Computers and Geosciences 127 (2019) 99–110

Contents lists available at ScienceDirect

Computers and Geosciences

journal homepage: www.elsevier.com/locate/cageo

Research paper

Calibration of random fields by FFTMA-SA

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:
Calibration
FFTMA
Simulated annealing
Conditional simulation
Local perturbation

ABSTRACT

A new algorithm is presented aiming to generate realizations calibrated to desired or observed field response functions. The Fast Fourier Transform Moving Average - simulated annealing algorithm (FFTMA-SA) combines the FFTMA method of simulation and the simulated annealing (SA) method for calibration. SA is performed on the uncorrelated random numbers used within the FFTMA algorithm, keeping by construction the field covariance function unchanged. Hard data exact conditioning by kriging is easily introduced into the algorithm main loop. An interesting advantage of FFTMA-SA is the capability to perturb the field locally or globally, simultaneously on any desired number of points. We therefore introduce a decreasing schedule on the number of points to perturb mimicking the cooling schedule of the temperature parameter in SA. Two cases studies are presented. The first one is a spatially asymmetric synthetic field where FFTMA-SA is shown to better reproduce the spatial asymmetry than calibrated realizations obtained either by classical SA or by FFTMA combined with gradual deformation. The second example is the Walker Lake data where regularized row and columns facies proportions are imposed. It is shown that few iterations (100) of FFTMA-SA are enough to reproduce the main characteristics of the proportion curves. Increasing the number of iterations only reduces rather uniformly the variability of realization proportions. The same data is used to illustrate the local calibration capability of the method with a local data assimilation example. The proposed algorithm offers a simple and flexible tool for calibration. It is easily adaptable for any moving average method of simulation.

1. Introduction

There exist numerous geostatistical methods aiming to simulate conditional Gaussian fields with prescribed covariance function (Chilès and Delfiner, 2012; Lantuéjoul, 2002). However, one common difficulty arising in practical applications is to ensure realizations do comply with additional information that might be available. Typical examples of such information are connectivity measures (Renard and Allard, 2013), third or higher order spatial statistics (Hörning and Bárdossy, 2018), proportions above thresholds, travel time between wells in tracer tests (Sudicky, 1986), pressure drop or water cut in petroleum wells (see Oliver et al. (1997); Ravalec-Dupin (2005); Hu et al. (2013); Oliver et al. (2010); Rezaee and Marcotte (2018), among many others) and, more generally, any measure that relates non-linearly to the simulated field properties or that are simply not compatible with a Gaussian field. The enforcement of additional field characteristics to comply to non-linearly related data is known under various names like history matching, data assimilation and inversion. For simplicity, all these variants are identified hereafter under the generic name of calibration. Available calibration methods can require optimization of an objective function (OF) or be based on sequential updating of the field (e.g. history matching with Ensemble Kalman Filtering (Evensen, 2009)). Optimization methods can be gradient based (e.g. pilot or master point method (de Marsily et al., 1984; Gómez-Hernánez et al., 1997)) or based on field perturbation (Deutsch, 1992; Hu, 2000; Rezaee and Marcotte, 2018), or sometimes a combination of both (Hu and Ravalec-Dupin, 2004). Desired characteristics of the method used for calibration are the following: realistic computing times, reproduction of hard data (HD), preservation of the field spatial covariance and capacity to act locally on the field. We stress that calibration is understood here as any optimization method seeking to optimize parameters of the field so as to approximate any static data or dynamic response measured directly or indirectly on the field.

One commonly used iterative calibration method is simulated annealing (SA) (Kirkpatrick et al., 1983; Geman and Geman, 1984; Deutsch, 1992; Deutsch and Journel, 1994). SA suffers a few drawbacks: it is usually slow to converge and hard data points (HD) can be poorly embedded with the neighbouring cells (Hörning and Bárdossy, 2018) creating undesired discontinuities. Hörning and Bárdossy (2018) and Yao (1998) proposed instead to apply SA on the phase component of the Fourier transform of the simulated field. This method preserves the spatial covariance of the field at all iterations and ensure proper HD

* Corresponding author.
E-mail address: denis.marcotte@polymtl.ca (D. Marcotte).

https://doi.org/10.1016/j.cageo.2019.03.003
Received 19 December 2018; Received in revised form 3 March 2019; Accepted 9 March 2019
Available online 14 March 2019
0098-3004/

embedding with its neighbours. However, in the phase annealing (PA) method, conditioning to HD is ensured by inserting a HD-misfit component in the objective function (OF). Hence, the HD conditioning is only approximate. Moreover, PA is a global method as any phase modification affects simultaneously the entire field. It seems preferable to adopt a method allowing exact conditioning and capability to perturb field locally as well as globally.

A simulation method that allows local modifications is the FFTMA method of Le Ravalec et al. (2000) and more generally moving average methods (Chilès and Delfiner, 2012). In these methods, the weighted average of independent Gaussian variables is done over a window (circular, elliptic or else) which size is a function of the correlation ranges of the covariance components. Hence, modifying a single random variable mostly modify the area around the perturbed cell. Hu and Ravalec-Dupin (2004) proposed to combine FFTMA with gradual deformation method (FFTMA-GD) for inversion of permeability field from pressure data. One possible drawback of FFTMA-GD is the tendency of the OF to stabilize after a few iterations. As the convergence properties of SA are established for a sufficiently slow cooling schedule (Geman and Geman, 1984), it might be profitable to replace GD by SA in conjunction with FFTMA. This contribution compares performances of FFTMA-SA to SA and FFTMA-GD and illustrate the main advantages and possible weaknesses of the proposal.

The methodology section recalls briefly the main characteristics of FFTMA and SA. Then, FFTMA-SA algorithm is presented. A synthetic case study showing clear spatial asymmetry is used to illustrate the use of FFTMA-SA. Results of FFTMA-SA are compared to SA, uncalibrated conditional realizations and FFTMA-GD. A second case study with Walker Lake data set compares uncalibrated conditional realizations and FFTMA-GD to FFTMA-SA for calibration to target vertical and horizontal facies proportions. The sensitivity to maximum number of iterations used is examined. A second application with the Walker Lake data illustrates how newly acquired data in a subarea can easily be assimilated by FFTMA-SA. A general discussion underlining the pros and cons of the proposed method follows.

2. Methodology

This section presents briefly the main characteristics of SA and FFTMA that are then combined for calibration purpose.

2.1. Simulated annealing (SA)

Kirkpatrick et al. (1983) and Geman and Geman (1984) introduced simulated annealing (SA), a general purpose optimization algorithm based on iterative perturbations of an existing field. The perturbations are accepted at iteration i when improving the objective function (i.e $O_i < O_{i-1}$) and accepted with probability $exp([O_{i-1} - O_i]/T_i)$ when deteriorating the objective (i.e. $O_i > O_{i-1}$). The temperature parameter T_i is decreased slowly with iterations such as to gradually render less likely a deterioration of the OF as the iterations progress. The slower the temperature decreases, the higher is the likelihood of convergence to the global minimum but the larger is the number of iterations required to eventually reach it. Note that perturbations in SA can be brought to one, two or more cells at a time. The algorithm runs until a suitable stopping criterion is reached.

As always in optimization, many stopping criteria can be adopted like the maximum number of iterations (our choice), the rate of acceptance of perturbations, the rate of decrease of the OF, or the reaching of a threshold on OF. The stopping criteria depends of the quality, precision and confidence in data to calibrate, of the goals of the study and of the available computing power and time. Although SA has good convergence properties (Geman and Geman, 1984), its weakness is to typically require many iterations. It is therefore considered a slow method. In cases where OF evaluation is costly in CPU time it can even become inapplicable in practice.

2.2. Fast Fourier transform - moving average (FFTMA)

The FFTMA method was originally proposed by Le Ravalec et al. (2000) as an efficient mean to perform unconditional Gaussian simulations on a regular grid of points. It seeks to simplify the determination of the weighting function g in moving average methods (Oliver, 1995; Chilès and Delfiner, 2012). The moving-average method simulates a Gaussian field Z by convolution of a suitably chosen weighting function g with a Gaussian white noise Y. The well-known convolution theorem (Sneddon, 1951) states that the convolution in spatial domain becomes a simple product in the spectral domain. Hence, writing $Z = Y * g$, one has $\mathcal{F}(Z) = \mathcal{F}(Y)\mathcal{F}(g)$. Assuming that $g(h)$ is symmetric around the origin, one has:

$$C = g * g \tag{1}$$

Where C is the covariance function. Hence,

$$\mathcal{F}(C) = \mathcal{F}(g)\mathcal{F}(g) = \mathcal{F}(g)^2 \rightarrow \mathcal{F}(g) = \mathcal{F}(C)^{0.5} \tag{2}$$

where $\mathcal{F}$ represents the Fourier transform. The weighting function g does not have to be found explicitly. It suffices to know the FFT of C. Note that C being real and symmetrical around the origin, $\mathcal{F}(C)$, $\mathcal{F}(g)$ and, consequently the function g are also real and symmetric functions.

Following Le Ravalec et al. (2000) the FFTMA simulation method comprises the following steps:

- Computation of covariance C over the regular grid with respect to the grid origin.
- Generation of the Gaussian white noise Y on the grid.
- Calculation of the Fast Fourier transforms $\eta = \mathcal{F}(Y)$ and $S = \mathcal{F}(C)$.
- Computation of $G = \sqrt{S}$.
- Computation of $\mathcal{F}(Z) = \eta G$.
- Inverse Fourier transform of $\mathcal{F}(Z)$ giving the correlated Gaussian field $Z(x)$ with covariance C.

