Computers & Geosciences 187 (2024) 105581

Contents lists available at ScienceDirect

Computers and Geosciences

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

Review Article

Conditioning of multiple-point statistics simulations to indirect geophysical data

Shiran Levy [a,*], Lea Friedli [a], Grégoire Mariéthoz [b], Niklas Linde [a]

[a] Institute of Earth Sciences, University of Lausanne, Lausanne, Switzerland
[b] Institute of Earth Surface Dynamics, University of Lausanne, Lausanne, Switzerland

ARTICLE INFO

Dataset link: https://github.com/ShiLevy/IDCS

Keywords:
Geophysics
Bayesian inversion
Geostatistics
Multiple-point statistics
Ground-penetrating radar

ABSTRACT

Multiple-point statistical (MPS) simulation methods have gained widespread adoption across various Earth science disciplines. They offer a versatile framework for simulating intricate spatial patterns and heterogeneity in both surface and subsurface structures. While these simulations adeptly incorporate conditioning to hard data, such as information from boreholes, conditioning to indirect data (e.g. geophysical data) is more challenging. A new methodology is introduced that provides geostatistical realisations honouring indirect geophysical data and complex prior knowledge described by a training image. An MPS simulation is iteratively built up pixel-by-pixel starting from an empty grid or with initial hard conditioning data if available. During each simulation step, a pixel value is selected from a set of candidates proposed by the MPS algorithm. This selection is made proportionally to an approximated likelihood that accounts for indirect geophysical data. The expected values and uncertainty quantification are obtained by simulating many complete field realisations. Our approach, which we name Indirect Data Conditional Simulations (IDCS), is tested for multi-Gaussian and complex subsurface structures with synthetic data from linear and non-linear crosshole ground-penetrating radar responses. The IDCS method is inherently approximate due to the finiteness of the training image, a limited number of MPS candidates at each simulation step and the need to approximate intractable likelihood functions. Nevertheless, the results demonstrate that the posterior approximations obtained by IDCS are often comparable to those obtained with a Markov chain Monte Carlo method, with IDCS being at least one order of magnitude faster. While the method performs the best when the underlying physics is modelled as a linear response, encouraging preliminary results considering non-linear physical responses are provided.

1. Introduction

Multiple-point statistics (MPS) is a non-parametric family of methods used to produce geostatistical realisations honouring higher-order spatial statistics present in so-called training images (TI; Guardiano and Srivastava, 1993; Strebelle, 2002; Zhang, 2006; Mariethoz and Caers, 2014). These methods proceed in a sequential manner by assigning parameter values to points on a simulation grid. This process entails scanning the training image and contrasting the patterns within it with the patterns surrounding the simulated point on the simulation grid, using various distance metrics. They are widely used for applications in geology and hydrogeology (Høyer et al., 2017; Le Coz et al., 2017; Barfod et al., 2018), remote sensing (Zakeri and Mariethoz, 2021 and references therein) and reservoir engineering (Zhang et al., 2006; Melnikova et al., 2015) to obtain ensembles of model realisations with the spatial statistics of the TI while also honouring available hard data (e.g. borehole information) or volume (linear averages, Straubhaar et al., 2016) measurements. Even if deep generative models offer highly competitive approaches to generate unconditional realisations (Laloy et al., 2017, 2018), MPS algorithms are still far superior in conditioning to hard data (Zhang, 2015; Hansen et al., 2018; Straubhaar and Renard, 2021).

Sequential geostatistical simulations, including MPS, are commonly constrained to hard data such as well measurements. Multiple-point statistics methods can additionally be utilised as a post-processing tool to refine a deterministic, smoothness-constrained solution obtained from a non-linear inversion. In such a setting, the resolution of the model realisations is enhanced (Lochbühler et al., 2014; Linde et al., 2015), but without any guarantees that the resulting realisations honour the original geophysical data. Other algorithms such as the Blocking Moving Window (BMW) method, introduced by Alcolea and Renard (2010), constrains MPS simulations to both hard data and connectivity information and while it introduces correlation with soft data

* Corresponding author.
E-mail address: shiran.levy@unil.ch (S. Levy).

https://doi.org/10.1016/j.cageo.2024.105581
Received 1 November 2023; Received in revised form 24 February 2024; Accepted 23 March 2024
Available online 26 March 2024

(e.g. geophysical models) through a secondary training image, it does not impose it as a constraint.

Geophysical inversion using MPS algorithms generally relies on Markov chain Monte Carlo (MCMC) methods (Mariethoz et al., 2010a; Hansen et al., 2012). At each model proposal step, an MPS algorithm performs sequential Gibbs sampling in which a subset of randomly chosen model cells (Mariethoz et al., 2010a) or a randomly selected patch (Hansen et al., 2012) within the model domain are re-simulated. The model proposals generated by the MPS algorithm are consistent with the patterns of the training image and conditional on the cell values that have not been re-simulated. The acceptance probability is given by the ratio of the likelihoods of the proposed model state and the previous model state in the chain. This extended Metropolis method (Mosegaard and Tarantola, 1995) will eventually locate the posterior probability density function (PDF) and sample proportionally to it. However, it is often very slow in practice, as geostatistical re-simulation and forward simulation times are non-negligible, and there is often a need to perform millions of MCMC iterations before the posterior PDF is sufficiently sampled. The latter is a result of very high correlation in the sampled MCMC states, implying that very long runs are needed to draw a sufficient number of independent samples (Ruggeri et al., 2015).

In the context of linear forward problems, Hansen et al. (2006) introduced a method for conditioning sequential simulations to noisy indirect data of mixed support (point- and volume-support). This method allows for the incorporation of geophysical measurements into the simulation process. In their implementation, the mean and covariance of the posterior PDF is obtained by solving a kriging system with an a priori mean and covariance as well as support volumes related to the physical response. In a subsequent step, posterior realisations are generated through sequential simulation using the kriging mean and covariance. This method was later extended by Hansen and Mosegaard (2008) to accommodate non-Gaussian marginal prior distributions, however, it captures only two-point spatial statistics, and its ability to accurately reproduce the prior is restricted by the kriging process. Applying the concept of averages over support volumes to MPS, Straubhaar et al. (2016) showed how simulations can be constrained to indirect geophysical data and the multiple-point statistics of a conditional prior. In their method, MPS candidate values for a simulated location are accepted according to an accumulated error considering the target value (mean value obtained from the data), a tolerance range and the mean over the support volume in the simulation grid. This method, however, does not sample the posterior PDF, as it does not account for the error statistics and is based on subjective tolerance values.

