions defined non-parametrically on the basis of indicator variables was the sparkle that initiated a surge of variants aimed at the simulation from random functions that were getting far from the standard stationary multiGaussian one. The co-authors of this paper witnessed the birth of sequential indicator simulation and were fundamental during the first stages of that new algorithm, an algorithm that is based in the simple concept of breaking a complex problem into many simple ones and then progressing one step at a time.

One of the great talents of André Journel is that, throughout his career, he was able to find the core of an idea, to understand why it worked well (or didn't work so well), and to construct a generalization that allows the idea to flourish in other applications. And so it was with sequential simulation: what began as a method for indicator simulation that was freed from the constraints of an unnecessary intermediate Gaussian random function soon became recognized by André as a rich paradigm, the idea that you can create a simulation simply by correctly sequencing a series of random drawings, one step at a time. When the first multipoint simulation algorithm was written, it was André who recognized that it was, at its core, a sequential algorithm which opened up many new branches of inquiry based on all that had already been learned about other sequential algorithms.

Acknowledgements The first author wishes to acknowledge the financial contribution of the Spanish Ministry of Science and Innovation through project number PID2019-109131RB-I00.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Alabert FG (1987) The practice of fast conditional simulations through the LU decomposition of the covariance matrix. *Math Geology* 19(5):369–387
- Anderson TW (1984) *Multivariate statistical analysis*. Wiley, New York
- Armstrong M, Galli A, Beucher H, Loc'h G, Renard D, Doligez B, Eschard R, Geffroy F (2011) *Plurigaussian simulations in geosciences*. Springer Science & Business Media
- Arpat GB (2005) *Sequential simulation with patterns*. PhD thesis, Stanford University
- Arpat GB, Caers J (2007) Conditional simulation with patterns. *Mathematical Geology* 39(2):177–203
- Bratley P, Fox BL, S LE (1983) *A Guide to Simulation*. Springer-Verlag, New York
- Carle SF (1999) T-progs: Transition probability geostatistical software. University of California, Davis, CA 84
- Carle SF, Fogg GE (1996) Transition probability-based indicator geostatistics. *Mathematical Geology* 28(4):453–476
- Davis MW (1987) Production of conditional simulations via the LU triangular decomposition of the covariance matrix. *Math Geol* 19(2):91–98
- Deutsch CV, Journel AG (1992) *GSLIB, Geostatistical Software Library and User's Guide*. Oxford University Press, New York
- Dimitrakopoulos R, Luo X (2004) Generalized sequential gaussian simulation on group size ν and screen-effect approximations for large field simulations. *Mathematical Geology* 36(5):567–591
- Dimitrakopoulos R, Mustapha H, Gloaguen E (2010) High-order statistics of spatial random fields: exploring spatial cumulants for modeling complex non-gaussian and non-linear phenomena. *Mathematical Geosciences* 42(1):65
- Durstenfeld R (1964) Algorithm 235: random permutation. *Communications of the ACM* 7(7):420
- Fu J, Gómez-Hernández JJ (2009) A blocking Markov chain Monte Carlo method for inverse stochastic hydrogeological modeling. *Mathematical Geosciences* 41:105–128, DOI 10.1007/s11004-008-9206-0
- Galli A, Beucher H, Le Loc'h G, Doligez B, et al. (1994) The pros and cons of the truncated gaussian method. In: *Geostatistical simulations*, Springer, pp 217–233
- Gómez-Hernández JJ (1989) Indicator conditional simulation of the architecture of hydraulic conductivity fields: application to a sand-shale sequence. In: Sahuquillo A, Andréu J, O'Donnell T (eds) *Groundwater Management:*