To account for periodicity of the FFT, C must be periodized over the field (see Le Ravalec et al. (2000); Chilès and Delfiner (2012) for details) and the simulated field must be embedded in a field extended by the finite range in all directions. For covariances with asymptotic ranges, the field must be enlarged enough to bring the covariance very close to zero (Chilès and Delfiner, 2012) to minimize aliasing. Two important drawbacks of FFTMA are worth mentioning. First, it works on regular grids. This can limit the applicability of the method, especially in 3D where data are located along boreholes or in 2D for highly clustered data. Second, it can be limited by computer memory, for example in 3D for covariances with a large range. The FFTMA computational complexity is $O(Nlog(N))$ compared to $O(N)$ for Turning Bands (TB). Our own experience however is that FFTMA can be faster than TB for small to medium size fields.

The strength of FFTMA is its flexibility and fast calculation. Any covariance model with finite range can be simulated exactly on the grid. The method clearly separates the random component Y from the deterministic weighting function g. The FFT of g can be computed once. Modifications can be brought to one or many random numbers to create new realizations without altering the simulated covariance.

2.3. FFTMA-SA

The random field of FFTMA being uncorrelated, any number of cells can be modified simultaneously without altering the spatial covariance, much like does phase annealing (PA) (Hörning and Bárdossy, 2018; Yao, 1998). This helps speeding convergence compared to classical SA. In our implementation of the method, we start with a large number of cells to perturb and decrease this number regularly up to a single cell according to a decreasing schedule inspired from temperature cooling schedule in SA. Secondly, the perturbed cells can be selected over the entire field or be concentrated in a particular subarea to leave

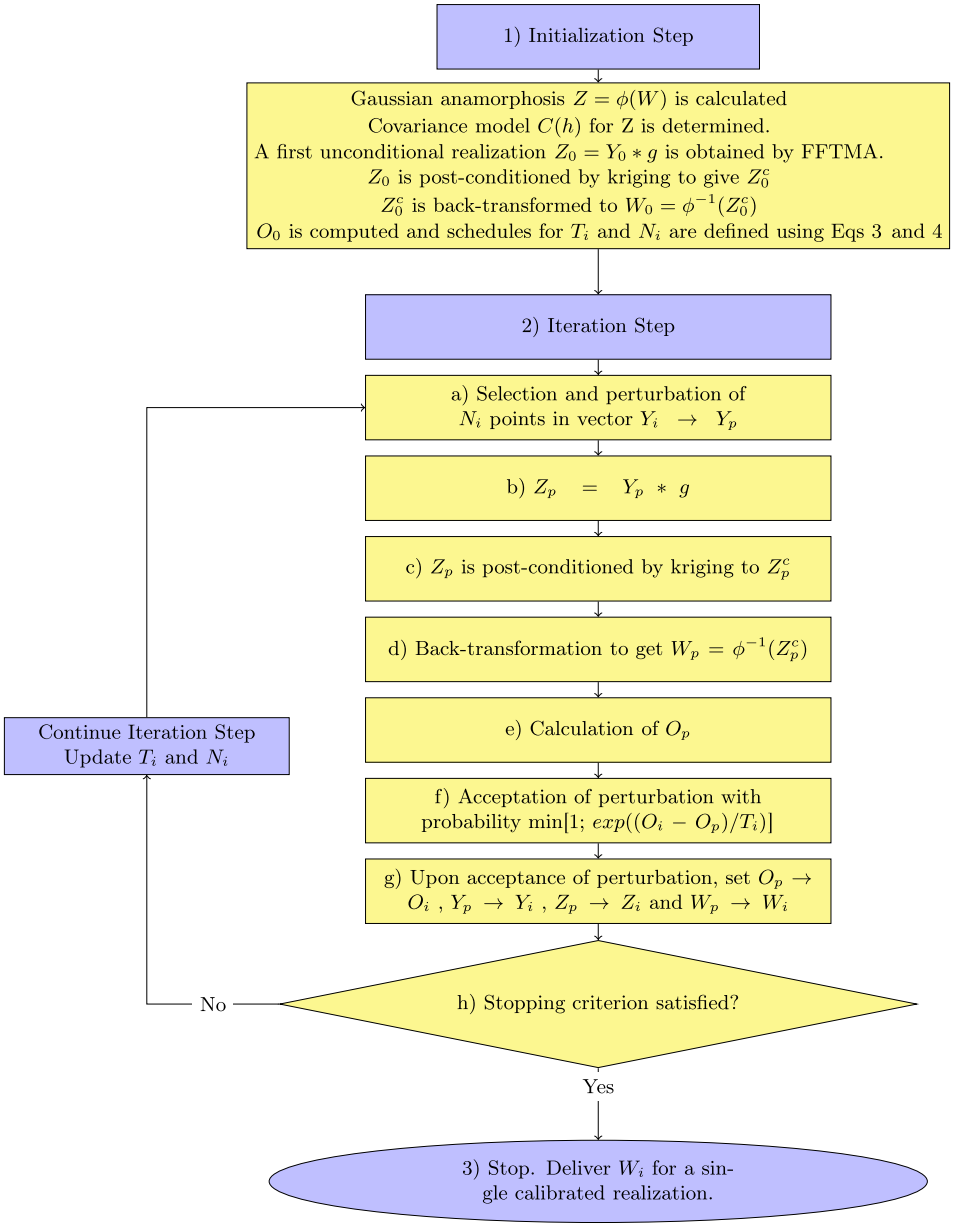


Fig. 1. Flowchart of FFTMA-SA algorithm. Z is N(0,1), W is the variable to simulate, ϕ is the Gaussian anamorphosis function, $*$ is the convolution operator.

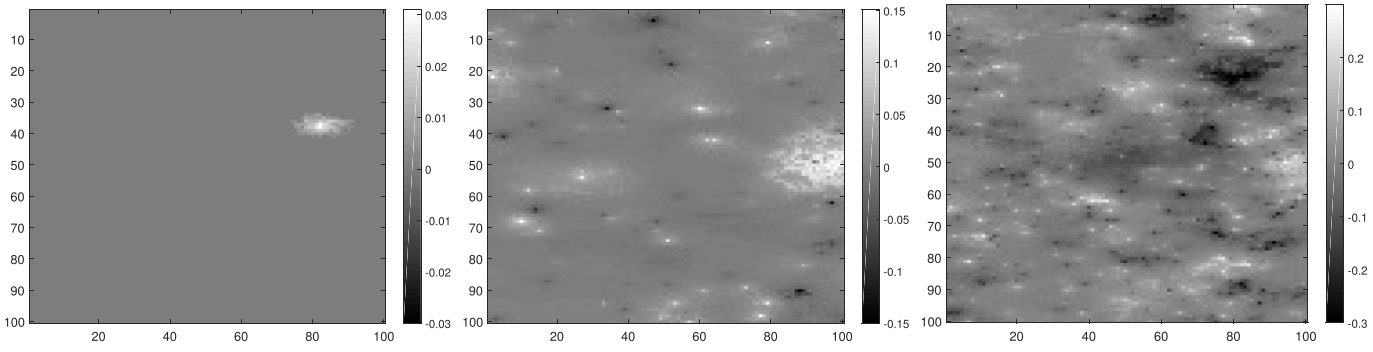


Fig. 2. Difference between a 100 × 100 Gaussian field and the same field with N random perturbations by FFTMA-SA. Variogram model: spherical, C = 1, a_x = 50 and a_y = 25.

unchanged areas already well calibrated. This capability sharply contrasts with PA where the entire field is modified at every iteration. Fig. 1 summarizes the FFTMA-SA algorithm.

After a few tests the schedule chosen for temperature and number of cells to perturb simultaneously are defined by a constant multiplicative reduction factor (in]0,1[) computed and applied at each iteration i:

$$T_i = T_0\left(\frac{T_{iter}}{T_0}\right)^{i/iter} \tag{3}$$

$$N_i = N_0\left(\frac{N_{iter}}{N_0}\right)^{i/iter} \tag{4}$$

where $iter$ is the chosen maximum number of iterations, T_0 and N_0 represent the initial temperature and the initial number of cell perturbed and T_{iter} and N_{iter} are the desired final temperature and final number of cells to perturb. In the case studies, the reduction factors were computed such that $T_0/T_{iter} = exp(15)$, N_0 is chosen between 5% and 20% of the total number of cells in the field and $N_{iter} = 1$. The ratio T_0/T_{iter} was chosen by trial-and-error and probably needs to be tailored to the particular OF studied. The schedule for number of points to perturb materializes the logical idea of working globally at the beginning of the algorithm and locally towards the end.

Fig. 2 illustrates the local behavior of FFTMA-SA. The figure shows the difference between a 100 × 100 Gaussian random fields with spherical variogram (C = 1, a_x = 50 and a_y = 25) and the same field for N random perturbations by FFTMA-SA for N = 1, 100 and 1000. When N = 1 (Fig. 2(a)), mostly cells within the correlation length are modified as a decreasing function of the distance to the perturbed point. With larger N, the entire field becomes affected by the perturbations. Selecting the points to perturb in a desired subarea as illustrated in Fig. 3 restore a more local behavior. Note that the perturbations extends beyond the target subarea. This ensures spatial continuity and preserves variogram reproduction when crossing subarea boundaries. Moreover, the weights defined by FFTMA extends farther than the (finite) correlation range as shown in Fig. 4 for the case of a spherical model.