In this paper, a methodology is proposed enabling geostatistical simulations honouring geophysical constraints under a linear system response and its extension to non-linear responses is explored. This approach can, for instance, be applied to potential-field methods such as gravity, magnetics and self-potential when prior knowledge is best represented by higher-order statistics (e.g. Jensen et al., 2022). Our approach involves gradually constructing an MPS realisation starting from an empty simulation grid or, if hard data is available, with the known local data values. Each simulated value is selected based on geostatistical constraints considering spatial patterns formed by the already informed values, as well as constraints offered by the geophysical data. Incorporating the latter constraint would normally involve calculating a likelihood by marginalising over the distributions of the uninformed model parameters (grid cells), something that is computationally impractical. Instead, at each simulation step, k conditional samples are drawn using the MPS algorithm and one is accepted with a probability that is proportional to an approximate likelihood. In the likelihood approximation, the uninformed model parameter values (mean and covariance) are estimated using kriging. Once the simulation grid is fully informed, it can be seen as a draw from an approximate posterior distribution. Conducting multiple independent simulations results in an ensemble approximating the posterior distribution.

The proposed approach is faster than sequential Gibbs sampling within MCMC, as simulations are built up conditionally to the data, at each simulation step, and no re-simulation steps are performed. Moreover, the approach can be easily parallelised since each full simulation can be performed independently of other simulations. Nonetheless, the method is approximate due to three factors: (1) the training image is finite, (2) the likelihood function is approximated in each simulation step using a limited number of MPS proposals and (3) the uninformed model parameters are assumed to be normally-distributed when approximating the likelihood. To assess the impacts of these approximations on the simulation results, we first consider a training image depicting a multi-Gaussian field for which the posterior is known analytically. We then consider more complex continuous and binary channelised training images for which comparisons are made, in terms of computational effort and accuracy, with respect to a sequential Gibbs sampler. Finally, we introduce an extension of our approach to non-linear physical responses and show preliminary results.

The paper is organised as follows: Section 2 provides a detailed explanation of the underlying theory; Section 3 details the metrics and the comparative approach used to assess the quality of the results; Section 4 presents the results obtained when considering a linear physical solver as well as initial results for a non-linear solver; Section 5 discusses the results, highlighting the limitations, advantages and possible future developments. Finally, Section 6 provides conclusions.

2. Methods

The proposed methodology allows conditioning MPS simulations to both point data (e.g., facies) and indirect data (geophysical measurements). Our implementation relies on QuickSampling (Gravey and Mariethoz, 2020) as the MPS algorithm, and the indirect geophysical data in the chosen examples are given by synthetic crosshole ground-penetrating radar (GPR) simulations. In the following subsections, a detailed description of the method (Section 2.1) is provided, focusing on the approximation of the likelihood and how to perform a fast update of the kriging mean and covariance. This description is then followed by a concise summary of the QuickSampling algorithm (Section 2.2) and the considered forward operators (Section 2.3).

2.1. Bayesian formulation for conditional sequential simulation

Consider a discretised random field parameterised using a grid with M cells $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_M)$. The true field is considered as a realisation $\boldsymbol{\vartheta}$ of the random field. The aim is to simulate realisations of $\boldsymbol{\theta}$ that are conditioned to indirect geophysical data $\mathbf{d}$. To achieve this, a simulation grid $S(x)$ is used with nodes $(x_1, \dots, x_M)$ defining the locations corresponding to the M grid cells in the discretised random field $\boldsymbol{\theta}$. The method begins with an empty simulation grid $S(x)$ and during the simulation, the cells are populated one by one with values. If any (hard) conditioning points $\theta_i = \vartheta_i$ are known, they are assigned to the simulation grid in $S(x_i)$ before the simulation begins. If the hard data is available at a finer scale than the resolution of the simulation grid, then it is necessary to apply upscaling techniques (Zhang, 2015). The simulation path $\mathbf{p}$ (order of simulated locations) is generated randomly to maximise the variability of the realisations.

At each simulation step, there are three types of grid cells: informed cells $\boldsymbol{x}_I$, corresponding to locations that were populated in previous steps with the values $S(\boldsymbol{x}_I) = \vartheta_I$ (or are related to hard data); a simulated cell x_S, corresponding to the cell that is simulated in the current step of the algorithm; uninformed cells $\boldsymbol{x}_U$, corresponding to the empty grid cells that have not yet been simulated. To select the value for the simulated cell x_S, we rely on the distribution of the underlying random field $\boldsymbol{\theta}$, introducing θ_S, $\boldsymbol{\theta}_I$ and $\boldsymbol{\theta}_U$ as the random variables associated with the cells at locations x_S, $\boldsymbol{x}_I$ and $\boldsymbol{x}_U$. In each simulation

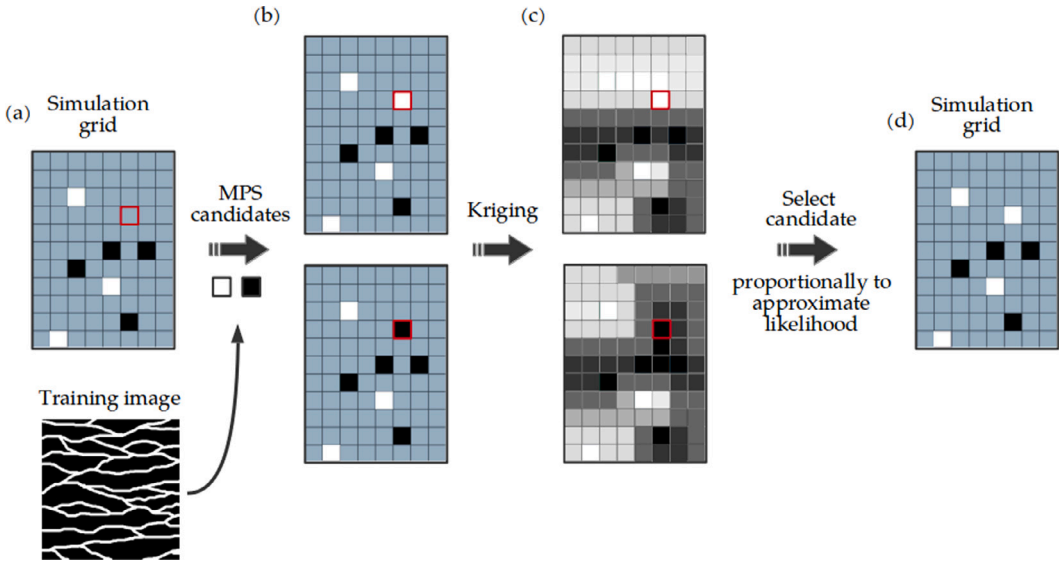

Fig. 1. Schematic illustration of one IDCS simulation step for a binary model. (a) The simulation grid with the simulated location marked by a red square, informed locations marked as either white or black filled squares and uninformed locations marked in blue. (b) The MPS algorithm proposes $k = 2$ candidates sampled from the training image that are conditional on the pattern in the simulation grid. (c) Kriging is used to estimate uninformed grid cells that are conditioned on informed and simulated (MPS candidates) locations. Kriging provides k kriging means and a single kriging covariance $\tilde{\Sigma}_\theta$, which are then used to calculate an approximate likelihood using Eqs. (9)–(11) given the geophysical data. (d) One MPS candidate is drawn proportionally to the approximate likelihood and assigned to the simulated location.

