



Research paper

Calibration of random fields by a sequential spectral turning bands method

Dany Lauzon*, Denis Marcotte

Civil, Geological and Mining Department, Polytechnique Montréal, C.P. 6079 Succ. Centre-ville, Montréal, (Qc), H3C 3A7, Canada

ARTICLE INFO

Keywords:

Constructive calibration
 Conditional simulation
 Spectral turning bands
 High-order statistics

ABSTRACT

A new algorithm for calibration of conditional realizations to measured or desired response functions is presented. The Sequential-Spectral Turning Bands Method (S-STBM) builds the field by choosing the phase of each new cosine function such that the observed field response functions become increasingly calibrated. The phase selection has little influence on the spatial correlation structure but can help to meet other objectives. Conditioning by kriging is used in the algorithm main loop to impose exact hard data reproduction. A first case study illustrates the performance of the algorithm for a cyclic and asymmetric field. S-STBM is shown to reproduce similarly or better the directional asymmetry than calibrated realizations obtained by FFTMA-SA. A training image (TI) with connected low values provides the second case study where the target is the reproduction of non-centered third-order spatial moments. A third case study shows the effectiveness of the S-STBM algorithm to calibrate a Gaussian field to tracer tests. Contrary to FFTMA-SA, S-STBM works on irregular grids. Its computational complexity of $\mathcal{O}(n)$ and small memory requirement makes it an attractive method for calibration.

1. Introduction

Several geostatistical methods exist to simulate conditional random Gaussian fields (Lantuéjoul, 2002; Chilès and Delfiner, 2012). These methods allow the reproduction of second-order spatial statistics but do not guarantee the reproduction of field characteristics such as high-order spatial statistics (Tsatsanis and Giannakis, 1992; Journel, 1993; Delopoulos et al., 1994; Dimitrakopoulos et al., 2010; Hörning and Bárdossy, 2018), connectivity measures (Renard and Allard, 2013), travel time between wells in tracer tests (Sudicky, 1986), or generally any information that relates non-linearly to the properties of the simulated fields.

Many different approaches have been proposed to incorporate non-Gaussian characteristics in simulated fields or to calibrate the field to extraneous information. History matching with Ensemble Kalman Filtering (Liu and Oliver, 2005; Evensen, 2009) use sequential updating of a field using the covariance between the field properties and the response variables. Other approaches optimize an objective function (OF) either using gradients (Marsily et al., 1984; Gómez-Hernández et al., 1997), or by iteratively perturbing the field (Kirkpatrick et al., 1983; Hörning and Bárdossy, 2018; Lauzon and Marcotte, 2019) or a combination of both (Hu and Ravalec-Dupin, 2004).

A commonly used perturbation method for calibration is the Simulated Annealing (SA) algorithm (Kirkpatrick et al., 1983; Geman and Geman, 1984). However, SA converges slowly and hard data points (HD) can present undesired discontinuities with the neighboring

cells (Hörning and Bárdossy, 2018). A method proposed by Hörning and Bárdossy (2018) called phase annealing (PA) improves embedding by applying SA on the Fourier transform (FFT) phase component. PA suffers of two important drawbacks. First, the size of problems that can be tackled is limited due to the use of the FFT. Secondly, the FFT algorithm can only be used for regular grids. Algorithms combining the FFT-Moving Average with gradual deformation (FFTMA-GD, Hu and Ravalec-Dupin (2004)) or with simulated annealing (FFTMA-SA, Lauzon and Marcotte (2019)) also suffer from the same limitations.

The Spectral Turning Bands Method (STBM) and the Turning Bands Method (TBM) both allow to obtain realizations over irregular grids and solve the memory issue. The simulation in \mathbb{R}^n is replaced by a series of simulations on \mathbb{R} . The STBM developed by Shinozuka (1971) and Shinozuka and Jan (1972) and studied by Emery et al. (2015) is based on the sequential addition of several one-dimensional cosine waves with specific spectral densities and random phases. The main idea of our proposed algorithm is to select the phase of each cosine function such as to minimize an OF that includes elements to calibrate. The calibration is performed sequentially, wave by wave, during the construction of the random field. Besides grid and memory advantages, it does not require to set extra parameters such as a cooling schedule.

The paper is structured as follows. First, the main characteristics of the STBM are summarized. Then, Sequential-STBM (S-STBM) algorithm is described. Directional asymmetry and third-order statistics

* Corresponding author.

E-mail address: dany.lauzon@polymtl.ca (D. Lauzon).

used in our synthetic examples are introduced. The first case study shows cyclical and asymmetric structures and is used to compare S-STBM to FFTMA-SA. A second case study taken from a training image illustrates the use and performance of S-STBM with different OFs. A third example aiming to calibrate a Gaussian porosity field to travel time between wells in a tracer test shows the effectiveness of the algorithm. A discussion on strengths and weaknesses of the proposed method follows.

2. Methodology

This section begins with a general presentation of the STBM algorithm followed by our proposed adaptation for sequential calibration (S-STBM). Thereafter, some implementation details of the algorithm are presented.

2.1. Spectral Turning Band Method

As for its spatial counterpart, the main idea of the STBM is to replace the simulation process in \mathbb{R}^n (in general, $n = 2$ or 3) into a simpler series of simulations in \mathbb{R} . Let us consider a n -dimensional stationary continuous covariance $C_n(h)$. Its spectral representation is (Chilès and Delfiner, 2012, p. 66) :

$$C_n(h) = \int_{\mathbb{R}^n} e^{i\langle h, \omega \rangle} d\chi(\omega) = \int_{\mathbb{R}^n} \cos(\langle h, \omega \rangle) d\chi(\omega) \quad (1)$$

where $h = (h_1, h_2, \dots, h_n)$ is a n -dimensional vector specifying a direction, $\omega = (\omega_1, \omega_2, \dots, \omega_n)$ denotes a n -dimensional frequency vector, $\langle h, \omega \rangle = h_1\omega_1 + h_2\omega_2 + \dots + h_n\omega_n$ is the inner scalar product in \mathbb{R}^n and $d\chi(\omega)$ is the spectral measure. Note that $C_n(h)$ being real and symmetrical around the origin, the spectral measure is also real and symmetrical around the origin. Further, when $C_n(h)$ is square integrable the spectral measure can be expressed as $d\chi(\omega) = f(\omega)d\omega$ where $f(\omega)$ is the spectral density (Lantuéjoul, 2002, p. 191). Thereafter let us consider a stationary zero-mean random field $Y(x)$ defined by:

$$Y(x) = \sqrt{2}\cos(\langle V, x \rangle + 2\pi U) \quad (2)$$

where V is a random frequency vector with distribution given by the spectral density $f(\omega)$ and U is uniformly distributed on $[0, 1]$. The random process $Y(x)$ is zero-mean and possesses covariance $C_n(h)$ as:

$$E[Y(x)] = \int_0^1 dU \int_{\mathbb{R}^n} d\chi(V) \sqrt{2}\cos(\langle V, x \rangle + 2\pi U) = 0 \quad (3)$$

$$\begin{aligned} E[Y(x)Y(x+h)] &= \int_0^1 dU \int_{\mathbb{R}^n} d\chi(V) 2\cos(\langle V, x \rangle + 2\pi U) \\ &\quad \times \cos(\langle V, x+h \rangle + 2\pi U) \\ &= \int_0^1 dU \int_{\mathbb{R}^n} d\chi(V) \cos(\langle V, h \rangle) \\ &\quad + \cos(2\langle V, x \rangle + \langle V, h \rangle + 4\pi U) \\ &= \int_{\mathbb{R}^n} d\chi(V) \cos(\langle V, h \rangle) = C_n(h) \end{aligned} \quad (4)$$

Eq. (4) clearly establishes that random variable V is responsible for the covariance reproduction and U is a uniformly distributed random phase shift bearing no effect on C_n . Although $Y(x)$ in Eq. (2) is zero mean and has the desired covariance, it is non-ergodic and non-Gaussian as it is made of a single cosine function. To enforce ergodicity and Gaussian field, a sum of N independent random processes $Y(x)$ is used:

$$Z(x) = \sqrt{\frac{1}{N}} \sum_{i=1}^N Y_i(x) \quad (5)$$

$Z(x)$ becomes more ergodic and Gaussian as N increases. Eq. (5) can be implemented sequentially or parallelized. S-STBM uses a sequential implementation for the cosine functions where, instead of being chosen randomly, the phase U of each added process $Y(x)$ is selected so as to minimize the desired OF at the current step. Note that parallelization

can still be done for each cosine function by computing the coordinate projections on the line with the GPU as suggested by Räss et al. (2019).

The spectral density defines the distribution of V in Eq. (2). For isotropic 3D fields the spectral density is radially symmetric. It can be easily determined by taking the Fourier transform of the line covariance in the TB operator. The line covariance is (Matheron, 1973):

$$C_1(h) = \frac{d(hC_3(h))}{dh} \quad (6)$$

The one-dimensional spectral density $f_1(\omega)$ is then:

$$f_1(\omega) = 2 \int_0^\infty C_1(h) \cos(\omega h) dh \quad (7)$$

Another possibility is to compute the Fourier transform of $C_3(h)$ to obtain the spectral density $f_3(\omega)$ and from it obtain the one-dimensional spectral density $f_1(\omega)$.

$$f_3(\omega) = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} C_3(h) e^{-i\langle \omega, x \rangle} dx \quad (8)$$

$$f_1(\omega) = (2\pi\omega)^2 f_3(\omega) \quad (9)$$

Spectral densities can be found for most common covariance models in Lantuéjoul (2002), Emery and Lantuéjoul (2006), Chilès and Delfiner (2012), Marcotte (2015, 2016) and Marcotte and Allard (2017).