2.3.1. Gaussian anamorphosis

The initial Gaussian anamorphosis of step 1 in Fig. 1 is required because FFTMA produces Gaussian fields. The anamorphosis is done graphically by associating Gaussian quantiles to the same experimental quantiles of HD. This only ensures marginal Gaussian distribution. Higher order characteristics of HD can be enforced by incorporating these in the OF definition.

2.3.2. Post-conditioning by kriging

Given an unconditional realization $Z_u(x)$ and HD $Z(x_i)$, $\forall\, i = 1..n$, one obtains a conditional realization by:

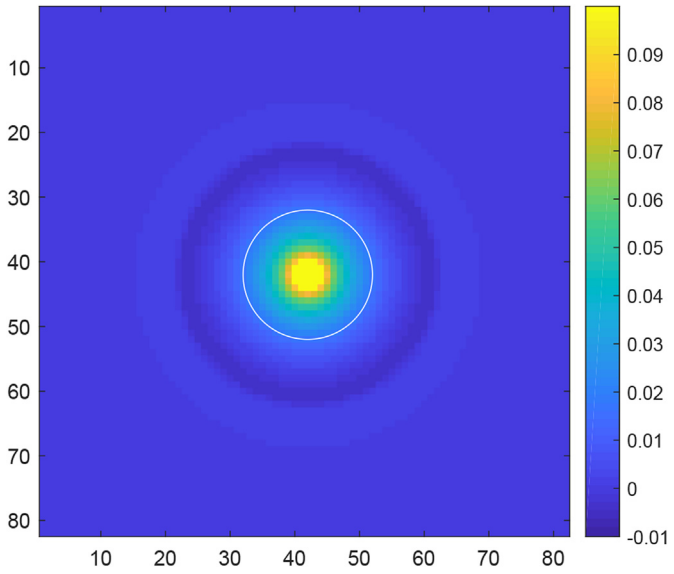


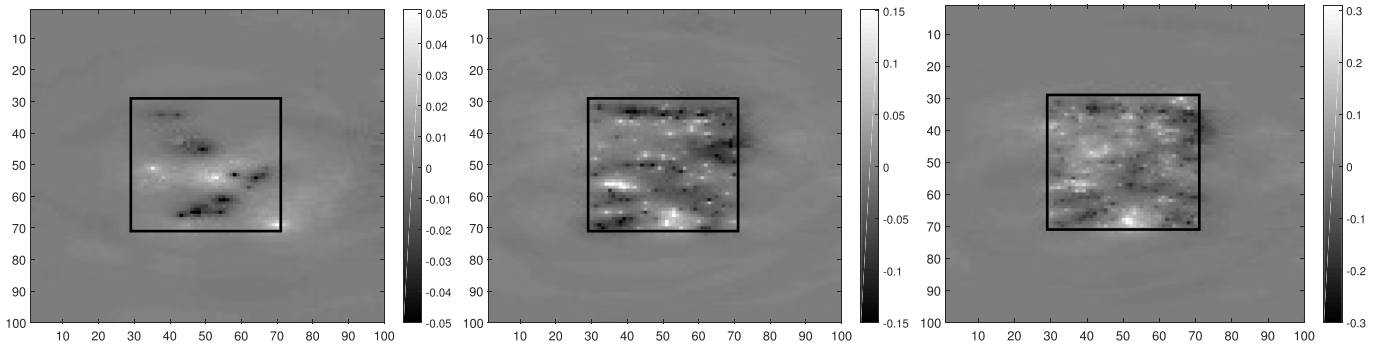
Fig. 4. MA weights obtained by FFTMA for a spherical model of range 20 (white circle). The weights produced by FFTMA extend beyond the correlation range.

$$Z_c(x) = Z^*(x) + (Z_u(x) - Z_u^*(x)) = Z_u(x) + (Z^*(x) - Z_u^*(x)) \tag{5}$$

where $Z_c(x)$ is the conditional realization at point x, $Z^*(x)$ is the kriging estimate at point x using HD at points x_i, $\forall\, i = 1..n$, $Z_u^*(x)$ is the kriging estimate obtained at x using unconditionally simulated values at points x_i. Hence, Eq. (5) clearly shows that post-conditioning by kriging can be seen as adding a spatially correlated error $(Z_u(x) - Z_u^*(x))$ to a kriged estimate, or alternatively to add the kriging of the difference $(Z^*(x) - Z_u^*(x))$ to the unconditional realization. Note that any form of kriging can be used for the post-conditioning and that post-conditioning is exact. Further details about this classical method are found in Chilès and Delfiner (2012).

3. Results

A synthetic case showing spatial asymmetry is constructed. FFTMA-SA is compared to classical simulated annealing (SA), uncalibrated conditional realizations and FFTMA-GD. The OF measures the dissimilarity between experimental third order spatial moments of the reference and the calibrated simulation. The second case study tests FFTMA-SA on the well known Walker Lake dataset (Isaaks and Srivastava, 1989; Mariethoz and Caers, 2014) with an OF giving the difference in facies proportions along vertical and horizontal directions. The third example illustrates the capability of FFTMA-SA to assimilate



(a) N=20 (b) N=200 (c) N=500

Fig. 3. Difference between a 100 × 100 Gaussian random field and the same field with N random perturbations by FFTMA-SA in a centered window 40 × 40. Variogram model: spherical, C = 1, a_x = 50 and a_y = 25.

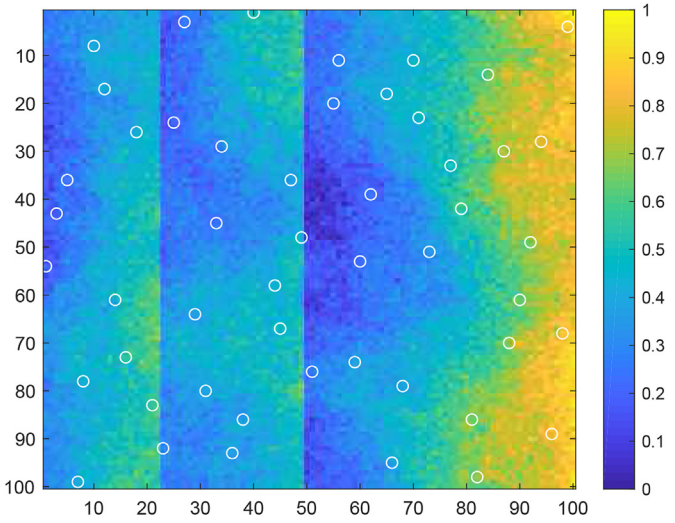


Fig. 5. Synthetic case. Reference field and location of conditioning data (white circles).

newly acquired data on a subarea of the field for the Walker Lake example.

3.1. Synthetic case

The first example (Fig. 5) is a simple synthetic field that could represent three successive volcanic sequences with silica concentration increasing within each sequence from West to East according to the Bowen's reaction series. At the beginning of a volcanic sequence, the silica concentration is low and mafic igneous rocks are crystallized. At the end, the silica concentration is higher and more felsic igneous rock appear. A new volcanic cycle reproduces typically the same sequence with possible variations in power and completeness of each cycle.

The dimension of the synthetic field is 100 × 100 cells. Fifty samples are randomly selected for conditioning data (white circles). The objective function is the misfit (quantified by RMSE) between the reference and simulated asymmetry function as defined by Bárdossy and Hörning (2017) and Hörning and Bárdossy (2018):

$$A(h) = \frac{1}{N(h)} \sum_{x_i - x_j \approx \mathbf{h}} (F(Z(x_i)) - F(Z(x_j)))^3 \tag{6}$$

where $N(h)$ is the number of pairs and $F(Z(x))$ is the *cdf* evaluated at $Z(x)$. It can be computed very efficiently on a regular grid using FFTs by generalizing the approach presented in Marcotte (1996). To avoid sampling fluctuations, and without loss of generality of the method, we used the entire reference field to compute the target asymmetry function. SA and FFTMA-SA algorithms are applied to get 100 realizations calibrated to the reference asymmetry functions. The stopping criteria is 5000 iterations. Note that the iteration number in our algorithm represents the number of evaluations. With FFTMA-SA, the number of points perturbed can be much higher as we simultaneously perturb many points of the field each time. Initially, 500 cells of the field are perturbed and this number is progressively reduced to one in the last iteration. A Tabu strategy is followed for node selection to prevent considering again a cell before all cells have been visited. Fig. 6(a) illustrates a realization obtained by SA and by FFTMA-SA. For reference, Fig. 6(b) shows a simple conditional realization without calibration.

Comparing SA and uncalibrated conditional realization shows, as expected, that SA better reproduces the alternating sequences. However, this is achieved at the expense of a larger noise variance (Fig. 7(a)), and globally a loss of spatial structure (Fig. 8(a)).

3.1.1. Results with FFTMA-SA

FFTMA-SA is applied on the same synthetic field. Fig. 8 present variograms along East and South directions and Fig. 9 displays the spatial asymmetry function. Fig. 9 shows that all FFTMA-SA realizations (dark gray) reproduce the reference asymmetry, contrary to uncalibrated realizations (light gray). The good calibration to asymmetry is done without any noticeable loss of the variogram reproduction obtained by FFTMA-SA (Fig. 8) despite that the variogram was not included in the objective function. Note that the calibrated fields are no more Gaussian as the spatial asymmetry is incompatible with a Gaussian field. Hence, the method provides a way to simulate non-Gaussian fields with imposed covariance function and higher order spatial moments different from the Gaussian case.

As a side effect, the calibration of asymmetry function by FFTMA-SA method enables globally an improved reproduction of the reference field as indicated by the cell-based histogram of mean errors between realizations-reference (Fig. 7(b)).