step, one of k candidate values $\mathcal{Q}_{MPS} = \{\vartheta_S^1, \dots, \vartheta_S^k\}$ proposed for the simulated cell x_S is chosen proportionally to the unnormalised posterior distribution of θ_S that is conditional on observed data $\mathbf{d}$ and the values of the previously informed cells,

$$p(\vartheta_S|\mathbf{d}, \boldsymbol{\theta}_I = \vartheta_I) \propto p(\mathbf{d}|\theta_S = \vartheta_S, \boldsymbol{\theta}_I = \boldsymbol{\vartheta}_I) p(\vartheta_S|\boldsymbol{\theta}_I = \boldsymbol{\vartheta}_I). \tag{1}$$

As the conditional prior $p(\vartheta_S|\boldsymbol{\theta}_I = \boldsymbol{\vartheta}_I)$ cannot be computed explicitly, an MPS algorithm is relied upon to sample from it. In general, MPS algorithms generate samples from a conditional distribution that preserves higher-order statistics. These algorithms scan the training image and compare its patterns to that of a defined neighbourhood around the simulated location x_S in the simulation grid $S(x)$. However, the proposed methodology is not confined to MPS and it is adaptable to any algorithm capable of generating multiple samples from a conditional prior.

The likelihood in Eq. (1), which is a marginalised likelihood over all uninformed parameters $\boldsymbol{\theta}_U$:

$$p(\mathbf{d}|\theta_S = \vartheta_S, \boldsymbol{\theta}_I = \boldsymbol{\vartheta}_I) = \int p(\mathbf{d}|\theta_S = \vartheta_S, \boldsymbol{\theta}_I = \boldsymbol{\vartheta}_I, \boldsymbol{\theta}_U = \boldsymbol{\vartheta}_U) p(\boldsymbol{\vartheta}_U|\theta_S = \vartheta_S, \boldsymbol{\theta}_I = \boldsymbol{\vartheta}_I) d\boldsymbol{\vartheta}_U, \tag{2}$$

is intractable as the forward response depends on the whole property field, but the response contribution from the uninformed parameters $\boldsymbol{\theta}_U$ is unknown. To circumvent this problem, an approximation of the likelihood is derived below by estimating the distribution of the uninformed $\boldsymbol{\theta}_U$ parameters conditional on informed $\boldsymbol{\theta}_I = \boldsymbol{\vartheta}_I$ and simulated $\theta_S = \vartheta_S$ parameters.

2.1.1. Likelihood approximation

The distribution of the uninformed parameters $\boldsymbol{\theta}_U$ needed for the likelihood approximation is approximated by kriging-based interpolation (Matheron, 1963). Kriging assumes a mean $m_\theta(x)$ and a stationary covariance $\sigma_\theta(x_i, x_j)$ function describing the correlation between locations x_i and x_j that are separated by some distance and angle. To build the covariance model, kriging relies on theoretical variograms (Oliver and Webster, 1990). Here simple kriging is used, in which the mean of the property of interest is assumed to be known and the values of uninformed locations are estimated based on conditioning to informed and simulated (MPS candidates) locations (Chilès and Desassis, 2018).

To estimate the distribution of $\boldsymbol{\theta}_U$, both the property field $\boldsymbol{\theta}$ and the observational noise are assumed to follow a normal distribution. Note that the Gaussian assumption on the property field is only made to approximate the likelihood (Eq. (2)), while the candidates are provided by draws from MPS-based priors. Given a multi-Gaussian field with the following prior and likelihood distributions

$$\boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta) \tag{3}$$

$$\mathbf{d}|(\boldsymbol{\theta} = \boldsymbol{\vartheta}) \sim \mathcal{N}(\mathbf{G}\boldsymbol{\vartheta}, \boldsymbol{\Sigma}_d), \tag{4}$$

there exists an analytical solution both for the likelihood $p(\mathbf{d}|\boldsymbol{\theta} = \boldsymbol{\vartheta})$ and posterior $p(\boldsymbol{\vartheta}|\mathbf{d})$ PDF (see Appendix A). In Eq. (3), $\boldsymbol{\mu}_\theta = (m_\theta(x_i))_{1\le i\le M}$ and $\boldsymbol{\Sigma}_\theta = (\sigma_\theta(x_i, x_j))_{1\le i,j\le M}$ represent the mean vector and covariance matrix of the grid cells. In Eq. (4) it is assumed that the forward response can be simplified such that a general forward operator $g(\cdot)$ can be replaced by $\mathbf{G}$, the linear forward operator of the physical response, wherein the term $\mathbf{G}\boldsymbol{\vartheta}$ corresponds to the expected value of the data. $\boldsymbol{\Sigma}_d$ in Eq. (4) is the covariance matrix of the data errors. The $\boldsymbol{\theta}_c = (\theta_S, \boldsymbol{\theta}_I)$ is introduced for locations $\boldsymbol{x}_c = (x_S, \boldsymbol{x}_I)$ and the distribution of the random field is re-written as

$$\begin{pmatrix} \boldsymbol{\theta}_c \\ \boldsymbol{\theta} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \boldsymbol{\mu}_{\theta_c} \\ \boldsymbol{\mu}_\theta \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{\theta_c} & \boldsymbol{\Sigma}_{\theta_c\theta} \\ \boldsymbol{\Sigma}_{\theta\theta_c} & \boldsymbol{\Sigma}_\theta \end{pmatrix} \right), \tag{5}$$

where $\boldsymbol{\mu}_{\theta_c}$ and $\boldsymbol{\mu}_\theta$ are the mean vectors for the respective positions, $\boldsymbol{\Sigma}_{\theta_c}$ and $\boldsymbol{\Sigma}_\theta$ are the covariance matrices whose (i, j) entries are the covariances between the ith and jth positions in $\boldsymbol{\theta}_c$ or $\boldsymbol{\theta}$, respectively, and $\boldsymbol{\Sigma}_{\theta_c\theta}$ refers to the covariance matrix consisting of the covariance values between $\boldsymbol{\theta}_c$ and $\boldsymbol{\theta}$. The conditional distribution $\boldsymbol{\theta}|(\boldsymbol{\theta}_c = \boldsymbol{\vartheta}_c) \sim \mathcal{N}(\tilde{\boldsymbol{\mu}}_\theta, \tilde{\boldsymbol{\Sigma}}_\theta)$ can then be calculated as follows (Prince, 2012):