Drawing V from a given spectral density can be done by sampling uniformly the inverse cumulative distribution. First, a random direction is selected in 3D. This can be done by generating three independent Gaussian random values and normalizing the resulting vector to a unit vector giving u the random direction. The radial spectral density is then integrated to provide the cumulative distribution function (cdf):

$$F_1(\omega) = \frac{1}{\pi} \int_0^\omega f_1(s) ds \quad (10)$$

Then, a uniform 0-1 value t is drawn and V is taken as $V = F_1^{-1}(t)u$.

2.2. S-STBM algorithm

The S-STBM works essentially as STBM except that U is chosen to respect as much as possible to desired additional information on the field like high-order statistics (Dimitrakopoulos et al., 2010; Hörning and Bárdossy, 2018), connectivity measures (Renard and Allard, 2013), travel time between wells in tracer tests or pressure drop or water cut in petroleum wells (see : Oliver et al. (1997), Le Ravalec-Dupin and Hu (2005), Oliver et al. (2010), Hu et al. (2013), Rezaee and Marcotte (2018), among many others).

In summary, the S-STBM algorithm is:

(a) Initialization step

- Computation of Gaussian anamorphosis $Y = \phi(Z)$ where Z is the data in original units, Y is the Gaussian equivalent
- Determination of covariance model $C_3(h)$ for Y
- Computation of $f_1(\omega)$ with Eq. (6) and (7) or with Eq. (8) and (9).
- Integration of $f_1(\omega)$ to get $F_1(\omega)$

(b) Iteration step $\forall i = 1 \dots N$

- Choose a line direction u_i over the unit half-sphere (e.g. using van der Corput sequence, see Section 2.2.1)
- Draw uniformly t_i in $[0, 1]$ and define $w_i = F_1^{-1}(t_i)$
- Define $V_i = w_i u_i$
- Phase optimization
 - Search step for U in $[0, 1]$. For candidate value U_{can} obtained from the linesearch

- i. Compute the random field $Y_i = \sqrt{\frac{i-1}{i}}Y_{i-1} + \sqrt{\frac{1}{i}}\left(\sqrt{2}\cos(\langle V_i, x \rangle) + 2\pi U_{can}\right)$ where the first term on the right side represents the contribution of the field at the previous iteration (invariant) and the right term is the perturbation brought by the new cosine function defined by V_i and candidate phase U_{can}
- ii. Unconditional random field Y_i is post-conditioned by kriging to get Y_{c_i}
- iii. Backtransformation to get $Z_{c_i} = \phi^{-1}(Y_{c_i})$
- iv. The objective function OF_i is computed with Z_{c_i}
- v. Keep the best candidate U_{can} and corresponding Y_i minimizing the OF

(c) Evaluate stopping criteria.

Note that a Gaussian anamorphosis is performed as a first step because STBM converges to Gaussian field. Gaussian quantiles are associated to HD experimental quantiles so as to ensure a marginal Gaussian distribution. The exact representation of HD is obtained by a post conditioning by kriging right before the backtransformation to the original unit field and the evaluation of the OF (see step b.). This step is classical in geostatistics, for further details see [Chilès and Delfiner \(2012\)](#) and [Lauzon and Marcotte \(2019\)](#).

We stress that in the algorithm Y_i represents the entire field adding the contributions of i cosine functions. The current field Y_i is thus formed of the $i - 1$ already determined cosine functions forming field Y_{i-1} and the i th cosine function defined by the couple (V_i, U_{can}) with

U_{can} the phase candidate value being currently examined in the line-search. The current field $Y_i(x)$ is then post-conditioned exactly by kriging using known HD values $Y(x_k)$ ($k = 1 \dots n$) and backtransformed before evaluating the OF. Hence OF evaluation is always done using a field that is exactly conditioned to HD. The line search for optimal U fully integrates the effect of HD conditioning on the field. This step helps to strategically choose the phase by taking fully into account the information contained in the HD, allowing a better convergence of the OF. A similar strategy was applied with the FFTMA-SA algorithm presented in [Lauzon and Marcotte \(2019\)](#). The number of trials for U_{can} (see Section 2.2.3 for details) can correspond to a few steps in a golden search line optimization ([Luenberger and Ye, 2008](#)). The stopping criteria can be, as examples, the number of iterations (therefore of optimized cosine functions) N is reached or the OF does not improve sufficiently after a series of iterations.

2.2.1. Distribution of lines over \mathbb{R}^n

The first implementations of TBM used the maximum number in 3D of 15 exactly regularly spaced lines obtained by joining the edges of an icosahedron. It was quickly acknowledged that many more lines were required to avoid banding artifacts in the simulated fields ([Chilès and Delfiner, 2012](#), pp 506–508). As shown in [Figs. 1\(a\) and 1\(c\)](#), a completely random distribution over the half-sphere does not provide a very regular discretization. [Freulon and de Fouquet \(1991\)](#) studied different methods to obtain quasi-regular distribution of lines. The Halton sequence 1964 and the van der Corput sequence 1935 ([Figs. 1\(b\) and 1\(d\)](#)) provide more uniform sampling ([Chilès and Delfiner, 2012](#)). The van der Corput sequence is used in our algorithm.

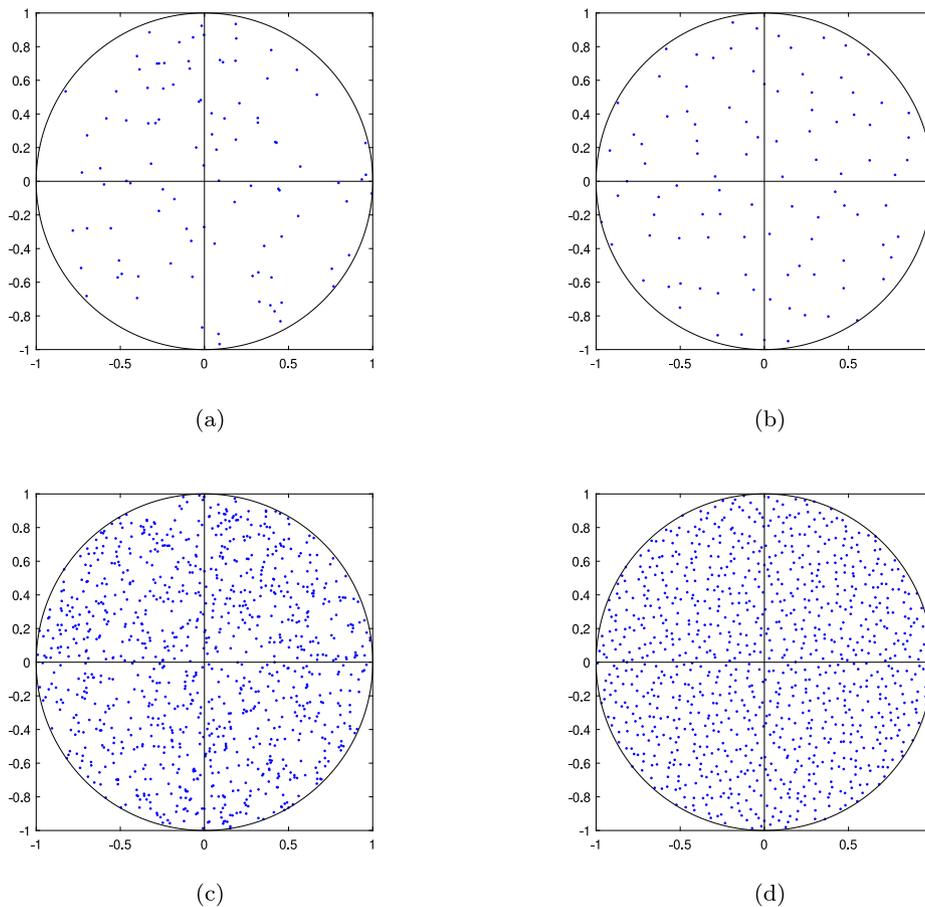


Fig. 1. Lambert azimuthal equiareal projection of 100 points (top row) and 1000 points (bottom row). Left column: random; right column: quasi-random (van der Corput sequence).