3.1.2. Comparison of FFTMA-GD with FFTMA-SA

Gradual deformation (GD) is applied on the same synthetic case. The OF obtained by FFTMA-SA and FFTMA-GD with 1 000, 5000 and 10 000 iterations are shown as a function of the iteration index. Fig. 10 presents the mean common logarithm of the of 100 realizations of FFTMA-SA (red) and FFTMA-GD (blue). Each iteration corresponds to an evaluation of the OF. Hence, the computation times are similar for the two methods as the CPU most intensive part is computation of the OF. FFTMA-GD decreases initially faster than FFTMA-SA, but FFTMA-SA decreases at a globally higher rate. As a result, in the last iteration, FFTMA-SA presents in average a two to three-fold reduction of the objective function with respect to FFTMA-GD.

The correlations of the four methods (SA, uncalibrated realization, FFTMA-GD and FFTMA-SA) with the reference have been computed for each of 100 realizations. The mean correlation and standard deviation of the correlations are shown in Table 1. SA realizations show poor correlation with the reference due to the loss of spatial structure (introducing the variogram in the OF helped increase the correlation but it remained smaller than for FFTMA-GD and FFTMA-SA). The mean correlations of FFTMA-SA and FFTMA-GD are similar and significantly larger than those obtained with SA or with uncalibrated conditional realizations. This indicates that the asymmetry function is an important feature of the synthetic case that should be duly considered.

3.2. Walker Lake data

The second example bears on the well known Walker Lake data set (Isaaks and Srivastava, 1989). The complete field is 260 × 300 cells (Fig. 11(a)). A regular grid of 100 data (10 × 10) form the conditioning HD. We define an indicator variable $T(x, y)$ taking the value one when variable $V(x, y) > 500$ (see Fig. 11(b)). The mean of $T(x, y)$ is computed along X- and Y-axis and then the moving averages over 50 consecutive lines or columns are computed to obtain the proportion curves $\overline{T_r}(x)$ and $\overline{T_r}(y)$. The OF to minimize is the RMSE between the reference and the realization facies proportions.

Fig. 11 presents realizations of variable V (uncalibrated: Fig. 11(c) and calibrated: Fig. 11(e)) and the e-type of variable T computed after truncation of the 100 realizations of variable V for uncalibrated (Fig. 11(d)) and calibrated realizations (Fig. 11(f)). As expected, calibrated e-type estimates of T better represent areas of high values than uncalibrated ones (compare with Fig. 11(b)). The mean correlation on V between realizations and reference is higher for calibrated realizations than uncalibrated ones for both FFTMA-GD and FFTMA-SA with a slight advantage for FFTMA-SA (see Table 2).

Fig. 12 presents the moving average along the X-axis (left) and the Y-axis (right) for the uncalibrated and FFTMA-SA realizations. After 2000 iterations, the calibrated realizations by FFTMA-SA (dark gray) reproduce almost perfectly the reference curves (solid red). In fact, 500

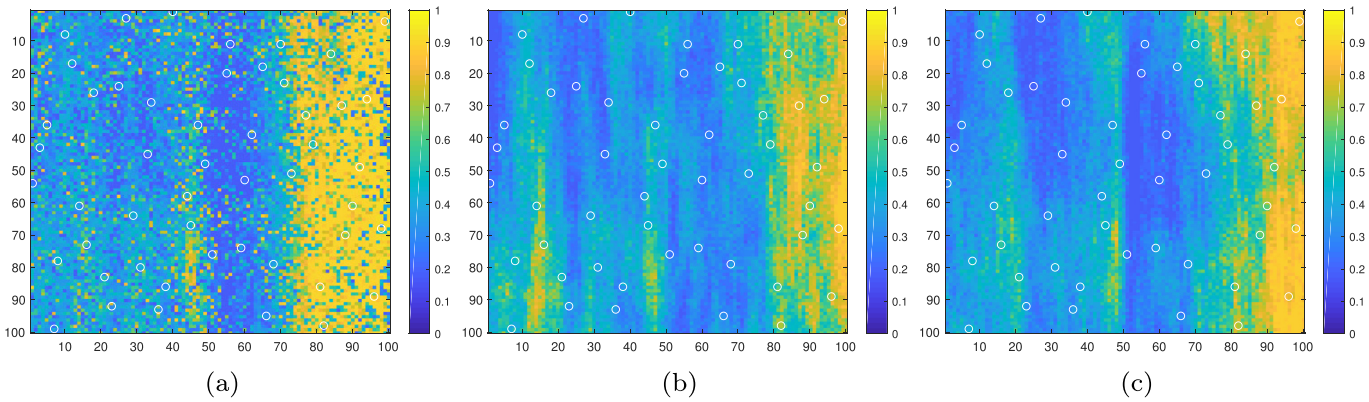


Fig. 6. Example of simulation by SA (a), uncalibrated conditional realization (b), and FFTMA-SA (c).

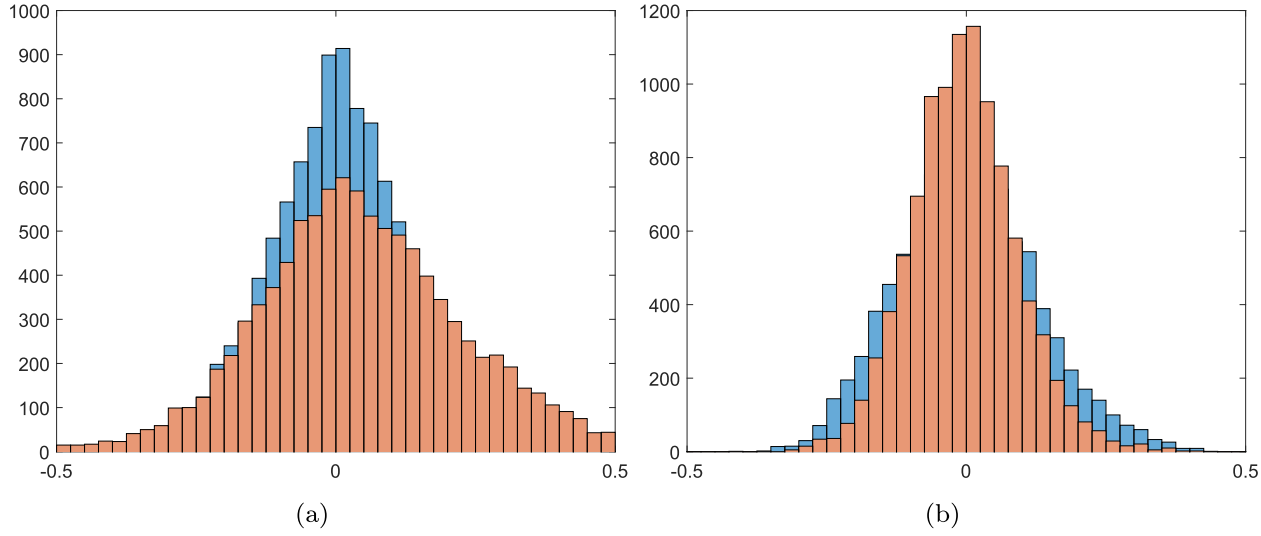


Fig. 7. Cell-based histogram of the mean difference reference-realizations over 100 uncalibrated conditional realizations (blue) and (a) SA (red) and (b) FFTMA-SA (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

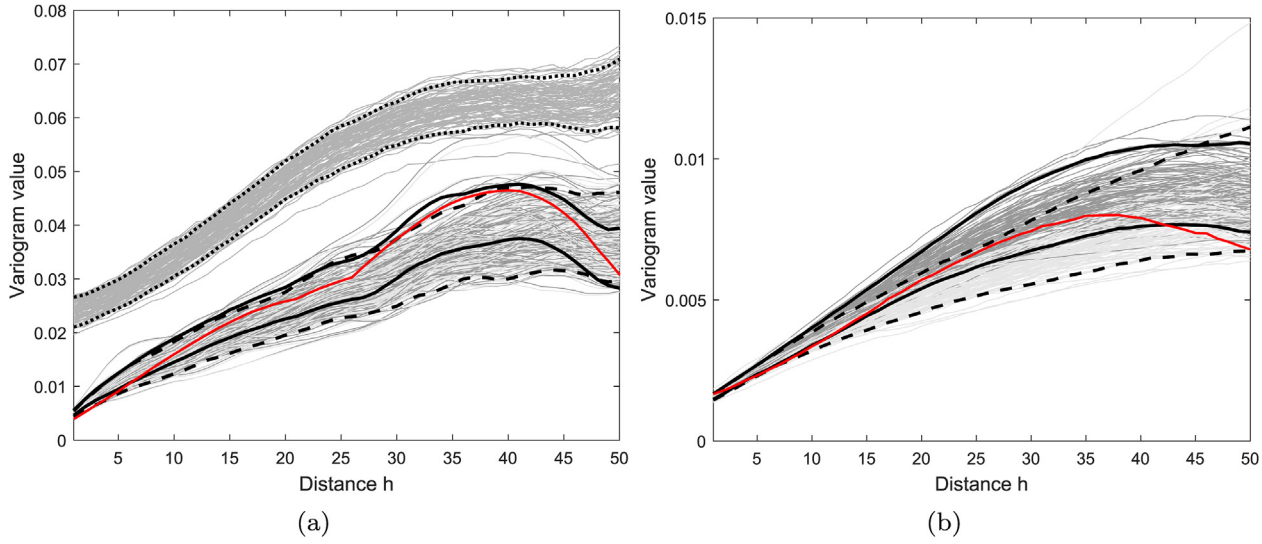


Fig. 8. Experimental variograms and 90% confidence intervals along (a) East direction for SA (dottted lines), uncalibrated conditional realizations (dashed lines) and calibrated by FFTMA-SA (solid black lines), (b) South direction for uncalibrated conditional realizations (dashed lines) and calibrated by FFTMA-SA (solid black lines). Reference field variograms (solid red lines). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