$$\tilde{\boldsymbol{\mu}}_\theta = \boldsymbol{\mu}_\theta + \boldsymbol{\Sigma}_{\theta\theta_c} \boldsymbol{\Sigma}_{\theta_c}^{-1} (\boldsymbol{\vartheta}_c - \boldsymbol{\mu}_{\theta_c}), \tag{6}$$

$$\tilde{\boldsymbol{\Sigma}}_\theta = \boldsymbol{\Sigma}_\theta - \boldsymbol{\Sigma}_{\theta\theta_c} \boldsymbol{\Sigma}_{\theta_c}^{-1} \boldsymbol{\Sigma}_{\theta_c\theta}. \tag{7}$$

For informed and simulated parameters θ_c, the entries in the kriging mean $\tilde{\mu}_\theta$ retain the same values as prior to kriging and their corresponding entries in the kriging covariance matrix $\tilde{\Sigma}_\theta$ are zero. The multiplication $\Sigma_{\theta\theta_c}\Sigma_{\theta_c}^{-1}$ yields the kriging weights that provide the necessary information for interpolating from known grid points (at locations $\boldsymbol{x}_I$ and x_S) to unknown points (at location $\boldsymbol{x}_U$).

At each simulation step, k candidate values $\mathcal{Q}_{MPS}$ are considered and k kriging means and a single covariance matrix (see Fig. 1 for illustration) are obtained. The likelihood of each candidate value $p(\mathbf{d}|\theta_c = \vartheta_c)$ is estimated as a Gaussian PDF with mean $\tilde{\mu}_L$ and covariance matrix $\tilde{\Sigma}_L$ (Bishop and Nasrabadi, 2006):

$$\mathbf{d}|(\theta_c = \vartheta_c) \sim \mathcal{N}(\tilde{\mu}_L, \tilde{\Sigma}_L), \tag{8}$$

$$\tilde{\mu}_L = \mathbf{G}\tilde{\mu}_\theta, \tag{9}$$

$$\tilde{\Sigma}_L = \Sigma_d + \mathbf{G}\tilde{\Sigma}_\theta\mathbf{G}^T. \tag{10}$$

For each candidate value and corresponding kriging mean, the forward response is calculated to generate $\tilde{\mu}_L$ (Eq. (9)). Additionally, the kriging error is incorporated into the error covariance of the likelihood $\tilde{\Sigma}_L$ (Eq. (10)). Finally, the likelihood can be approximated as:

$$p(\mathbf{d}|\theta_c = \vartheta_c) = p(\mathbf{d}|\theta_S = \vartheta_S, \boldsymbol{\theta}_I = \boldsymbol{\vartheta}_I) \approx \frac{1}{\sqrt{(2\pi)^{N_d}|\tilde{\Sigma}_L|}} \exp\left(-\frac{1}{2}\left[\mathbf{d} - \tilde{\mu}_L\right]^T \tilde{\Sigma}_L^{-1}\left[\mathbf{d} - \tilde{\mu}_L\right]\right), \tag{11}$$

where N_d is the number of observed indirect data and $|\tilde{\Sigma}_L|$ is the determinant of $\tilde{\Sigma}_L$. The value assigned to the simulated location is drawn proportionally to the approximate likelihoods of the k proposed values. This is achieved by drawing randomly from the cumulative distribution function (CDF) of the approximate likelihoods.

To estimate the covariance structure of the Gaussian prior distribution in Eq. (5), the GSTools Python library (Müller et al., 2022) is employed to automatically fit a theoretical variogram using samples from the training image. Based on these samples, which in this study are limited to 30 000, GSTools provides the standard deviation, integral scale in two directions and the shape parameter of the fitted model. In this paper, the term "training image" refers to either a complete image or a portion of an image used in the simulation process.

2.1.2. Fast update of the conditional mean and covariance

To gain computational efficiency by avoiding re-computing Eqs. (6)–(7) at each simulation step, the fast kriging update approach is adopted. This approach was introduced by Emery (2009) and later extended by Chevalier et al. (2014, 2015). Chevalier et al. (2015) used it to assimilate new observation points to sequential simulations. Their technique enables a fast update of the kriging mean $\tilde{\mu}_\theta$ and the kriging covariance $\tilde{\Sigma}_\theta$ given new conditioning data points. Instead of calculating the conditional mean and covariance from scratch at each simulation step, the previous estimate is perturbed given the newly simulated value. To maintain consistency, the general notation used in the previous subsection is adopted and the kriging update equations are expressed as a function of the simulation step, denoted as t. In this notation, $\boldsymbol{\theta}$ becomes $\boldsymbol{\theta}^{(t)}$ and $\boldsymbol{\theta}_c^{(t)} = (\boldsymbol{\vartheta}_I^{(t)}, \vartheta_S^{(t)})$. The conditional distribution then becomes $\boldsymbol{\theta}^{(t)}|(\boldsymbol{\theta}_c^{(t)} = \boldsymbol{\vartheta}_c^{(t)}) \sim \mathcal{N}(\tilde{\boldsymbol{\mu}}_\theta^{(t)}, \tilde{\boldsymbol{\Sigma}}_\theta^{(t)})$.

Expressing $\boldsymbol{\mu}_\theta = (m_\theta(x_i))_{1\le i\le M}$ and $\tilde{\boldsymbol{\mu}}_\theta = (\tilde{m}_\theta(x_i))_{1\le i\le M}$, $\boldsymbol{\Sigma}_\theta = (\sigma_\theta(x_i, x_j))_{1\le i,j\le M}$ and $\tilde{\boldsymbol{\Sigma}}_\theta = (\tilde{\sigma}_\theta(x_i, x_j))_{1\le i,j\le M}$ and re-writing Eqs. (6)–(7) with respect to step t and the location x lead to,

$$\tilde{m}_\theta^{(t)}(x) = m_\theta(x) + \boldsymbol{\sigma}_\theta(\boldsymbol{x}_c^{(t)}, x)^T \boldsymbol{\sigma}_\theta(\boldsymbol{x}_c^{(t)}, \boldsymbol{x}_c^{(t)})^{-1}(\boldsymbol{\vartheta}_c^{(t)} - m_\theta(\boldsymbol{x}_c^{(t)})), \tag{12}$$

$$\tilde{\sigma}_\theta^{(t)}(x_i, x_j) = \sigma_\theta(x_i, x_j) - \boldsymbol{\sigma}_\theta(\boldsymbol{x}_c^{(t)}, x_i)^T \boldsymbol{\sigma}_\theta(\boldsymbol{x}_c^{(t)}, \boldsymbol{x}_c^{(t)})^{-1}\boldsymbol{\sigma}_\theta(\boldsymbol{x}_c^{(t)}, x_j). \tag{13}$$