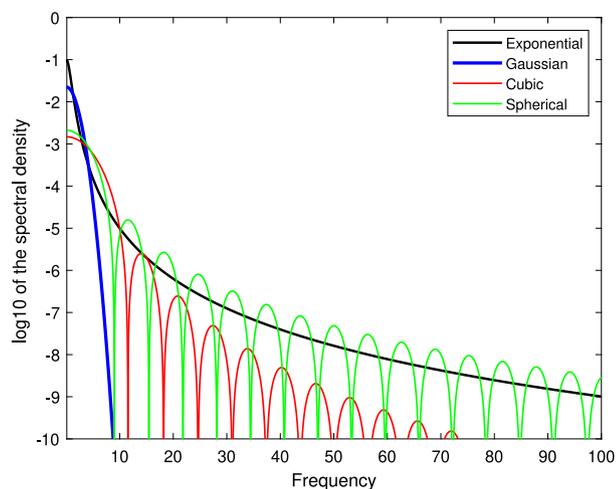


Fig. 2. Log-spectral density of four covariance models. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

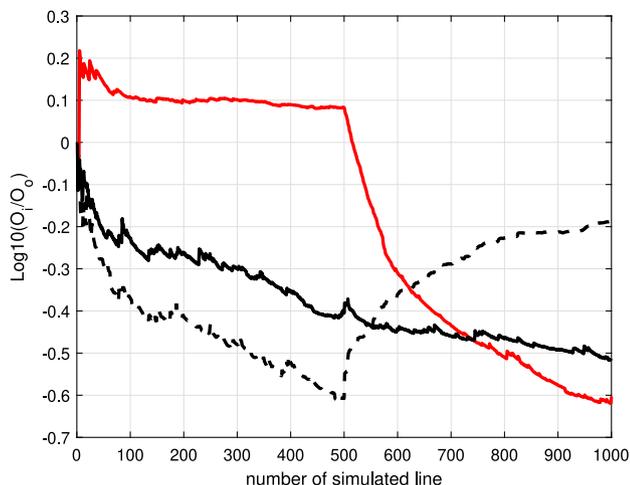


Fig. 3. Mean OF over 100 realizations for three calibration strategies: 500 last lines (solid red), 500 first lines (dashed black), one every two lines (solid black) are calibrated. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

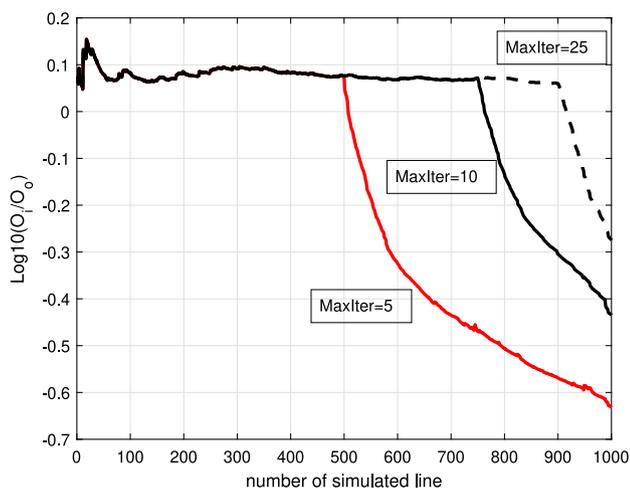


Fig. 4. Mean OF over 100 realizations for 2500 total OF evaluations and maximum number of iterations per line of 5 (solid red), 10 (solid black) and 25 (dashed black). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.2.2. Number of lines used

Tompson et al. (1989) suggest to generate 100 randomly oriented lines for the spatial TBM. Freulon and de Fouquet (1991) have simulated spherical and linear covariance model with 100 quasi-random lines (van der Corput sequence) and their results suggest that it could suffice for some covariances with the spatial turning band. Other authors (Lantuéjoul, 2002; Emery and Lantuéjoul, 2006; Chilès and Delfiner, 2012; Marcotte, 2016) propose to use several hundred to several thousand lines, especially for the spectral version of TBM which needs to sample the high frequencies for covariance models with linear behavior at the origin (see Fig. 2).

2.2.3. Calibration with the phase

Since the number of lines to be simulated can reach several hundred or several thousand lines, the computational cost of S-STBM can become large when each OF evaluation is itself CPU intensive. One idea to save some CPU is to optimize the phase component only for some iterations and randomly select the phase for the others. Fig. 3 compares three different strategies. The lines to optimize are either selected 1) all at the end (red line), 2) all at the beginning (black dashed line) or 3) alternately (black line). The test was carried on a total of 1000 lines with 500 of them being optimized. When optimization is performed at the beginning, the last not optimized lines lose all the gain obtained in the initial iterations. The alternating solution provides good results but the optimization at the end provides overall better calibrated realizations with this example. This is our retained method as it seems logical to create first a field with the desired covariance that is then slowly modified so as to meet other targets. At each iteration where phase is optimized, it is probably advisable to stop early the optimization to limit the number of OF evaluations for a single line. Fig. 4 shows how the OF behaves as a function of the number of phase values considered (5, 10 and 25) for each line. The total number of calls to the OF is kept constant at 2500. Clearly, it is better to optimize lightly the phase on more lines (red line) than deeply on less lines (black line and black dashed line). Thus, the maximum number of iterations for a single line is set to 5 in the following sections.

3. Objective function

Variogram based simulated fields bear an implicit Gaussian collocation, especially when the conditioning data are scarce. Non-Gaussian characteristics can however be included in the objective function to optimize (Guthke and Bárdossy, 2017; Bárdossy and Hörning, 2017; Hörning and Bárdossy, 2018; Lauzon and Marcotte, 2019). Two useful statistics to describe non-Gaussian features are the directional asymmetry function (Hörning and Bárdossy, 2018) and the spatial high order moments (Dimitrakopoulos et al., 2010).

3.1. Directional asymmetry

Directional asymmetry occurs in many instances. A classical example is provided by the Barchan sand dune. The softer slope faces the wind and the steeper slope occurs on the opposite side. This induces a clear directional asymmetry (Bárdossy and Hörning, 2017). The directional asymmetry can be evaluated by (Hörning and Bárdossy, 2018):

$$A(h) = \frac{1}{N(h)} \sum_{x_i - x_j \approx h} (F(Z(x_i)) - F(Z(x_j)))^3 \quad (11)$$

where $N(h)$ is the number of pairs with separation vector h and $F(Z(x))$ is the marginal *cdf* evaluated at $Z(x)$. This function can be computed very efficiently using FFTs (Marcotte, 1996).

3.2. Non-centered third order spatial moment

For a stationary random field $Z(x)$, the spatial non-centered third-order moment is defined as:

$$M(h_1, h_2) = \frac{1}{N_{h_1, h_2}} \sum_{i=1}^{N_{h_1, h_2}} Z(x)Z(x + h_1)Z(x + h_2) \quad (12)$$

where N_{h_1, h_2} is the number of triplets with separation vectors h_1, h_2 . This statistic can be computed efficiently using the bispectrum function (Horikawa, 2000) that can also be obtained using FFTs. However computing the spatial third moment for all combinations of lag vectors is quite demanding in memory. For a field with k cells along each of its d dimensions, the memory required to store all the triplets is k^{2d} , which quickly becomes prohibitive for $d \geq 2$. It is then advisable in practice to select only a few h_1 or h_2 lag vectors for the computation. For a fixed

h_2 in Eq. (12), we define $f(x) = Z(x)Z(x + h_2)$ and $g(x + h) = Z(x + h_1)$. Then, Eq. (12) amounts to computing the covariance of f and g . This computation is done by FFT in $O(k^d \log(k^d))$ operations and requires only $O(k^d)$ in memory space for each selected h_2 .

4. Results

Three synthetic cases are studied. S-STBM is compared to uncalibrated and FFTMA-SA realizations. The experimental spatial asymmetry and the third-order non-centered moments are used as targets in the OFs of the first two case studies. The last case study presents a calibration of a tracer test. The OF is formed by the MSRE for travel times obtained between an injection well and pumping wells in the reference and the simulation.