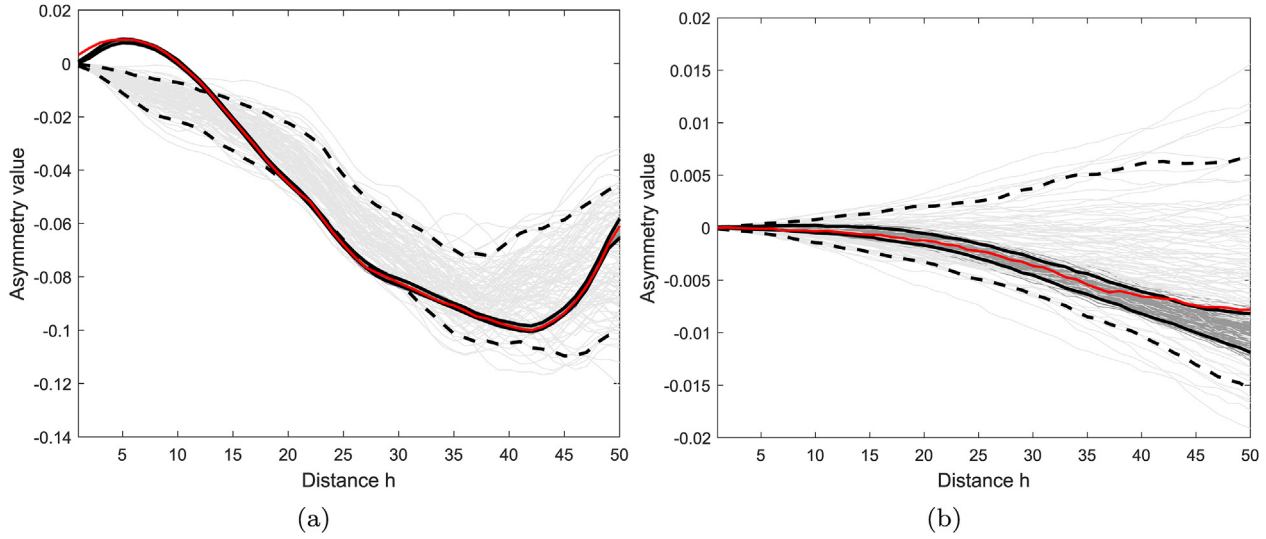


Fig. 9. (a) Asymmetry function along East direction, (b) asymmetry function along South direction. Reference field (solid red), FFTMA-SA realizations (dark gray), uncalibrated conditional realizations (light gray); 90% confidence intervals shown (solid dark: calibrated and dashed lines: uncalibrated). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

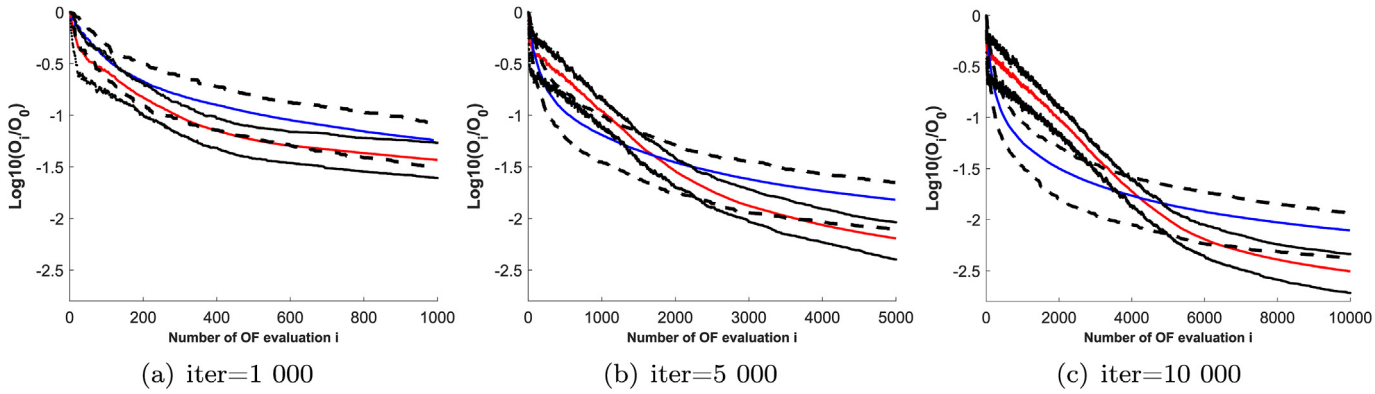


Fig. 10. Mean of objective functions (red: FFTMA-SA, blue: FFTMA-GD) and 75% confidence intervals shown (dark lines: FFTMA-SA, dashed lines: FFTMA-GD); 100 realizations of both methods. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 1
Correlations of simulated fields with synthetic reference field; 100 realizations.

Method	1000 iterations		5000 iterations		10 000 iterations	
	Mean ρ	Std ρ	Mean ρ	Std ρ	Mean ρ	Std ρ
SA	0.64	0.03	0.57	0.05	0.44	0.08
Uncalibrated	0.78	0.03	0.78	0.02	0.78	0.03
FFTMA-GD	0.83	0.02	0.84	0.02	0.85	0.02
FFTMA-SA	0.83	0.02	0.85	0.02	0.85	0.03

iterations were enough to obtain an almost equivalent calibration with this OF. By comparison, uncalibrated conditional realizations (light gray) show much larger deviations with respect to target functions.

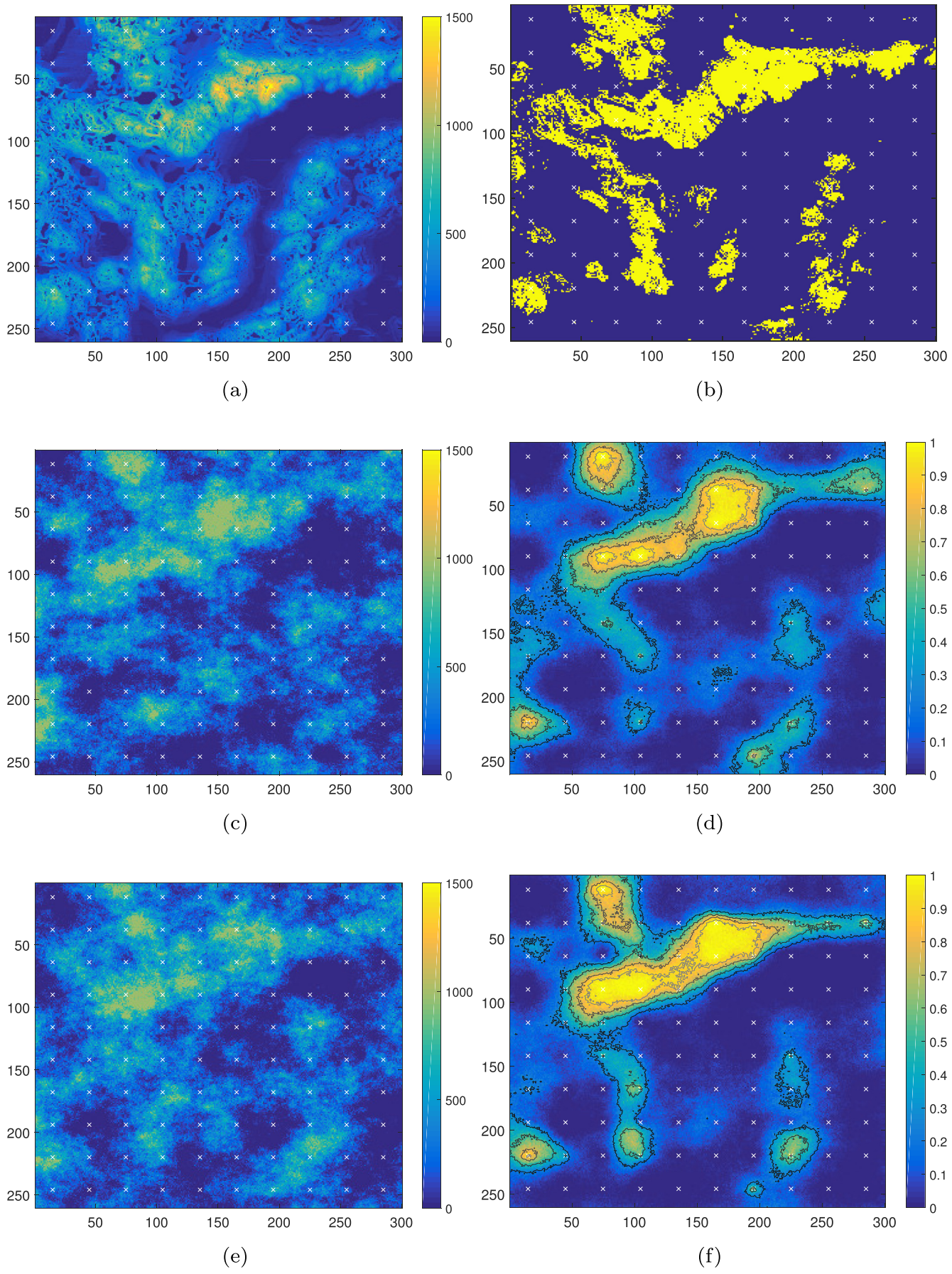
Fig. 13 compares the evolution of the OF for FFTMA-GD and FFTMA-SA. As for the previous synthetic case study, FFTMA-GD stalls early after a quick decrease of the OF compared to FFTMA-SA which decreases initially at a slower pace but is able to get better calibration at the last iteration. FFTMA-SA shows a similar behavior on all subfigures, with no obvious sign of getting trapped in a local minimum. The relative stabilization observed at iteration approximately equal to *iter*/2 corresponds to $N_i \approx 20$, a relatively small number of simultaneously perturbed cells in a field comprising 78 000 cells.