Using the fast update, $\tilde{m}_\theta^{(t)}(\boldsymbol{x})$ is computed by perturbing the kriging mean resulting from the previous step $\tilde{m}_\theta^{(t-1)}(x)$, according to the difference between the value of the MPS candidate $\vartheta_S^{(t)}$ and the value at location x_S in $\tilde{m}_\theta^{(t-1)}$:

$$\tilde{m}_\theta^{(t)}(x) = \tilde{m}_\theta^{(t-1)}(x) + \tilde{\sigma}_\theta^{(t-1)}(x_S^{(t)}, x)^T(\tilde{\sigma}_\theta^{(t-1)}(x_S^{(t)}, x_S^{(t)}))^{-1}(\vartheta_S^{(t)} - \tilde{m}_\theta^{(t-1)}(x_S^{(t)})), \tag{14}$$

where $\tilde{\sigma}_\theta^{(t-1)}(x_S^{(t)}, x)^T(\tilde{\sigma}_\theta^{(t-1)}(x_S^{(t)}, x_S^{(t)}))^{-1}$ represents the kriging weight. The update to the conditional covariance is based on the same kriging weight

$$\tilde{\sigma}_\theta^{(t)}(x_i, x_j) = \tilde{\sigma}_\theta^{(t-1)}(x_i, x_j) - \tilde{\sigma}_\theta^{(t-1)}(x_S^{(t)}, x_i)^T(\tilde{\sigma}_\theta^{(t-1)}(x_S^{(t)}, x_S^{(t)}))^{-1}\tilde{\sigma}_\theta^{(t-1)}(x_S^{(t)}, x_j). \tag{15}$$

Once the conditional mean and covariance are updated, they are plugged into Eqs. (9) and (10). For a summary of our full conditioning algorithm, see Algorithm 1. The algorithm describes a single run of the methodology which is referred to as Indirect Data Conditional Simulations (IDCS) for the rest of the paper.

Algorithm 1: Indirect Data Conditional Simulations (IDCS) with a general MPS algorithm

```
Input: simulation grid S(x) with M grid cells (either empty or informed by hard conditioning data), training image T, simulation path p, observed data d, number of candidates k and algorithm-specific MPS parameters b.
Output: fully informed grid with property field ϑ
set simulation step t = 1
for t <=M do
    x_S = p_t
    Function MPS(S(x), T, x_S, k, b)
        Sample candidate values from T that are conditional on the ϑ_I around location x_S in the simulation grid.
        return Q_MPS
    if t=1 then
        Compute μ̃_θ^(t) and Σ̃_θ^(t) using Eqs. (5)-(7) for all k candidates
    else
        Update μ̃_θ^(t) and Σ̃_θ^(t) using Eqs. (14)-(15) for all k candidates
    end
    μ̃_L, Σ̃_L ← compute Eqs. (9)-(10) for all k candidates
    Approximate p(d|θ_c = ϑ_c) ← compute Eq. (11) for all k candidates
    Calculate cumulative distribution function (CDF) of k likelihoods and draw one value from Q_MPS
    Populate S(x_S) with the selected candidate value ϑ_S
    t = t + 1
end
```

2.2. QuickSampling algorithm

QuickSampling (QS) is used to demonstrate the proposed IDCS methodology. It is a computationally efficient pixel-based algorithm that in contrast to many other pixel-based MPS algorithms, does not store conditional distributions (Strebelle, 2002; Straubhaar et al., 2011) or rely on threshold criteria for choosing a candidate (Mariethoz et al., 2010b). In this algorithm, the training image, denoted as T, is scanned to find a close match to the pattern around the simulated location x_S in the simulation grid $S(x)$. The pattern is represented by values and their relative positions with respect to x_S. At each simulation step a neighbourhood in S, denoted as N, is considered; $N(x)$ is centred around the currently simulated grid cell and contains within a specified radius the locations in the simulation grid that are previously informed (either previously simulated or conditioning data points).

In QS, the cross-correlation between $N(x)$ and T is calculated to generate a dissimilarity (mismatch) map E:

$$E(T, N(x)) \propto \mathcal{F}^{-1}\left\{\sum_{i\in I}\sum_{j\in J}\overline{\mathcal{F}\{\mathbb{1}(T_i)\circ f_j(T_i)\}}\circ\mathcal{F}\{w_i\circ\mathbb{1}(x_i)\circ h_j(N(x)_i)\}\right\}, \tag{16}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ are the fast Fourier transform and its inverse, $\overline{\mathcal{F}\{\}}$ is the conjugate and $\circ$ indicates an element-wise multiplication. Furthermore, w is a weighting matrix and it can be set to assign different weights as a function of the distance from x_S. The indicator function $\mathbb{1}$ equals 1 at informed grid cells and 0 everywhere else, that is,

$$\mathbb{1}(x) = \begin{cases} 1, & \text{if } N(x) \text{ is informed} \\ 0, & \text{otherwise} \end{cases}. \tag{17}$$

The variables f_j and h_j represent components of a series of decomposed functions that depend on the distance metric used (see Gravey and Mariethoz (2020) for more information). In the original implementation, candidates are sorted in increasing order of mismatch and the simulated value is sampled proportionally to a probability determined by a user-defined rank k_{rank}. In addition to k_{rank}, the QS algorithm requires a user-defined parameter n that determines the number of neighbouring locations around x_S, and effectively the size of the neighbourhood $N(x)$ to be considered for MPS conditioning. In our implementation, the QS algorithm functions solely as a means to sample k conditional draws from the prior that are evaluated using the approximate likelihood. Therefore, the QS parameter k_{rank} is replaced in IDCS by k_{cand} which represents the number of candidates provided by the QS algorithm ($k = k_{cand}$).

2.3. Forward response

To test the IDCS, a crosshole geometry is considered in which GPR antennas placed in two boreholes are used to emit and receive electromagnetic signals and the first-arrival travel-times between different source and receiver pairs are measured. In this setting, $\boldsymbol{\theta}$ is a random slowness field (inverse of the velocity). Specifically, a ray-based formulation is used in which the travel-time t_{ray} is an integration of slowness over the ray path l:

$$t_{ray} = \int \vartheta(l)dl. \tag{18}$$

The aforementioned physical response can be written in a general form as

$$\mathbf{d} = g(\boldsymbol{\vartheta}) + \boldsymbol{\epsilon}, \tag{19}$$

where $g(\cdot)$ is the forward operator projecting the parameters $\boldsymbol{\vartheta}$ in the model space into a vector $\mathbf{d}$ in the data space and the process typically involves some type of error $\boldsymbol{\epsilon}$. Here only uncorrelated, randomly distributed Gaussian (measurement) noise under $\boldsymbol{\epsilon}$ is considered.