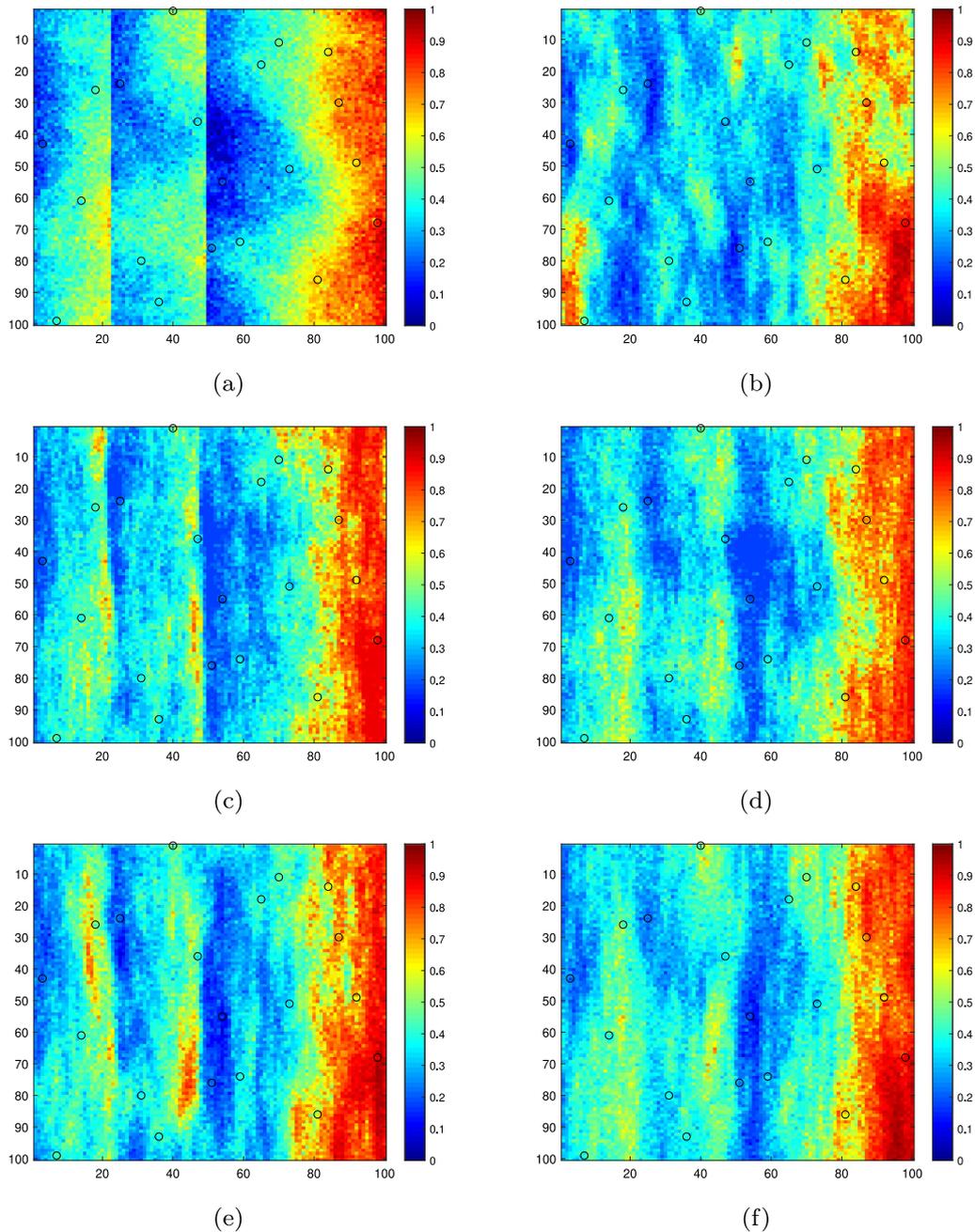


Fig. 5. (a) Reference field and location of conditioning data (black circles); (b) Uncalibrated STBM; (c) FFTMA-SA with OF directional asymmetry; (d) FFTMA-SA with OF third-order moments; (e) S-STBM with OF directional asymmetry; (f) S-STBM with OF third-order moments.

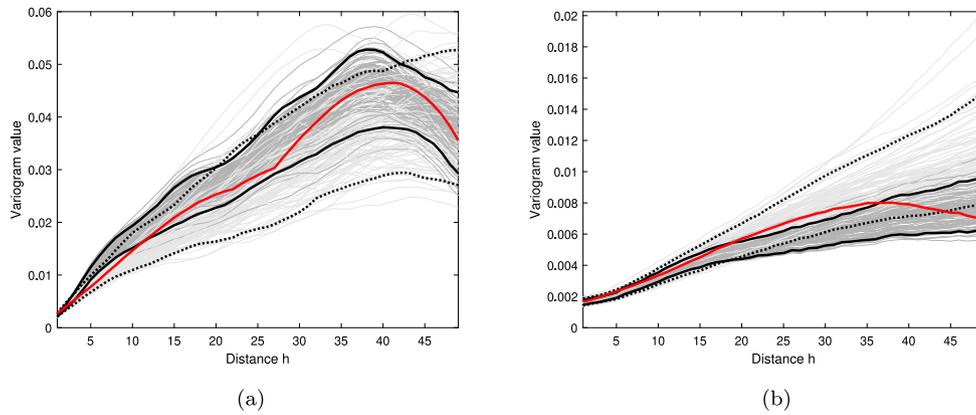


Fig. 6. Synthetic case. (a) Variograms along East direction, (b) along South direction. Reference field (solid red), S-STBM realizations (dark gray), uncalibrated STBM (light gray); 90% confidence intervals shown (solid dark: calibrated and dashed lines: uncalibrated). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

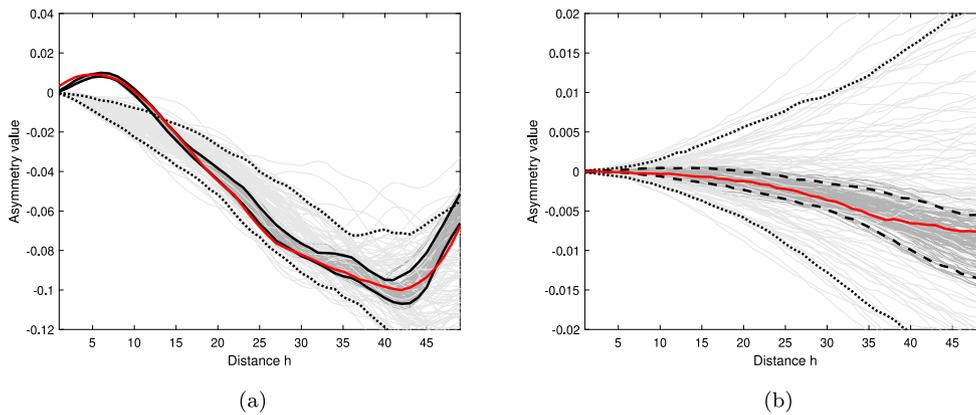


Fig. 7. Synthetic case. (a) Asymmetry function along East direction, (b) along South direction. Reference field (solid red), S-STBM realizations (dark gray), uncalibrated STBM (light gray); 90% confidence intervals shown (solid dark: calibrated and dashed lines: uncalibrated). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Synthetic case: Correlations of simulated fields with reference field; 100 realizations. (OF: Directional asymmetry.)

Number of OF calls	200		500		1000		2500		5000	
	Mean ρ	Std ρ								
No calibration	0.71	0.04	0.71	0.04	0.71	0.04	0.71	0.04	0.71	0.04
FFTMA-SA	0.77	0.04	0.78	0.04	0.80	0.04	0.81	0.05	0.83	0.03
S-STBM	0.75	0.04	0.77	0.04	0.78	0.06	0.81	0.06	0.82	0.02

Table 2
Synthetic case: Correlations of simulated fields with reference field; 100 realizations. (OF : Non-centered third order spatial moments, \vec{h}_1 =all direction, \vec{h}_2 = fixed along East direction for lags 20, 40, 60.)

Number of OF calls	200		500		1000		2 00		5000	
	Mean ρ	Std ρ								
FFTMA-SA	0.79	0.03	0.82	0.02	0.85	0.02	0.87	0.01	0.88	0.01
S-STBM	0.76	0.04	0.79	0.03	0.81	0.03	0.85	0.02	0.87	0.01

4.1. Synthetic case

The first synthetic example (Fig. 5(a)) is taken from Lauzon and Marcotte (2019). It aims to mimic rock silica concentration of volcanic sequences piling along East direction. The dimension of the synthetic field is 100 × 100 cells. Twenty samples are randomly selected for conditioning data (black circles). We compare results of FFTMA-SA algorithm (Lauzon and Marcotte, 2019) to the new S-STBM algorithm. Then, we compare results obtained using either directional asymmetry or non-centered third-order spatial moments.

4.1.1. Comparison of S-STBM with FFTMA-SA

S-STBM is applied on the first synthetic case. The OF is formed by the misfit of the directional asymmetry (Eq. (11)) of the realization and the reference. The stopping criteria is the total number of OF evaluations (here 5000). Fig. 6 presents the variogram along East and South directions and Fig. 7 displays the directional asymmetry along East and South directions. Fig. 6 shows that the calibration performed by S-STBM preserves the variogram reproduction (dark gray) while allowing a good calibration to the directional asymmetry of the reference image (Fig. 7, dark gray: S-STBM realization). Due to the calibration of the random field toward the directional asymmetry, the