The evolution of the calibration with choice of *iter* parameter differs significantly for FFTMA-SA and FFTMA-GD. Fig. 14 (a) to (c) shows the calibrated function along the y-axis for *iter*= 100, 200 and 500 respectively. The sets of FFTMA-SA realizations define intervals around the target proportion whose width decreases with increase of *iter*. The target curve is well contained within the confidence interval even with *iter* = 100. The result with *iter* = 500 is close to the one obtained with *iter* = 2000 (Fig. 12). The lack of fit is rather homogeneously distributed over the whole target curve, so the maximum number of iterations (100 vs 200 vs 500 or more) applied in FFTMA-SA globally controls fairly well the variability of FFTMA-SA realizations around the target response.

The behavior of FFTMA-GD in this example differs. Fig. 14 (d) to (f) shows that important portions of the target function are well outside the confidence interval defined by realizations, even with 500 iterations. The spread of responses is much smaller than for FFTMA-SA in the 100 and 200 iteration cases. In fact, the confidence intervals of FFTMA-GD show almost the same width irrespective of the number of iterations applied. Here, contrary to the FFTMA-SA case, the number of iterations does not control well the variability of realizations and globally the confidence intervals are not credible.

3.3. Local FFTMA-SA

The third example illustrates the capability of FFTMA-SA to work locally and to perform data assimilation. New local data informing on



(caption on next page)

Fig. 11. Walker-Lake data: (a) Variable V and location of conditioning data (white crosses), (b) variable T (Indicator of V > 500, in yellow). (c) uncalibrated conditional realization, (d) e-type of T for 100 uncalibrated conditional realizations, (e) calibrated conditional realization by FFTMA-SA, (f) e-type of T for 100 FFTMA-SA realizations, contours at 0.25, 0.5, 0.75 and 0.9. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 2
Walker lake Data: Correlations of simulated fields with reference field for *V* variable; 100 realizations.

Method	100 iterations		200 iterations		500 iterations		1000 iterations	
	Mean ρ	Std ρ	Mean ρ	Std ρ	Mean ρ	Std ρ	Mean ρ	Std ρ
Uncalibrated	0.50	0.03	0.49	0.03	0.50	0.03	0.49	0.03
FFTMA-GD	0.51	0.01	0.50	0.004	0.53	0.004	0.52	0.002
FFTMA-SA	0.53	0.03	0.55	0.02	0.55	0.02	0.55	0.02

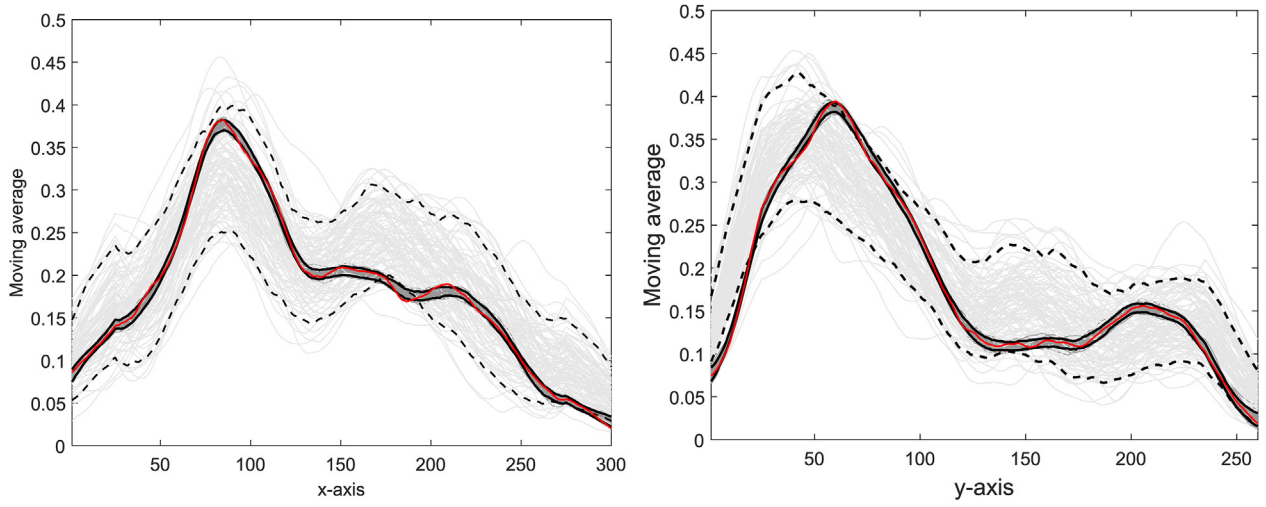


Fig. 12. Walker Lake data. Moving average of variable T over 50 lines along x-axis (a) and y-axis (b). Reference field (solid red), FFTMA-SA realizations after 2000 iterations (dark gray), uncalibrated conditional realizations (light gray). 90% confidence intervals shown (dashed lines: uncalibrated, solid black: calibrated). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

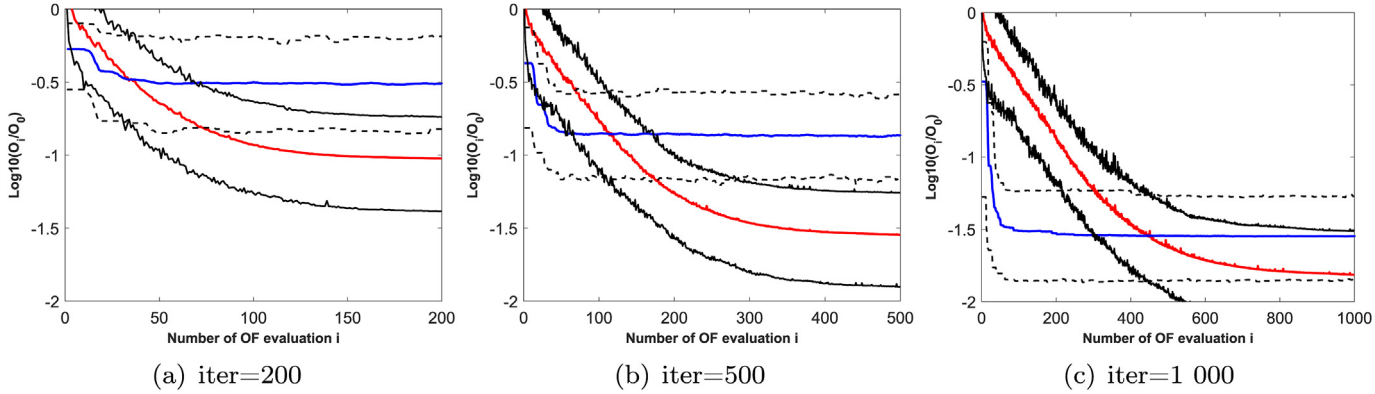


Fig. 13. Walker Lake data. Mean of objective functions (red: FFTMA-SA, blue: FFTMA-GD) and 90% confidence intervals shown (solid lines: FFTMA-SA, dashed lines: FFTMA-GD), 100 realizations. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

local averages of variable V (over 21 × 21 blocks) become available (see Fig. 16). One hundred realizations calibrated globally as in section 3.2 are obtained first. Then, we update the realizations so as to assimilate the newly acquired local data. For this, we form a new OF composed of two terms, the original global OF and the misfit to the new local data. The realizations are updated to local data by a new round of iterations (with cooling and number of points schedules reinitialized) with the global-local OF. The global part of the OF keeps the updated realizations compatible with the entire field whereas the local part improves fit to the new local data. In this second phase, the points are selected within the local window (although we could have use a mix of local and global perturbations). Fig. 15 shows that the assimilation phase continues improving the global calibration and succeeds to fit the local data as well. Fig. 16(d) and (e) show the e-type of variable $T(x, y)$ for the global OF and the global-local OF respectively. Fig. 16(f) shows the absolute difference between the two figures. The local data