2.3.1. Linear physical response

Considering linear physics, Eq. (18) can be simplified into $t_{ray} \approx \sum_i l_i \cdot \vartheta_i$ and the response becomes a matrix–vector multiplication operation

$$\boldsymbol{d}^{sim} = \boldsymbol{G}\boldsymbol{\vartheta}, \tag{20}$$

where $\mathbf{G}$ (also referred to as the sensitivity matrix) contains the ray length in each grid cell considering only a straight path between the source and receiver. The simulated data $\mathbf{d}^{sim}$ are represented in a vector containing each source–receiver response in the form of first-arrival travel-times.

2.3.2. Non-linear physical response

When dealing with a non-linear physical response, an approximation to Eq. (19) is required in order to calculate Eqs. (9)–(10). This is achieved by linearising the forward operator $g(\boldsymbol{\vartheta})$ around a given subsurface model to obtain the sensitivity matrix (Jacobian). In general, the Jacobian represents the gradient around the point of linearisation and, therefore, the forward response is calculated with respect to a reference point $\boldsymbol{\vartheta}_0$:

$$\boldsymbol{d}^{sim} = g(\boldsymbol{\vartheta}_0) + \boldsymbol{J}(\boldsymbol{\vartheta}_0)(\boldsymbol{\vartheta} - \boldsymbol{\vartheta}_0), \tag{21}$$

where $\boldsymbol{J}(\boldsymbol{\vartheta}_0)$ is the Jacobian calculated for the subsurface model $\boldsymbol{\vartheta}_0$. In travel-time tomography, the elements in a row of the Jacobian represent the length of the ray segment in each grid cell of the model for a specific ray path. Therefore, the forward operator can be replaced by the Jacobian to calculate the physical response $\boldsymbol{d}^{sim} = \boldsymbol{J}\boldsymbol{\vartheta}$, where $\boldsymbol{J}$ is the Jacobian given the slowness field $\boldsymbol{\vartheta}$. In this case, $\mathbf{G}$ in Eqs. (9) and (10) is simply replaced with $\boldsymbol{J}$ to obtain:

$$\tilde{\boldsymbol{\mu}}_L = \boldsymbol{J}\tilde{\boldsymbol{\mu}}_\theta \tag{22}$$

$$\tilde{\boldsymbol{\Sigma}}_L = \boldsymbol{\Sigma}_d + \mathbf{J}\tilde{\boldsymbol{\Sigma}}_\theta\mathbf{J}^T. \tag{23}$$

3. Comparative approach and quality assessment criteria

3.1. Sequential Gibbs sampling

To assess the quality and performance of the IDCS method when no analytical solution is available, it is compared against results obtained with the extended Metropolis algorithm (Mosegaard and Tarantola, 1995) using a sequential Gibbs sampler (Hansen et al., 2012; Cordua et al., 2012).

The extended Metropolis algorithm allows exploring the posterior PDF when dealing with a prior distribution of arbitrary complexity from which only samples can be drawn. In this algorithm, the acceptance or rejection of a model proposal $\boldsymbol{\vartheta}'$ is determined by the acceptance probability $P_{accept} = \min\left(1, \frac{p(\mathbf{d}|\theta'=\vartheta')}{p(\mathbf{d}|\theta^{(t)}=\vartheta^{(t)})}\right)$, where $\vartheta^{(t)}$ represents the current model. If accepted $\theta^{(t+1)} = \vartheta'$, else $\theta^{(t+1)} = \vartheta^{(t)}$. Gibbs sampling takes a given realisation $\boldsymbol{\vartheta}$ and at each iteration computes the conditional distribution at a random position i

$$p(\theta_i|\theta_1 = \vartheta_1, \ldots, \theta_{i-1} = \vartheta_{i-1}, \theta_{i+1} = \vartheta_{i+1}, \ldots, \theta_M = \vartheta_M). \tag{24}$$

A value for θ_i is then drawn from the conditional distribution forming the new realisation. Sequential Gibbs sampling combines sequential simulations with the Gibbs sampler such that one can generate realisations from the conditional distribution.

In this study, the extended Metropolis algorithm is used to estimate the posterior PDF and compare it to the approximate posterior computed based on realisations obtained from running the IDCS. The QS algorithm is used to initialise the MCMC chains with unconditional MPS realisations and to generate conditional model proposals in subsequent MCMC steps. At each MCMC step, a random subset of the model domain is re-simulated while being conditioned on the remaining part of the domain. The size of the subset is adapted during the first 2000 MCMC steps, within the so-called "burn-in" period (a period characterised by low likelihoods and dependence on the initial model state), using the parameter δ. After burn-in, the value of δ remains constant to maintain detailed balance of the Markov chain. The parameter δ represents half of the side-lengths of a square centred at a grid point chosen at random. The size of δ is used to control the step length of the sequential Gibbs sampler, where a small value leads to a high acceptance rate with highly correlated chains and larger values lead to lower acceptance rates but less correlated chains (Hansen et al., 2012). During burn-in, the value of δ is adjusted every 20 iterations according to

$$\delta_{new} = \delta * \frac{\bar{P}_{acc}}{P_{target}}, \tag{25}$$

with the aim of maintaining a reasonable acceptance rate (Gelman et al., 1996; Cordua et al., 2012). The variable $\bar{P}_{acc}$ is the average acceptance rate between adjustment steps and P_{target} is the target acceptance

rate, which is set to 0.3. The Gelman–Rubin diagnostic (Gelman and Rubin, 1992), denoted as $\hat{R}$, is used. This metric compares the variance between the independent chains and within the chains, to assess the convergence of the MCMC chains to a stationary distribution for each of the model parameters. Convergence is declared when $\hat{R} \leq 1.2$ for all considered parameters (grid cell values).