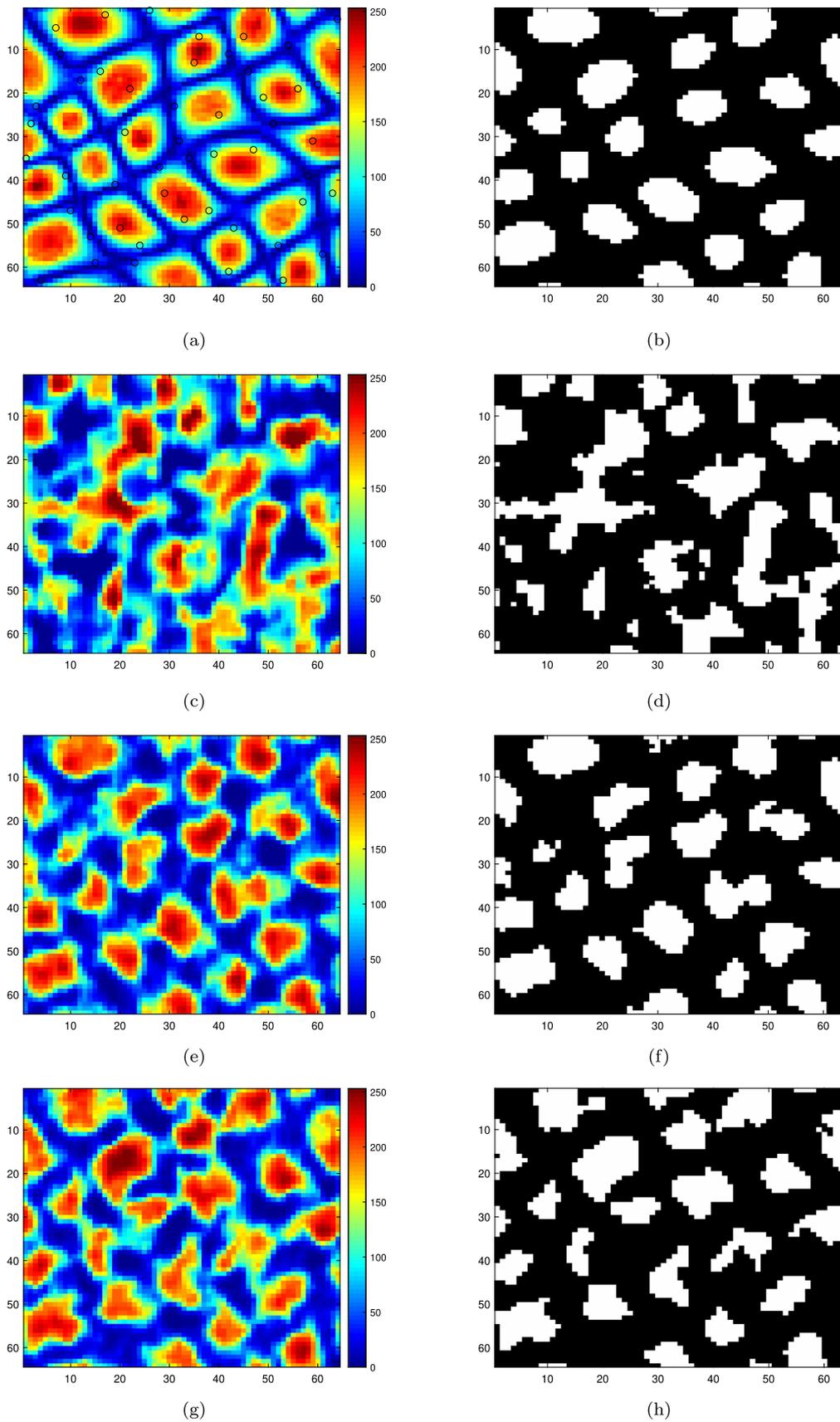


Fig. 8. Left column: Reference and realizations; Right column: thresholded at 70th percentile; First row: Reference field and 50 HD (black circle); Second row: Uncalibrated STBM with 50 HD; Third row: S-STBM with 50 HD; Fourth row: S-STBM with no HD. (20 000 OF calls, third-order spatial moments).

simulated field is no more Gaussian as its third-order statistics are not compatible with Gaussian field. The correlations of the three methods (uncalibrated realization, FFTMA-SA and S-STBM) with the reference have been computed for each of the 100 realizations. The mean correlation and standard deviation of the correlations are shown in Table 1. As expected, calibrated realizations (FFTMA-SA and S-STBM) present a significantly larger correlation than uncalibrated realizations. Also, FFTMA-SA and S-STBM realizations show globally similar correlations. Note that the 5000 iterations correspond for S-STBM to a calibration on 1000 lines with 5 line-search optimization of the phase for each line. For FFTMA-SA, the 5000 iterations correspond to 5000 field perturbations. The perturbations touch simultaneously a number of grid nodes according to a decreasing schedule as described in Lauzon and Marcotte (2019).

4.1.2. Directional asymmetry compared to non-centered high order spatial moments

The same synthetic case is used with OF defined this time by the non-centered third-order moments where the direction vector h_1 sweeps all the field and vector h_2 is oriented East for lags 20, 40 and 60. Again, Table 2 shows comparable results for FFTMA-SA and S-STBM. However, for this example, comparing Table 2 to Table 1 reveals that higher correlations are obtained with the non-centered third-order spatial moment. We stress that increasing the number of conditioning data reduces significantly the differences between the two tables. For example, a test with fifty conditioning data (not shown) presents almost the same correlations for the non-centered third-order spatial moment and the directional asymmetry.

4.2. Continuous image

The continuous “giraffe” image is taken from Wei and Levoy (2000). This TI was used as test case in Rezaee et al. (2015) and Rezaee et al. (2013). The dimension of the TI is 64×64 cells. Fifty samples are randomly selected for conditioning data (black circles). Table 3 presents the correlation between realizations and the reference. Fig. 8 illustrates that calibration to the non-centered third-order moment (e

and f) better reproduce the connectivity of low values present in the TI (a and b) compared to the case with no calibration (c and d). Moreover, the conditioning data helps reproducing a texture more similar to the TI. Fig. 9 shows variogram reproduction.

4.2.1. Further comparison

The continuous reference and realizations obtained with various methods are thresholded at their respective 70th percentile (Fig. 8 b, d, f and h). Some simple geometric characteristics on the white objects are computed: area, length of major axis of an ellipse fitted to the object, eccentricity (ratio of the distance between the foci of the ellipse and its major axis length) and the solidity (area/convex area). The distributions of geometric characteristics in a realization are compared to those in the reference image. The distance between distributions is measured by the Kolmogorov–Smirnov (KS) statistic (Smirnov, 1939). Repeating the exercise for each realization provides a distribution of KS statistics for each measured geometric characteristic and simulation method.

Fig. 10 shows the boxplots obtained for the different geometric characteristics. When the two-sample KS statistic (D^*) is above the critical value, there is significant difference between the TI and the realization. The green lines represents the 5th and 95th percentile of the critical value at level $\alpha = 0.05$ (the critical values varies slightly for the different realizations because it depends on the number of objects found in the realizations). The boxplots of geometric characteristics reflecting the size of the objects (area and major axis) (Fig. 10a–f) show a regular decrease of D^* until the 95th percentile settles close to its expected location in the critical zone. However, the boxplots of geometric characteristics reflecting the shape of the objects (eccentricity, solidity) show no particular decrease of D^* (Fig. 10g–l). Comparing the left and middle columns of Fig. 10 shows that conditioning data (left column) are useful to bring size geometrical characteristics of the realizations closer to the ones of the TI. In particular, for $N \geq 1000$ lines the 95th percentiles in (a) and (d) are closer to the critical zone compared to (b) and (e). On the other hand, shape characteristics do not reveal any clear improvement due to conditioning data.

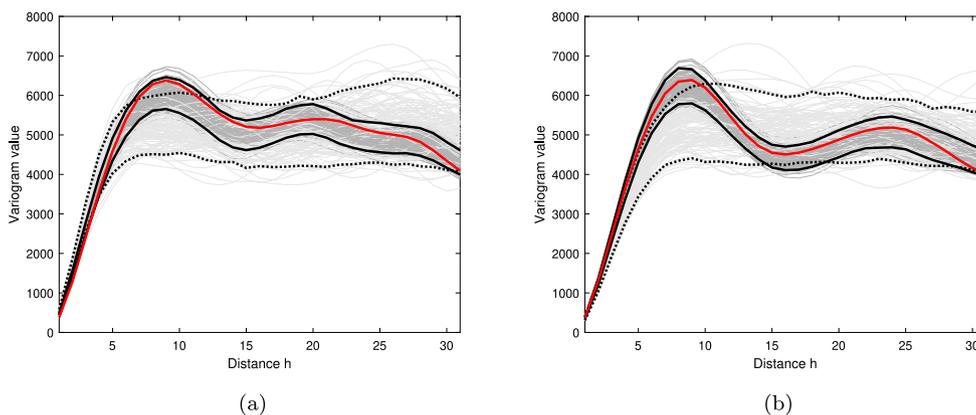


Fig. 9. “Giraffe” TI. (a) Variograms along East direction, (b) along South direction. Reference field (solid red), S-STBM realizations (dark gray), uncalibrated STBM (light gray); 90% confidence intervals shown (solid dark: calibrated and dashed lines: uncalibrated). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3

TI: Correlations of simulated fields with reference field; 100 realizations. OF: non-centered third order spatial moments with \vec{h}_2 in 9 directions ((0,0), (0,8), (8,0), (6,6), (-3,3), (2,6), (6,2), (-2,6),(-6,2)) and all available \vec{h}_1 .

Method	1000		2500		5000		10 000		20 000	
	Mean ρ	Std ρ								
No calibration	0.23	0.06	0.23	0.06	0.23	0.06	0.23	0.06	0.23	0.06
FFTMA-SA	0.55	0.04	0.68	0.03	0.76	0.02	0.82	0.01	0.85	0.01
S-STBM	0.36	0.05	0.59	0.05	0.77	0.02	0.82	0.01	0.86	0.01