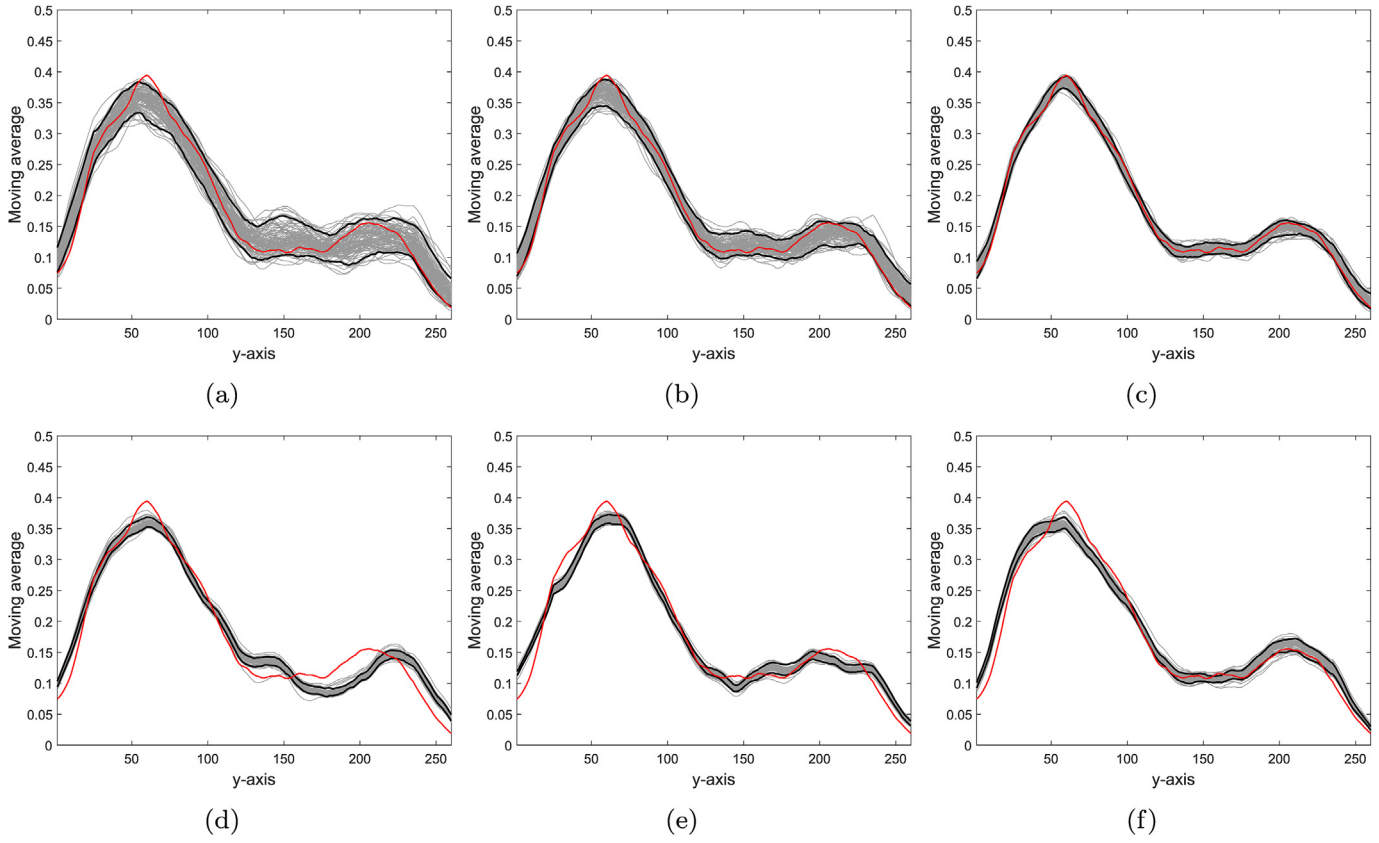


Fig. 14. Moving average of T over 50 lines along y-axis for FFTMA-SA (top row) and FFTMA-GD (bottom row). From left to right, calibration done using 100, 200 and 500 iterations. Reference field (solid red) and 90% confidence intervals shown (solid dark). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

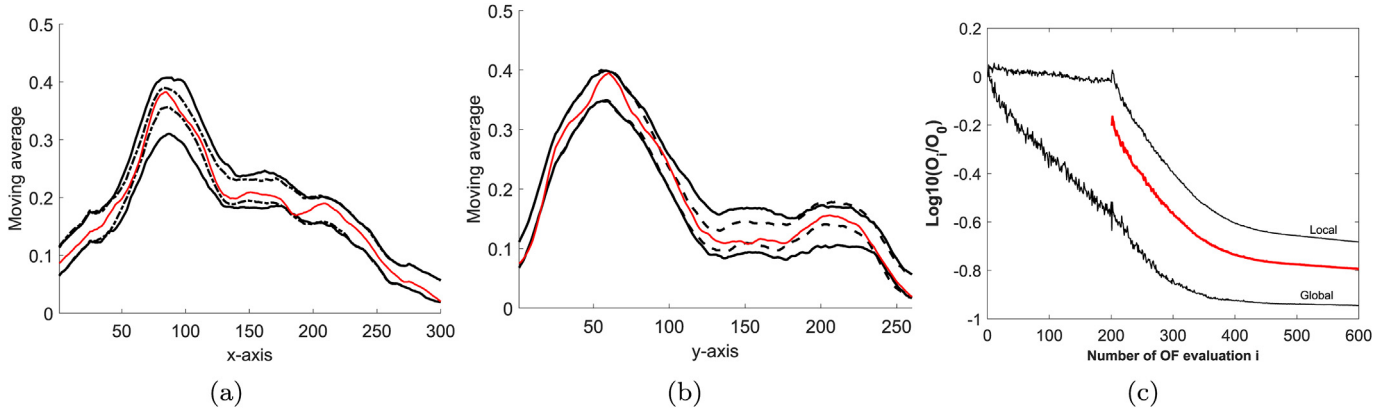


Fig. 15. Walker Lake data. Moving average of variable T over 50 lines along x-axis (a) and y-axis (b). Reference field (solid red). 90% confidence intervals obtained by global-local calibration (dashed black), by global calibration (solid black). (c) Evolution of the OF. Mean (red line), Mean of global OF and local OF (black lines); 100 realizations. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

assimilation increased the correlation with the reference field (see Table 3).

4. Discussion

A new algorithm for calibration combining FFTMA and SA is introduced. FFTMA-SA works by perturbing the independent random numbers of the FFTMA generator using a simulated annealing approach modified to allow simultaneous perturbation of many cells combined with a Tabu selection of cells to avoid revisiting the same cells too soon.

Compared to the classical SA, FFTMA-SA preserves by construction the simulated covariance after each iteration without having to introduce the variogram function in the OF, a major advantage as this typically produce artefacts in SA (Groleau and Marcotte, 1997). Similarly, due to the post-conditioning step, there is no need to include conditioning points in the OF as done with PA (Hörning and Bárdossy, 2018; Yao, 1998). Contrary to PA, FFTMA-SA exactly honours the HD. Moreover, PA works always globally whereas FFTMA-SA is by construction global-local thanks to the schedule for the number of points to perturb. We stress that the global-local behavior is built-in our approach contrary to FFTMA-GD where local perturbations rely on prior identification of the most sensitive areas, a task that can be difficult for some OF. With both FFTMA-SA and FFTMA-GD, local perturbations do not disturb the field covariance and continuity.

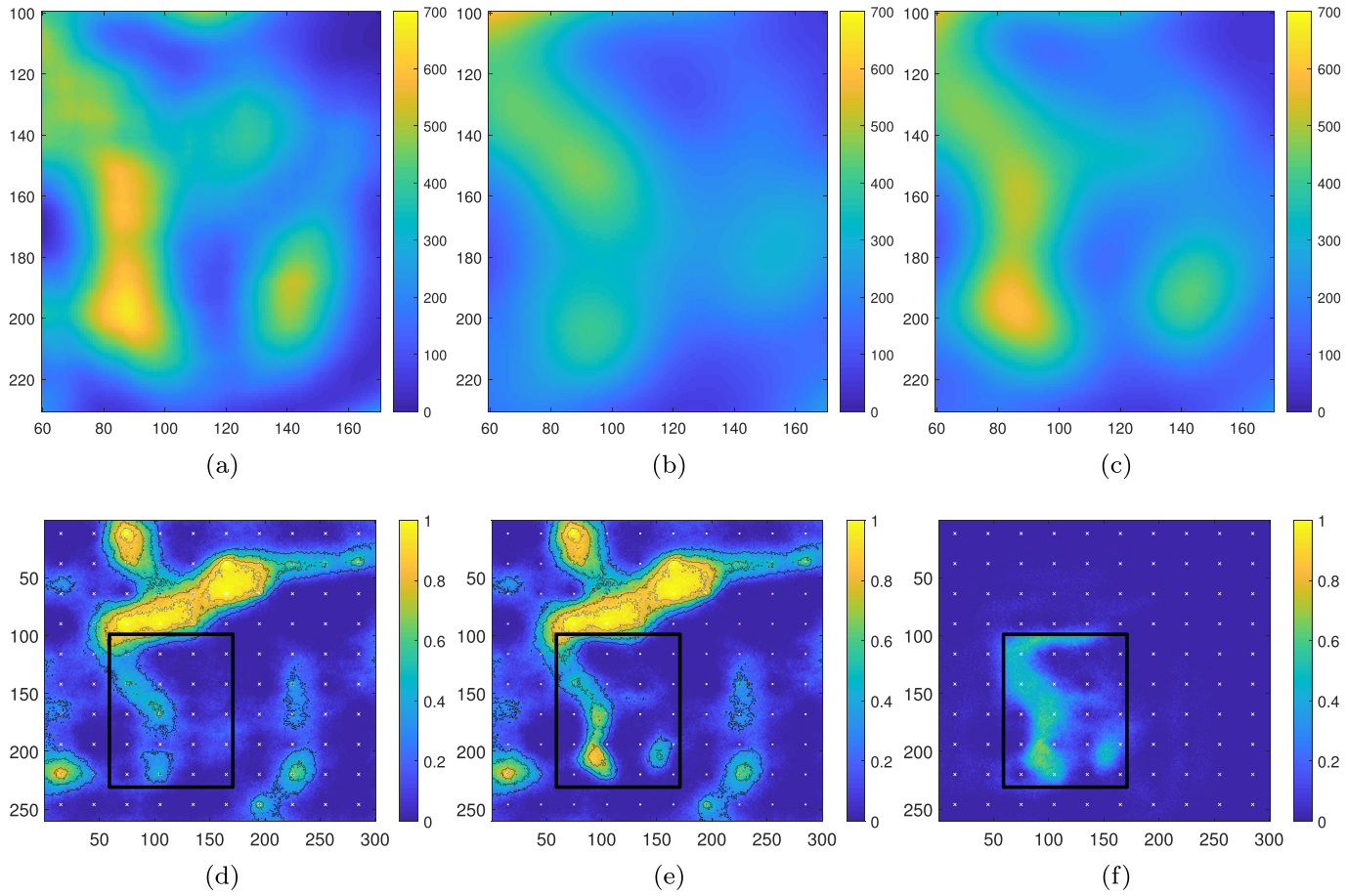


Fig. 16. Top row: Calibration of local averages (block of 21 × 21) of variable V (a) reference, (b) global calibration, (c) global-local calibration. Bottom row: conditioning data (white crosses), (d) calibrated conditional realization by FFTMA-SA and global OF, (e) calibrated realization by FFTMA-SA and global-local OF, (f) absolute difference between (d) and (e). e-type contours at 0.25, 0.5, 0.75 and 0.9. Local calibration window outlined.