3.2. Performance metrics

To determine the optimal algorithmic settings and to assess the quality of the resulting posterior realisations, the structural similarity index (SSIM; Wang et al., 2004) with respect to the reference subsurface model, and the weighted root-mean-squared error (WRMSE) with respect to the synthetic data are calculated. The SSIM metric indicates the structural similarity between two images. It is defined as:

$$SSIM(\mathbf{u},\mathbf{v}) = \frac{(2\mu_{\mathbf{u}}\mu_{\mathbf{v}} + C_1)(2\sigma_{\mathbf{uv}} + C_2)}{(2\mu_{\mathbf{u}}^2 + \mu_{\mathbf{v}}^2 + C_1)(2\sigma_{\mathbf{u}}^2 + \sigma_{\mathbf{v}}^2 + C_2)}, \tag{26}$$

where $\mathbf{u}$ and $\mathbf{v}$ are $W \times W$ sliding windows of their respective $[0,1]$ normalised image, $\mu_{\mathbf{u}}$ and $\mu_{\mathbf{v}}$ are the mean values over $\mathbf{u}$ and $\mathbf{v}$, $\sigma_{\mathbf{u}}^2$ and $\sigma_{\mathbf{v}}^2$ are the respective variances of $\mathbf{u}$ and $\mathbf{v}$, $\sigma_{\mathbf{uv}}$ is the covariance between $\mathbf{u}$ and $\mathbf{v}$ and C_1 and C_2 are constants. The size of the sliding window is set to 7×7 and C_1 and C_2 are set to 0.01 and 0.03, respectively. The SSIM metric ranges from -1 to 1, where 1 indicates perfectly matching images. It is reported as a mean value across all posterior realisations. The data fit is evaluated with respect to the standard deviation of the observational noise σ_d using the WRMSE

$$WRMSE = \sqrt{\frac{1}{N_d}\sum_{i=1}^{N_d}\left[\frac{d_i - d_i^{sim}}{\sigma_{d,i}}\right]^2}, \tag{27}$$

between the observed data $\mathbf{d}$ associated with the reference model and the simulated data $\mathbf{d}^{sim}$ corresponding to IDCS realisations. The reported WRMSE is the final value averaged over all simulations. A WRMSE value close to one indicates an appropriate data fit while values larger than one indicate that the data residuals are too large compared to the data noise. In addition to those two metrics, and for each test case the model realisation with the lowest root-mean-squared error (RMSE) with respect to the reference subsurface model is shown.

4. Results

Three different training images are considered: (1) a multi-Gaussian field, (2) an isotropic field with connected high values and (3) binary channels, each corresponding to a full stationary image of size 2500×2500 pixels representing an area of 250×250 m. Rather than using the entire image as a training image, a smaller section of 50×50 m is selected, serving as the training image for the QS algorithm and variogram fitting. The three 50×50 m training images are shown in Fig. 2 and their true and estimated covariance model parameters (standard deviation σ, integral scale ℓ and shape parameter ν) are reported in Table 1. The multi-Gaussian image was generated using the fast Fourier transform moving average (FFT-MA; Ravalec et al., 2000) method with an exponential model ($\nu = 1$). After generation, it was re-scaled to have a mean of 0.07 m/ns (~14.3 ns/m in slowness units) and a standard deviation of 0.004 m/ns (~0.8 ns/m). To generate an image with connected high values (Fig. 2b), the transformation by Zinn and Harvey (2003) was used. This field manipulation transforms an isotropic random Gaussian field of mean zero and unit standard deviation to a field in which high values are connected. After the transformation, the image is re-scaled to have a mean of 0.07 m/ns (~14.3 ns/m) and a standard deviation of 0.004 m/ns (~0.8 ns/m). The image with binary channels is taken from Zahner et al. (2016) and velocity values of $[0.06, 0.08]$ m/ns($[12.5, 16.6]$ ns/m) are assigned to the channels and surrounding matrix material, respectively.

Table 1
True and estimated training image statistics. Estimated values are based on a variogram fitted to 30 000 samples drawn from a training image of size 500×500 pixels.

TI	True				Estimated		
	ℓ_x/ℓ_y	ν	μ [ns/m]	σ [ns/m]	ℓ_x/ℓ_y	ν	σ [ns/m]
Gaussian	10/5	1	14.33	0.83	10.57/5.60	1.26	0.81
Connected high values	–	–	14.34	0.83	9.46/9.66	0.92	0.81
Binary channels	–	–	13.58	1.83	27.22/6.42	1.30	1.90

Following preliminary tests to determine the QS parameters (see Appendix B), n is set to 10 for the continuous models and to 25 for the binary one as they result in the best simulation quality. To mitigate the potential risk of sampling unfavourable candidates in the presence of a finite and possibly small training image, k_{cand} is set to 100 for both continuous and binary models. When performing MCMC inversion, the original implementation of QS is used and k_{rank} is set to 1.2 (it represents a probability rather than the number of candidates) and $n = 30$ as those values lead to a good simulation quality (Gravey and Mariethoz, 2020; Meerschman et al., 2013). The QS parameters remain constant throughout the IDCS simulations and the MCMC inversion.

A model domain of size 5×10 m with 0.1 m discretisation is considered yielding a total of 5000 model parameters. The forward response is computed between different antenna locations in two boreholes separated by a distance of 5 m. The borehole on the right side of the domain contains 25 source locations and the borehole on the left contains 25 receiver locations (see Fig. 3a). The antennas are located between 0.2 and 9.8 m depth with 0.4 m separating subsequent antenna positions. Ray-paths between source–receiver pairs that exceed angles of $\pm 50°$ to the horizontal are filtered out and are not considered during inversion. Consequently, the number of data points is 515. The reference model (synthetic truth) is cropped from a portion of the full 2500×2500 pixels image that remains unused during the simulation process, ensuring that there is no overlap with the training image. In all case studies, the synthetic observed data corresponding to the reference model are contaminated with normally distributed noise with mean zero and standard deviation of 1 ns.

4.1. Linear physics

We first consider the results obtained from IDCS simulations considering different subsurface models and a linear physical response.

4.1.1. Multi-Gaussian random field

A total of 100 independent IDCS runs (each with a different simulation path) are considered given noise-contaminated synthetic data corresponding to the reference model in Fig. 3a. As this example considers a multi-Gaussian property field and linear physics, the analytical solution (see Appendix A) of the posterior distribution can be computed and used for comparison. The element-wise mean and standard deviation of the analytical solution and the approximate posterior obtained by IDCS are displayed in Fig. 3. The mean obtained from running 100 IDCS simulations (Fig. 3c) is almost identical to the analytical mean (Fig. 3b) and the three IDCS posterior samples, representing the best and worst data fit (Fig. 3d and e, respectively) as well as the closest matching subsurface model realisation (Fig. 3f), are all reproducing the patterns in the reference model. The standard deviation calculated on the conditional realisations (Fig. 3h) underestimates (by 17% on average) that of the analytical solution (Fig. 3g). This underestimation is likely a consequence of using a finite training image and only 100 IDCS runs. As the simulations are conditioned on the observed data, the data misfit is expected to gradually decrease during the simulation to a WRMSE of one, representing realisations that fit the data to the noise level. This behaviour is confirmed by our results, given that the