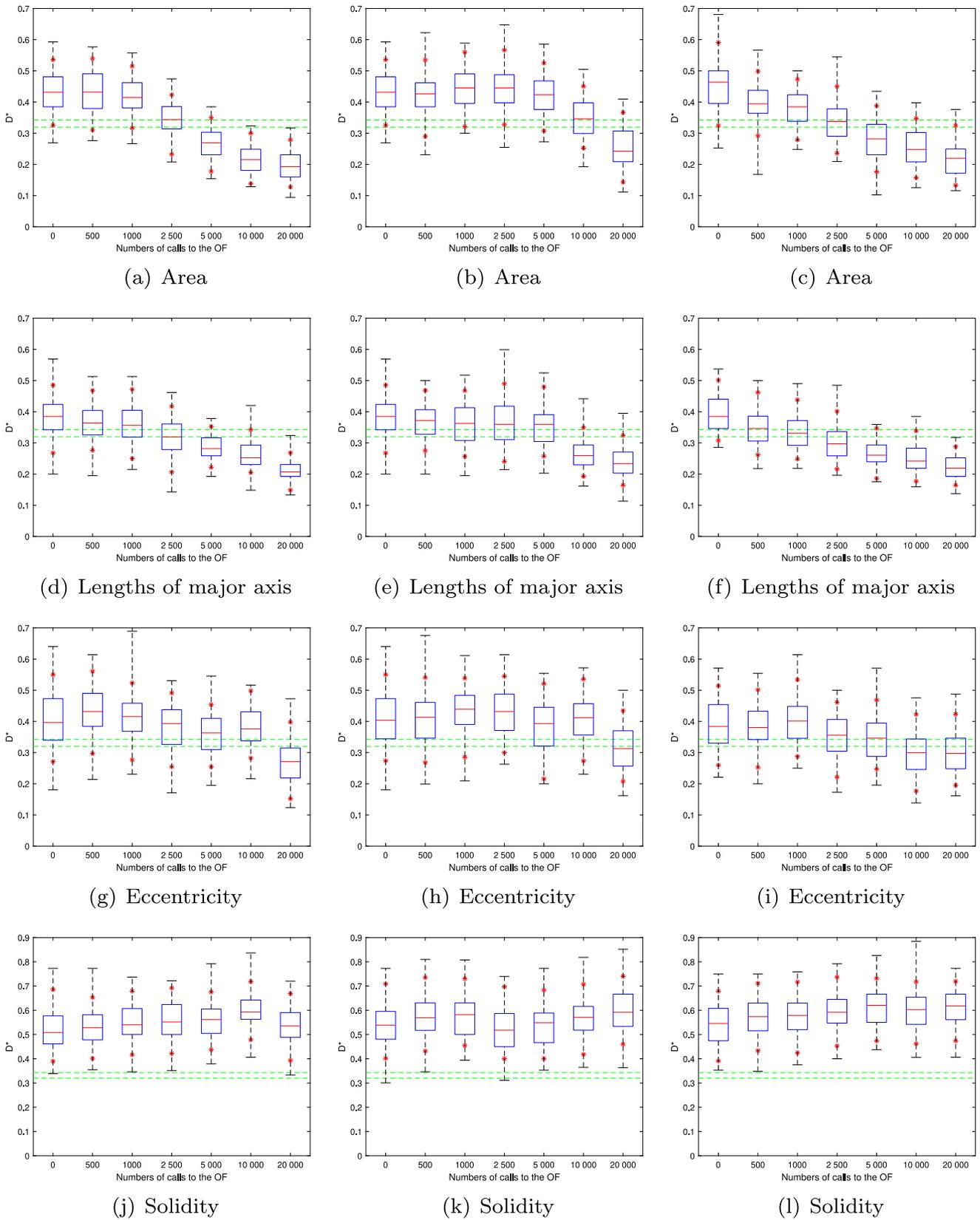


Fig. 10. Two-sample KS statistic (D^*) boxplots for different numbers of calibrated lines. First column: S-STBM with 50 HD; Second column: S-STBM with no HD; Third column: FFTMA-SA with 50 HD (same number of OF calls as for S-STBM). Green dashed lines: 5th and 95th percentiles of critical values at level $\alpha = 0.05$. Red asterisks: 5th and 95th percentiles of (D^*) distribution. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.3. Tracer test: travel time between wells

A synthetic Gaussian rock porosity field (Fig. 11(a)) is used to calibrate tracer tests using the Matlab Reservoir Simulation Toolbox (MSRT) (Lie, 2019). The wells are distributed into a quarter five-spot arrangement constituted of one injector well and four production wells (P1 to P4). The tracer is injected at flowrate of $0.0018 \text{ m}^3/\text{s}$ and is recovered at the production wells (P1 to P4). The dimension of the porosity field is 100×100 cells. Ten samples are randomly selected for the conditioning data (black circles). The OF is formed by the mean-squared error of the four travel times between wells of the reference and the realization (time for a particle to migrate from the injection well to a production well). The stopping criteria is the total number of OF evaluations (here 500 OF calls for 100 calibrated lines). Fig. 12 presents the variogram along East and South directions and Table 4 displays the effectiveness of the S-STBM algorithm to calibrate the tracer tests. The travel times between wells of the two S-STBM realizations (calibrated) are much closer to the reference than the two uncalibrated STBM realizations. Furthermore, the porosity spatial

Table 4

Tracer test : first-arrival traveltimes between injector and producers (P1 to P4).

	P1	P2	P3	P4
Reference	13.28	15.56	13.84	17.32
STBM No. 1	15.02	14.65	14.16	16.16
STBM No. 2	15.03	14.53	13.24	17.20
S-STBM No. 1	13.27	15.57	13.83	17.32
S-STBM No. 2	13.31	15.56	13.84	17.30

distribution of S-STBM realizations looks more like the reference than the STBM realizations (see bottom left corner of Fig. 11).

5. Discussion

A new algorithm is proposed for the calibration of random fields to any desired response functions. S-STBM works sequentially by selecting the phase on each line such as to optimize an OF measuring the departure of the realization to the field response. The two examples

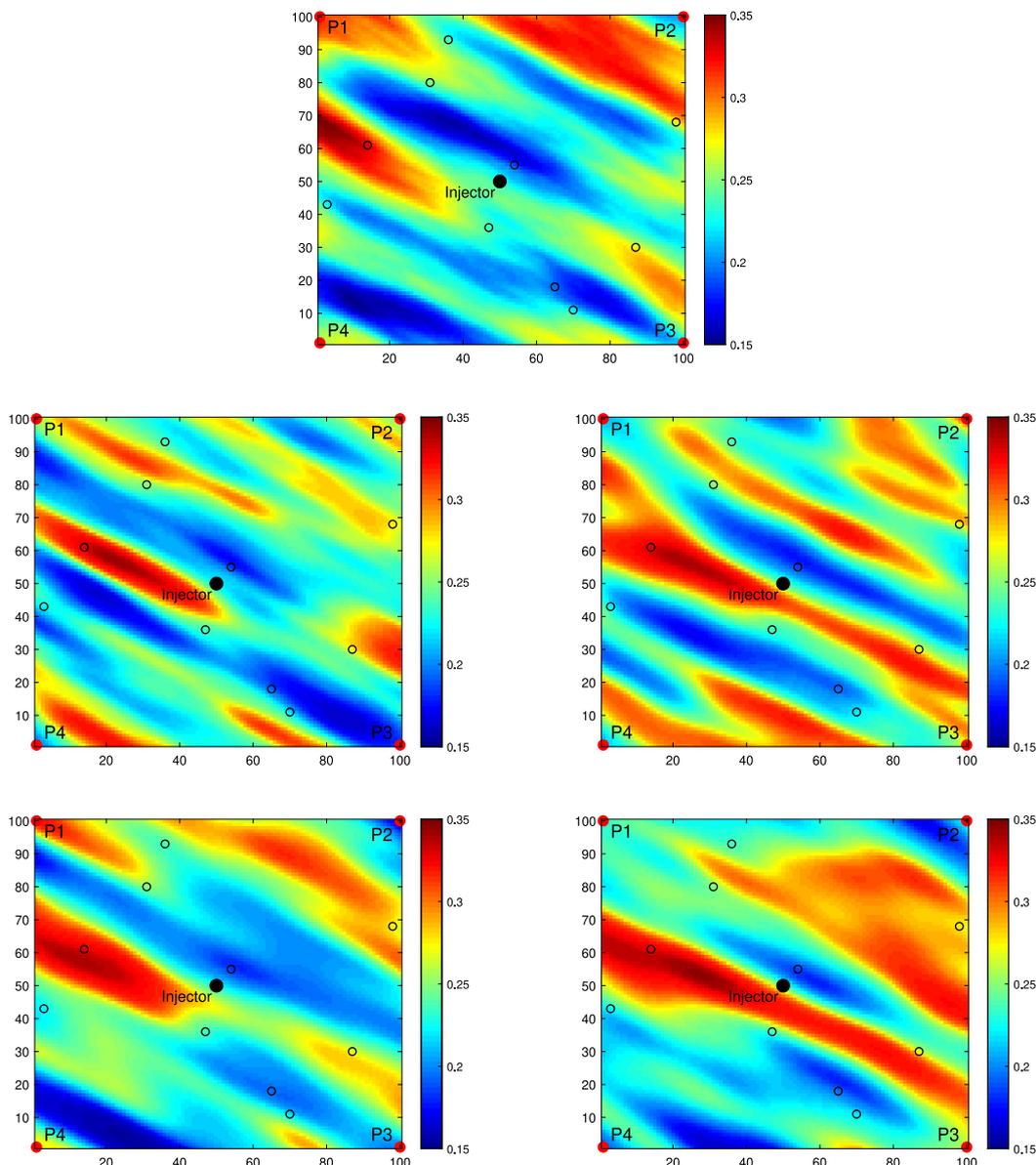


Fig. 11. Top: Reference porosity field, Middle: STBM realizations, uncalibrated, Bottom: S-STBM realizations, calibrated. (Red dots: production well, Black dot: injection well, black circles: conditioning data). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