Table 3
Walker lake Data: Correlations of simulated fields with reference field for *V* variable, 100 realizations.

Method	Mean ρ	Std of ρ
Uncalibrated	0.49	0.03
FFTMA-SA (Global OF)	0.53	0.02
FFTMA-SA (Global and local OF)	0.58	0.02

The two examples presented, synthetic and Walker Lake, illustrate the efficiency of the FFTMA-SA method to match each individual realization to the target response almost exactly when the number of iterations is selected large enough. This is a desirable feature when the response function is known with a high degree of precision. Typical examples of precise response functions are, in hydrogeology, the travel time between wells in a tracer test, or in petroleum reservoir engineering, the pressure drop or water cut in wells. However, some response functions can present significant uncertainties (e.g. facies proportion curves). To avoid overfitting, it is then advisable to stop iterations early in FFTMA-SA. One example showed (Fig. 14 (a) to (c)) that such strategy with FFTMA-SA increased variability of realizations and maintained credible confidence intervals around the target function. The width of the interval appeared well controlled by the number of iterations specified. A similar behavior (not shown) was observed for the synthetic case. Whether this is a general behavior of FFTMA-SA would require further study but we note that applying the same strategy with FFTMA-GD failed to produce credible confidence intervals for the Walker Lake case study. FFTMA-GD produced calibrated realizations lacking variability and unable to match some parts of the target proportion curves.

The main advantages of FFTMA-SA over FFTMA-GD are that FFTMA-SA is able, in the examples tested, to reach significantly lower OF for the same number of iterations. The variability of realizations appears better controlled by number of iterations with FFTMA-SA than with FFTMA-GD. The built-in global-local behavior of FFTMA-SA stems from the schedule on number of points to perturb. It confers a significant advantage to FFTMA-SA compared to FFTMA-GD. Finally, we note that convergence properties of SA are also theoretically better established than those of GD.

Two important drawbacks of FFTMA-SA (also shared by FFTMA-GD and PA) are worth mention. First, it works on a regular grid and second, it can be limited by computer memory for covariances with a large range. When HD are not on a regular grid, it is possible to use FFTMA on a sufficiently dense regular grid and then use sequential Gaussian simulations (SGS) or local Cholesky simulations to extend simulated values from the grid to the HD points to get the simulated values required for post-conditioning by kriging. This approach is adequate provided the screen effect of simulated points within the local neighbourhood around HD is strong enough. It would not involve significant increase in computing time and would leave unchanged all the other steps of the FFTMA-SA algorithm. Another simpler option classically adopted is to migrate each HD point to its closest cell, provided the displacement is small enough. The problem of memory limitation is more serious, one possible avenue is to replace FFTMA by a spatial or spectral turning band simulator which has no memory limitation. However, the price to pay would be the lost of capability to work

locally.

5. Conclusion

The proposed approach, FFTMA-SA, provides a fast and general method to produce realizations that can be calibrated to various response functions. The calibration is done efficiently without disturbing the desired field covariance function while honouring all HD by construction. Applications of the method on a synthetic example and the Walker Lake data demonstrated its capability to meet the imposed objectives and showed increased correlations with the reference field after calibration to the target responses. Comparisons with SA and FFTMA-GD showed better performance of FFTMA-SA. The ensemble of FFTMA-SA calibrated realizations defined credible confidence intervals around the reference, contrary to SA and FFTMA-GD. The capability of FFTMA-SA to work globally-locally was demonstrated with a data assimilation example. Good global and local calibrations were obtained in a relatively small number of iterations, illustrating FFTMA-SA constitutes an interesting and flexible alternative for model calibration.

Authorship statement

The main author Dany Lauzon did most of the programming work, all the computations and the initial manuscript writing. The second author Denis Marcotte provided the research idea, supervised the research, and did some of the programming and the final writing of the manuscript.

Computer code availability

The Matlab computer codes are available at https://github.com/Denis-Marcotte/FFTMA-SA.

Acknowledgments

This research was made possible by National Research Council of Canada thru NSERC grant (RGPIN-2015-06653) of D. Marcotte and NSERC undergraduate research grant (BRPC - 526627 - 2018) of D. Lauzon. The “Fonds de recherche nature et technologies du Québec” also contributed to financing the project. Constructive comments from three anonymous reviewers were helpful improving the paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cageo.2019.03.003.

References

Bárdossy, A., Hörning, S., 2017. Process-driven direction-dependent asymmetry: identification and quantification of directional dependence in spatial fields. Math. Geosci. 49, 871–891. https://doi.org/10.1007/s11004-017-9682-1.

Chilès, J., Delfiner, P., 2012. Geostatistics: Modeling Spatial Uncertainty, second ed. John Wiley & Sons.

de Marsily, G., Lavedan, G., Boucher, M., Fasanino, 1984. Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model. In: In: Verly, G., David, M., Journel, A., Marechal, A. (Eds.), Geostatistics for Natural Resources Characterization. NATO ASI Ser. C, vol. 182. Reidel Publishing Co, pp. 831–849.

Deutsch, C., 1992. Annealing Techniques Applied to Reservoir Modeling and the Integration of Geological and Engineering (Well Test) Data. Ph.D. thesis. Stanford University.

Deutsch, C.V., Journel, A.G., 1994. The application of simulated annealing to stochastic reservoir modeling. SPE Adv. Technol. 2, 222–227. https://doi.org/10.2118/23565-pa.

Evensen, G., 2009. Data Assimilation. Springer Berlin Heidelberghttps://doi.org/10.1007/978-3-642-03711-5.

Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. In: IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-6pp. 721–741. https://doi.org/10.1109/tpami.1984.4767596.

Gómez-Hernánez, J.J., Sahuquillo, A., Capilla, J., 1997. Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data—i. theory. J. Hydrol. 203, 162–174. https://doi.org/10.1016/s0022-1694(97)00098-x.

Groleau, P., Marcotte, D., 1997. The border effect of simulated annealing. Math. Geol. 29, 585–592. https://doi.org/10.1007/bf02775088.

Hörning, S., Bárdossy, A., 2018. Phase annealing for the conditional simulation of spatial random fields. Comput. Geosci. 112, 101–111.

Hu, 2000. Gradual deformation and iterative calibration of Gaussian-related stochastic models. Math. Geol. 32, 87–108.

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, 703–719. https://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.

Isaaks, E., Srivastava, R., 1989. An Introduction to Applied Geostatistics. Oxford University Press.

Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. Science 220, 671–680.

Lantuéjoul, C., 2002. Geostatistical Simulation. Springer, Berlin, Heidelberg.

Le Ravalec, M., Noetinger, B., Hu, L., 2000. The FFT moving average generator : an efficient numerical method for generating and conditioning Gaussian simulations. Math. Geol. 32, 701–722.

Marcotte, D., 1996. Fast variogram computation with FFT. Comput. Geosci. 22, 1175–1186. https://doi.org/10.1016/S0098-3004(96)00026-X.

Mariethoz, G., Caers, J., 2014. Multiple-point Geostatistics. Wiley-Blackwell. https://www.ebook.de/de/product/22554917/gregoire_mariethoz_multiple_point_geostatistics.html.

Oliver, D.S., 1995. Moving averages for Gaussian simulation in two and three dimensions. Math. Geol. 27, 939–960. https://doi.org/10.1007/bf02091660.

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, 61–91. https://doi.org/10.1007/bf02769620.

Oliver, D.S., Chen, Y., Nævdal, G., 2010. Updating Markov chain models using the ensemble Kalman filter. Comput. Geosci. 15, 325–344. https://doi.org/10.1007/s10596-010-9220-4.

Ravalec-Dupin, M.L., 2005. Inverse Stochastic Modeling of Flow in Porous Media (IFP Publications). Editions Technips.

Renard, P., Allard, D., 2013. Connectivity metrics for subsurface flow and transport. Adv. Water Resour. 168–196.

Rezaee, H., Marcotte, D., 2018. Calibration of categorical simulations by evolutionary gradual deformation method. Comput. Geosci. 22, 587–605. https://doi.org/10.1007/s10596-017-9711-7.

Sneddon, I., 1951. Fourier Transforms. McGraw-Hill.

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, 2069–2082. https://doi.org/10.1029/wr022i013p02069.

Yao, T., 1998. Conditional spectral simulation with phase identification. Math. Geol. 30, 285–308.