- Quantity and Quality, International Ass. of Hydrological Sciences, 188, pp 41–51
- Gómez-Hernández JJ, Cassiraga EF (1994) Theory and practice of sequential simulation. In: Armstrong M, Dowd P (eds) *Geostatistical Simulations*, Kluwer Academic Publishers, pp 111–124
- Gómez-Hernández JJ, Journel AG (1993) Joint sequential simulation of Multi-Gaussian fields. In: Soares A (ed) *Geostatistics Tróia '92*, Kluwer Academic Publishers, Dordrecht, vol 1, pp 85–94
- Gómez-Hernández JJ, Srivastava RM (1990) ISIM3D: An ANSI-C three dimensional multiple indicator conditional simulation program. *Computer and Geosciences* 16(4):395–440
- Gómez-Hernández JJ, Wen XH (1994) Probabilistic assessment of travel times in groundwater modeling. *J of Stochastic Hydrology and Hydraulics* 8(1):19–56
- Guardiano FB, Srivastava RM (1993) Multivariate geostatistics: Beyond bivariate models. In: Soares A (ed) *Geostatistics Tróia '92*, volume 1, Kluwer, pp 133–144
- Hu LY (2000) Gradual deformation and iterative calibration of gaussian-related stochastic models. *Mathematical Geology* 32(1):87–108
- Johnson ME (1987) *Multivariate statistical simulation: A guide to selecting and generating continuous multivariate distributions*, vol 192. John Wiley & Sons
- Journel A (1982) The indicator approach to estimation of spatial distributions. In: *Proceedings of the 17th APCOM International Symposium*, New York, pp 793–806
- Journel AG (1974) Geostatistics for conditional simulation of ore bodies. *Economic Geology* 69(5):673–687
- Journel AG (1983) Nonparametric estimation of spatial distributions. *Journal of the International Association for Mathematical Geology* 15(3):445–468
- Journel AG (1989) *Fundamentals of Geostatistics in Five Lessons*, Short Courses in Geology, vol 8. Agü, Washington D.C.
- Journel AG (1994) Modeling uncertainty: some conceptual thoughts. In: *Geostatistics for the next century*, Springer, pp 30–43
- Journel AG (2005) Beyond covariance: the advent of multiple-point geostatistics. In: *Geostatistics Banff 2004*, Springer, pp 225–233
- Journel AG, Gómez-Hernández JJ (1993) Stochastic imaging of the Wilmington clastic sequence. *SPE Formation Evaluation* March 93:33–40
- Journel AG, Huijbregts CJ (1978) *Mining Geostatistics*. Academic Press, London
- Journel AG, Isaaks EH (1984) Conditional indicator simulation: Application to a Saskatchewan uranium deposit. *Math Geology* 16(7):685–718
- Liu Y, Journel A (2004) Improving sequential simulation with a structured path guided by information content. *Mathematical Geology* 36(8):945–964
- Mahmud K, Mariethoz G, Caers J, Tahmasebi P, Baker A (2014) Simulation of earth textures by conditional image quilting. *Water Resources Research* 50(4):3088–3107

- Mariethoz G, Caers J (2014) Multiple-point geostatistics: stochastic modeling with training images. John Wiley & Sons
- Mariethoz G, Renard P, Straubhaar J (2010) The direct sampling method to perform multiple-point geostatistical simulations. *Water Resources Research* 46(11)
- Matérn B (1960) Spatial variation, volume 36 of. Lecture Notes in Statistics
- Matheron G (1973) The intrinsic random functions and their applications. *Advances in applied probability* 5(3):439–468
- Meerschman E, Pirot G, Mariethoz G, Straubhaar J, Van Meirvenne M, Renard P (2013) A practical guide to performing multiple-point statistical simulations with the direct sampling algorithm. *Computers & Geosciences* 52:307–324
- Minniakhmetov I, Dimitrakopoulos R (2017) Joint high-order simulation of spatially correlated variables using high-order spatial statistics. *Mathematical Geosciences* 49(1):39–66
- Minniakhmetov I, Dimitrakopoulos R, Godoy M (2018) High-order spatial simulation using legendre-like orthogonal splines. *Mathematical geosciences* 50(7):753–780
- Mustapha H, Dimitrakopoulos R (2010) High-order stochastic simulation of complex spatially distributed natural phenomena. *Mathematical Geosciences* 42(5):457–485
- Mustapha H, Dimitrakopoulos R (2011) Hosim: a high-order stochastic simulation algorithm for generating three-dimensional complex geological patterns. *Computers & geosciences* 37(9):1242–1253
- Nowak M, Srivastava R (1997) A geological conditional simulation algorithm that exactly honours a predefined grade-tonnage curve. *Proceedings of the Geostatistics Wollongong* 96:669–682
- Nussbaumer R, Mariethoz G, Gloaguen E, Holliger K (2018) Which path to choose in sequential gaussian simulation. *Mathematical Geosciences* 50(1):97–120
- Oz B, Deutsch CV, Tran TT, Xie Y (2003) DSSIM-HR: A FORTRAN 90 program for direct sequential simulation with histogram reproduction. *Computer and Geosciences* 29(2003):39–51, URL <http://www.iamg.org/CDEditor/index.htm>
- Rosenblatt M (1952) Remarks on a multivariate transformation. *The annals of mathematical statistics* 23(3):470–472
- Shinozuka M, Jan CM (1972) Digital simulation of random processes and its applications. *Journal of sound and vibration* 25(1):111–128
- Soares A (2001) Direct sequential simulation and cosimulation. *Math Geology* 33(8):911–926
- Strebelle S (2000) Sequential simulation drawing structures from training images. PhD thesis, Stanford University. 187pp
- Strebelle S (2002) Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical geology* 34(1):1–21
- Verly GW (1993) Sequential gaussian cosimulation: A simulation method integrating several types of information. In: Soares A (ed) *Geostatistics Tróia*

'92, volume 1, Kluwer, pp 543–554

Yao L, Dimitrakopoulos R, Gamache M (2020) Training image free high-order stochastic simulation based on aggregated kernel statistics. *Mathematical Geosciences* pp ?–?