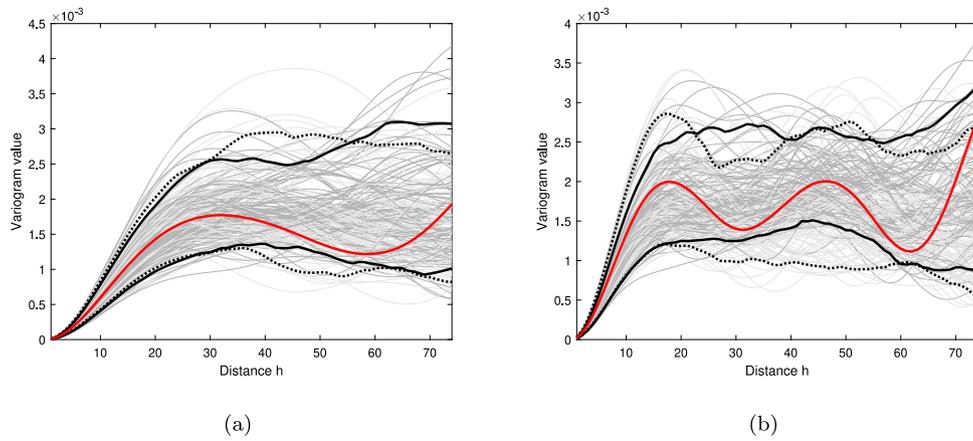


Fig. 12. Tracer test. (a) Variograms along East direction, (b) along South direction. Reference field (solid red), S-STBM realizations (dark gray), uncalibrated STBM (light gray); 90% confidence intervals shown (solid dark: calibrated and dashed lines: uncalibrated). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

presented, the synthetic case and the continuous ‘giraffe’ training image, show the ability of S-STBM to match well the realizations to the corresponding OF for a large number of iterations. If needed, the iterative process can be stopped early to increase variability between realizations. Fig. 13 shows that variability of calibrated realizations decrease regularly with the number of iterations while embedding early the true response. This good behavior permits choosing the number of iterations accounting for the uncertainty in the response function and thus to avoid overfitting the OF.

The main advantages of S-STBM over PA, FFTMA-SA or FFTMA-GD are that S-STBM can generate realizations on an irregular grid and has no memory limitations. Furthermore, FFTMA-SA and PA need a

cooling schedule for the SA process. For S-STBM, no parameter need to be specified else than the number of iterations. However, S-STBM, like PA, works globally contrary to FFTMA-SA and FFTMA-GD which both permit local or global calibration.

The continuous TI and realizations were further compared on the basis of distributions of object sizes and shapes they define after thresholding to their respective 70th percentile. We stress that the objects geometric characteristics were not included in the OF, only third order moments were used. Two-sample Kolmogorov–Smirnov (KS) statistic (Smirnov, 1939) was used to measure similarity of distributions. The size characteristics of the objects converged well to the size of objects found in the TI. However, the shape characteristics of objects

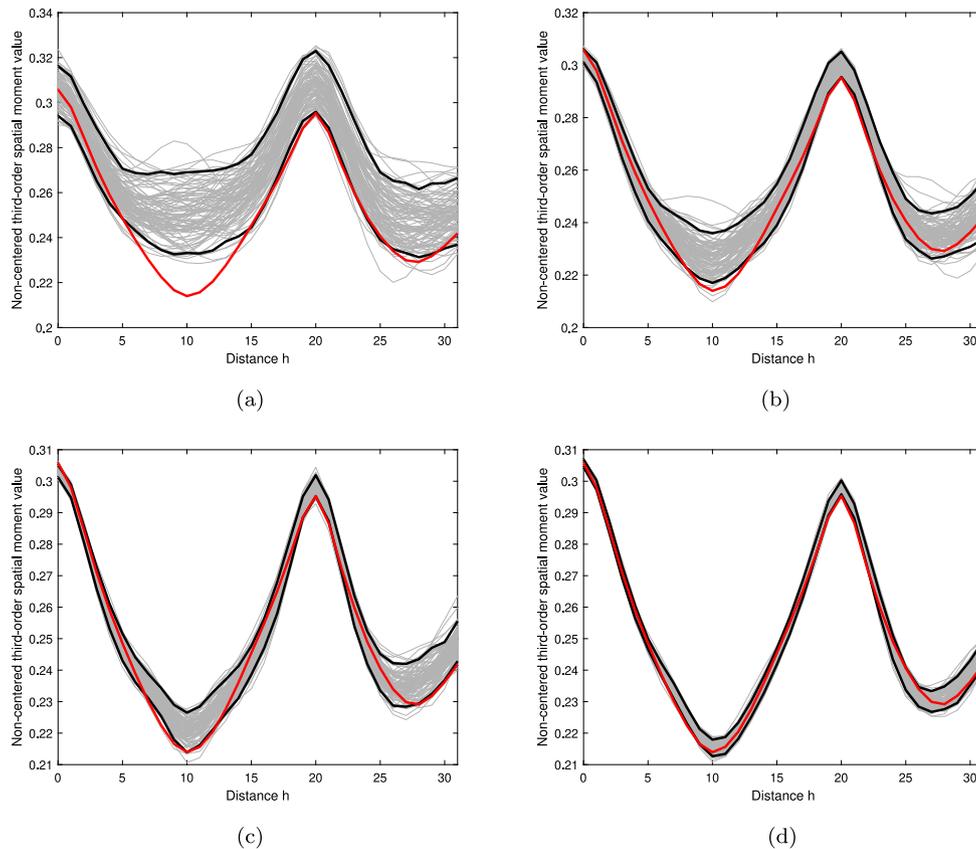


Fig. 13. Non-centered third-order spatial moment along East direction for $|h_2| = 20$. S-STBM with (a) 0, (b) 200, (c) 500, (d) 2000 calibrated lines. Reference field (solid red), 90% confidence interval (solid dark). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

were not well reproduced after calibration. This indicates the OF using third order moments is probably not rich enough in shape information. Using directly shape information in the OF is a possible option that warrants further research.

We note that, as for most iterative algorithms, convergence is not guaranteed with our sequential method. There is the possibility that the first iterations of calibration brings a state corresponding to a local minimum in the OF from which it will not be possible to escape. The methods using simulated annealing (PA, FFTMA-SA) have theoretical convergence guaranteed when the cooling schedule is slow enough but in practice too slow to be applicable (Geman and Geman, 1984).

The first two case studies present non-Gaussian high-order characteristics. Although the STBM algorithm (also TBM algorithm) simulates Gaussian fields, our algorithm allows to introduce non-Gaussian characteristics by the calibration process. We stress that other methods to simulate random fields with non-Gaussian characteristics exist. For example, multiple-point simulation (MPS) allow these characteristics to be directly integrated into a training image (TI). However conditioning to (continuous) HD with MPS is difficult. Moreover meaningful TIs can be difficult to obtain especially in 3D and MPS is confined to regular grids. On the contrary, our algorithm allows to integrate exactly a large amount of conditioning data on unstructured grid. It has a computational complexity of $\mathcal{O}(n)$ and small memory requirement allowing to perform simulation over very large fields. The idea of using a Gaussian simulation method to generate calibrated fields with non-Gaussian characteristics is akin to the usual practice where non-Gaussian features present in data are *de facto* transmitted to the realizations by the conditioning process.

A last remark concerns the distribution of the phase component after optimization. Eq. (4) requires U being uniformly distributed. Indeed, the distribution of U after optimization can depart from an uniform distribution. Hence, calibration can have some effect on the covariance function much like addition of a few HD also impacts the covariance reproduction. In our examples, this conditioning effect of calibration was minor as the realization variograms remained close to the experimental variograms (see Figs. 6, 9 and 12).

6. Conclusion

The S-STBM algorithm can calibrate realizations to various objective functions without disturbing significantly the field covariance of the calibrated fields and still ensuring exact HD reproduction. Two synthetic cases showed improved correlations of calibrated field with the reference field. Comparisons showed that S-STBM performances are similar to those of FFTMA-SA for a same number of OF evaluations. However, contrary to FFTMA-SA, S-STBM can handle arbitrary large simulated fields on both regular or irregular grids. Thanks to its $\mathcal{O}(n)$ complexity and small memory requirement, S-STBM constitutes a valuable tool for calibration.

Computer code availability

The Matlab computer codes are available at <https://github.com/Danlauz/S-STBM-Programs>.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRedit authorship contribution statement

Dany Lauzon: Methodology, Programming, Validation, Initiale Writing, Data curation. **Denis Marcotte:** Research idea, Methodology, Supervised the research, Programming, Final Writing.

Acknowledgments

This research was made possible by National Research Council of Canada thru NSERC grant (RGPIN-2015-06653) of D. Marcotte and NSERC grant (CGS M) of D. Lauzon.

References

- Bárdossy, A., Hörning, S., 2017. Process-driven direction-dependent asymmetry: Identification and quantification of directional dependence in spatial fields. *Math. Geosci.* 49 (7), 871–891. <http://dx.doi.org/10.1007/s11004-017-9682-1>.
- Chilès, J., Delfiner, P., 2012. *Geostatistics: Modeling Spatial Uncertainty*, second ed. John Wiley & Sons.
- Delopoulos, A., Tirakis, A., Kollias, S., 1994. Invariant image classification using triple-correlation-based neural networks. *IEEE Trans. Neural Netw.* 5 (3), 392–408. <http://dx.doi.org/10.1109/72.286911>.
- 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. *Math. Geosci.* 42 (1), 65–99. <http://dx.doi.org/10.1007/s11004-009-9258-9>.
- Emery, X., Arroyo, D., Porcu, E., 2015. An improved spectral turning-bands algorithm for simulating stationary vector Gaussian random fields. *Stoch. Environ. Res. Risk Assess.* 30 (7), 1863–1873. <http://dx.doi.org/10.1007/s00477-015-1151-0>.
- Emery, X., Lantuéjoul, C., 2006. TBSIM: A computer program for conditional simulation of three-dimensional Gaussian random fields via the turning bands method. *Comput. Geosci.* 32 (10), 1615–1628.
- Evensen, G., 2009. *Data Assimilation*. Springer, Berlin, Heidelberg, <http://dx.doi.org/10.1007/978-3-642-03711-5>.
- Freulon, X., de Fouquet, C., 1991. Remarques sur la pratique des bandes tournantes à trois dimensions. In: *Cahiers de géostatistique, Fascicule 1, Centre de Géostatistique, École des Mines de Paris, Fontainebleau*. pp. 101–117.
- Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intel. PAMI-6* (6), 721–741. <http://dx.doi.org/10.1109/tpami.1984.4767596>.
- Gómez-Hernández, J.J., Sahuquillo, A., Capilla, J., 1997. Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data—I. Theory. *J. Hydrol.* 203 (1–4), 162–174. [http://dx.doi.org/10.1016/S0022-1694\(97\)00098-X](http://dx.doi.org/10.1016/S0022-1694(97)00098-X).
- Guthke, P., Bárdossy, A., 2017. On the link between natural emergence and manifestation of a fundamental non-Gaussian geostatistical property: Asymmetry. *Spat. Stat.* 20, 1–29. <http://dx.doi.org/10.1016/j.spasta.2017.01.003>.
- Halton, J.H., 1964. Algorithm 247: Radical-inverse quasi-random point sequence. *Commun. ACM* 7 (12), 701–702. <http://dx.doi.org/10.1145/355588.365104>.
- Horikawa, Y., 2000. Bispectrum-based feature of 2D and 3D images invariant to similarity transformations. In: *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, vol. 2. IEEE Comput. Soc., pp. 511–514. <http://dx.doi.org/10.1109/icpr.2000.906124>.
- Hörning, S., Bárdossy, A., 2018. Phase annealing for the conditional simulation of spatial random fields. *Comput. Geosci.* 112, 101–111. <http://dx.doi.org/10.1016/j.cageo.2017.12.008>.
- Hu, L.Y., Ravalec-Dupin, M.L., 2004. An improved gradual deformation method for reconciling random and gradient searches in stochastic optimizations. *Math. Geol.* 36 (6), 703–719. <http://dx.doi.org/10.1023/b:matg.0000039542.73994.a2>.
- Hu, L., Zhao, Y., Liu, Y., Scheepens, C., Bouchard, A., 2013. Updating multipoint simulations using the ensemble Kalman filter. *Comput. Geosci.* 51, 7–15. <http://dx.doi.org/10.1016/j.cageo.2012.08.020>.
- Journel, A., 1993. *Geostatistics: Roadblocks and Challenges*. In: Soares A. (eds) *Geostatistics Tróia '92*, vol. 5. Quantitative Geology and Geostatistics.
- Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. *Science* 220 (4598), 671–680. <http://dx.doi.org/10.1126/science.220.4598.671>.
- Lantuéjoul, C., 2002. *Geostatistical Simulation*. Springer, Berlin, Heidelberg.
- Lauzon, D., Marcotte, D., 2019. Calibration of random fields by FFTMA-SA. *Comput. Geosci.* 127, 99–110. <http://dx.doi.org/10.1016/j.cageo.2019.03.003>, URL: <http://www.sciencedirect.com/science/article/pii/S0098300418312044>.
- Le Ravalec-Dupin, M., Hu, L.Y., 2005. Gradual deformation of boolean simulations. In: Leuangthong, O., Deutsch, C. (Eds.), *Geostatistics Banff 2004*. Springer, pp. 939–948.
- Lie, K., 2019. *An Introduction to Reservoir Simulation Using MATLAB/GNU Octave: User Guide for the MATLAB Reservoir Simulation Toolbox (MRST)*. Cambridge University Press.
- Liu, N., Oliver, D.S., 2005. Ensemble Kalman filter for automatic history matching of geologic facies. *J. Petrol. Sci. Eng.* 47 (3–4), 147–161. <http://dx.doi.org/10.1016/j.petrol.2005.03.006>.
- Luenberger, D., Ye, Y., 2008. *Linear and Nonlinear Programming*. Springer US, URL: https://www.ebook.de/de/product/19292757/david_g_luenberger_yinyu_ye_linear_and_nonlinear_programming.html.
- Marcotte, D., 1996. Fast variogram computation with FFT. *Comput. Geosci.* 22 (10), 1175–1186. [http://dx.doi.org/10.1016/S0098-3004\(96\)00026-X](http://dx.doi.org/10.1016/S0098-3004(96)00026-X).

- Marcotte, D., 2015. TASC3D : A program to test the admissibility in 3D of non-linear models of coregionalization. *Comput. Geosci.* 83, 168–175. <http://dx.doi.org/10.1016/j.cageo.2015.07.012>.
- Marcotte, D., 2016. Spatial turning bands simulation of anisotropic non-linear models of coregionalization with symmetric cross-covariances. *Comput. Geosci.* 89, 232–238. <http://dx.doi.org/10.1016/j.cageo.2016.01.004>.
- Marcotte, D., Allard, D., 2017. Half-tapering strategy for conditional simulation with large datasets. *Stoch. Res. Risk Assess.* <http://dx.doi.org/10.1007/s00477-017-1386-z>.
- Marsily, G., Lavedan, G., Boucher, M., Fasanino, G., 1984. Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model. *Geostat. Nat. Resour. Charact. Part 2*, 831–849.
- Matheron, G., 1973. The intrinsic random functions and their applications. *Adv. Appl. Probab.* 5 (3), 439–468. <http://dx.doi.org/10.2307/1425829>.
- Oliver, D.S., Chen, Y., Nævdal, G., 2010. Updating Markov chain models using the ensemble Kalman filter. *Comput. Geosci.* 15 (2), 325–344. <http://dx.doi.org/10.1007/s10596-010-9220-4>.
- Oliver, D.S., Cunha, L.B., Reynolds, A.C., 1997. Markov chain Monte-Carlo methods for conditioning a permeability field to pressure data. *Math. Geol.* 29 (1), 61–91. <http://dx.doi.org/10.1007/bf02769620>.
- Räss, L., Kolyukhin, D., Minakov, A., 2019. Efficient parallel random field generator for large 3-D geophysical problems. *Comput. Geosci.* <http://dx.doi.org/10.1016/j.cageo.2019.06.007>.
- Renard, P., Allard, D., 2013. Connectivity metrics for subsurface flow and transport. *Adv. Water Resour.* 51, 168–196. <http://dx.doi.org/10.1016/j.advwatres.2011.12.001>.
- Rezaee, H., Marcotte, D., 2018. Calibration of categorical simulations by evolutionary gradual deformation method. *Comput. Geosci.* 22 (2), 587–605. <http://dx.doi.org/10.1007/s10596-017-9711-7>.
- Rezaee, H., Marcotte, D., Tahmasebi, P., Saucier, A., 2015. Multiple-point geostatistical simulation using enriched pattern databases. *Stoch. Environ. Res. Risk Assess.* 29 (3), 893–913.
- Rezaee, H., Mariethoz, G., Koneshloo, M., Asghari, O., 2013. Multiple-point geostatistical simulation using the bunch-pasting direct sampling method. *Comput. Geosci.* 54, 293–308. <http://dx.doi.org/10.1016/j.cageo.2013.01.020>.
- Shinozuka, M., 1971. Simulation of Multivariate and Multidimensional Random Processes. *J. Acoust. Soc. Am.* 49 (1B), 357–368. <http://dx.doi.org/10.1121/1.1912338>.
- Shinozuka, M., Jan, C., 1972. Digital simulation of random processes and its applications. *J. Sound Vib.* 25, 111–128.
- Smirnov, N., 1939. On the estimation of the discrepancy between empirical curves of distribution for two independent samples. *Moscow Univ. Math. Bull.* 2, 2–14.
- Sudicky, E.A., 1986. A natural gradient experiment on solute transport in a sand aquifer: spatial variability of hydraulic conductivity and its role in the dispersion process. *Water Resour. Res.* 22 (13), 2069–2082. <http://dx.doi.org/10.1029/wr022i013p02069>.
- Tompson, A.F.B., Ababou, R., Gelhar, L.W., 1989. Implementation of the three-dimensional turning bands random field generator. *Water Resour. Res.* 25 (10), 2227–2243. <http://dx.doi.org/10.1029/wr025i010p02227>.
- Tsatsanis, M., Giannakis, G., 1992. Object and texture classification using higher order statistics. *IEEE Trans. Pattern Anal. Mach. Intell.* 14 (7), 733–750. <http://dx.doi.org/10.1109/34.142910>.
- van der Corput, J., 1935. Verteilungsfunktionen. *Akad. Wet.* (38), 813–821.
- Wei, L.-Y., Levoy, M., 2000. Fast texture synthesis using tree-structured vector quantization. In: *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH '00*. ACM Press, <http://dx.doi.org/10.1145/344779.345009>.