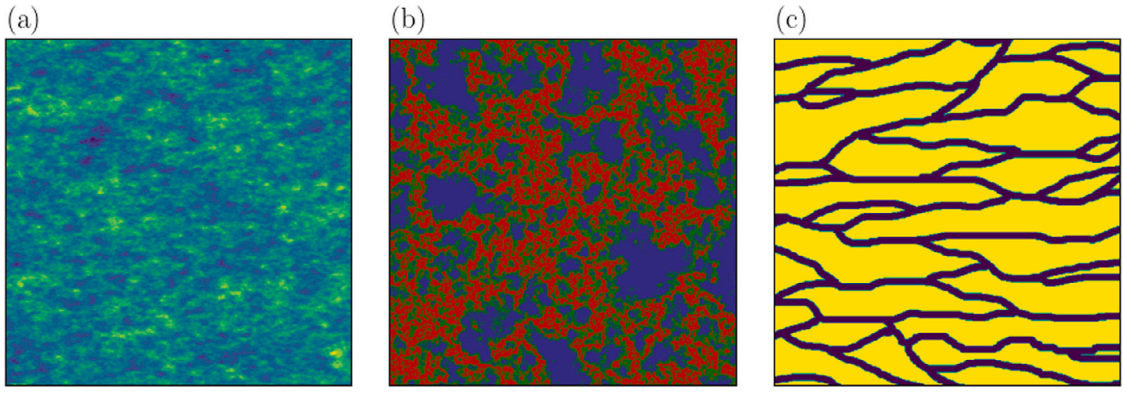

Fig. 2. Training images for the various tested models. 500 × 500 pixels (50 × 50 m) section of (a) anisotropic, random Gaussian field, (b) isotropic field with connected high values generated by applying the transformation in Zinn and Harvey (2003) and (c) binary channels.

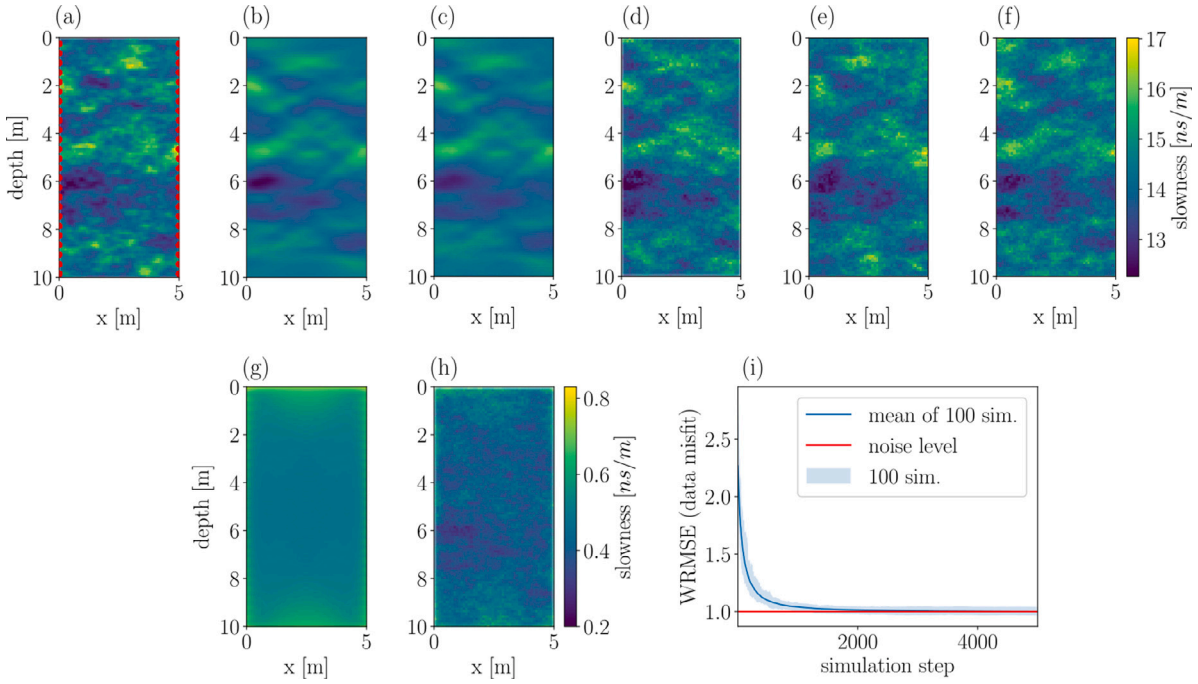

Fig. 3. IDCS results for a multi-Gaussian random field and a linear forward solver. (a) Reference model with source (left) and receiver (right) antenna locations as red filled dots, the (b) analytical and (c) approximate (IDCS) posterior pixel-wise mean, where the latter is computed on 100 IDCS realisations. IDCS realisations with the (d) lowest and (e) highest data WRMSE while (f) is the IDCS realisation with the lowest model RMSE. Pixel-wise standard deviation of the (g) analytical and (h) approximate (IDCS) posteriors. (i) The WRMSE curves of 100 IDCS runs as well as their mean.

Table 2
Summary of IDCS results for linear and non-linear physical responses. The mean SSIM as well as the median data WRMSE computed on 100 IDCS runs for the three types of models.

Physics	TI type	SSIM	WRMSE
Linear	Gaussian	0.50	1.00
	Connected high values	0.48	1.02
	Binary channels	0.86	1.02
Non-linear	Gaussian	0.46	1.03
	Connected high values	0.44	1.05
	Binary channels	0.55	2.44
	Binary channels (constant Jacobian)	0.60	2.56

median of the final WRMSE among the 100 realisations is 1.00 (Table 2) and it is already around 1.01 after simulating 1850 grid cells (Fig. 3i).

4.1.2. Isotropic field with connected high values

A total of 100 independent IDCS runs are now considered given noise-contaminated synthetic data corresponding to the reference model in Fig. 4a. Since no analytical solution is available for this case, the results are compared against eight independent MCMC chains (see Section 3.1). Computational resources are provided to permit the maximal performance of each method, namely, one CPU per chain for MCMC (eight in total) and one CPU per simulation for the conditional MPS simulations (100 in total). Both methods are executed on a cluster that is equipped with AMD EPYC™ 7402 CPUs. It took 100 min to run 100 conditional QS simulations in parallel and 26 hours to perform 20 000 MCMC steps (per chain, in total 160 000 samples). For the MCMC, a re-simulated sub-domain with a maximum size of 11 × 11 cells ($\delta = 5$) is used resulting in acceptance rate of 31% on average. The MCMC chains did not converge after 20 000 iterations per chain

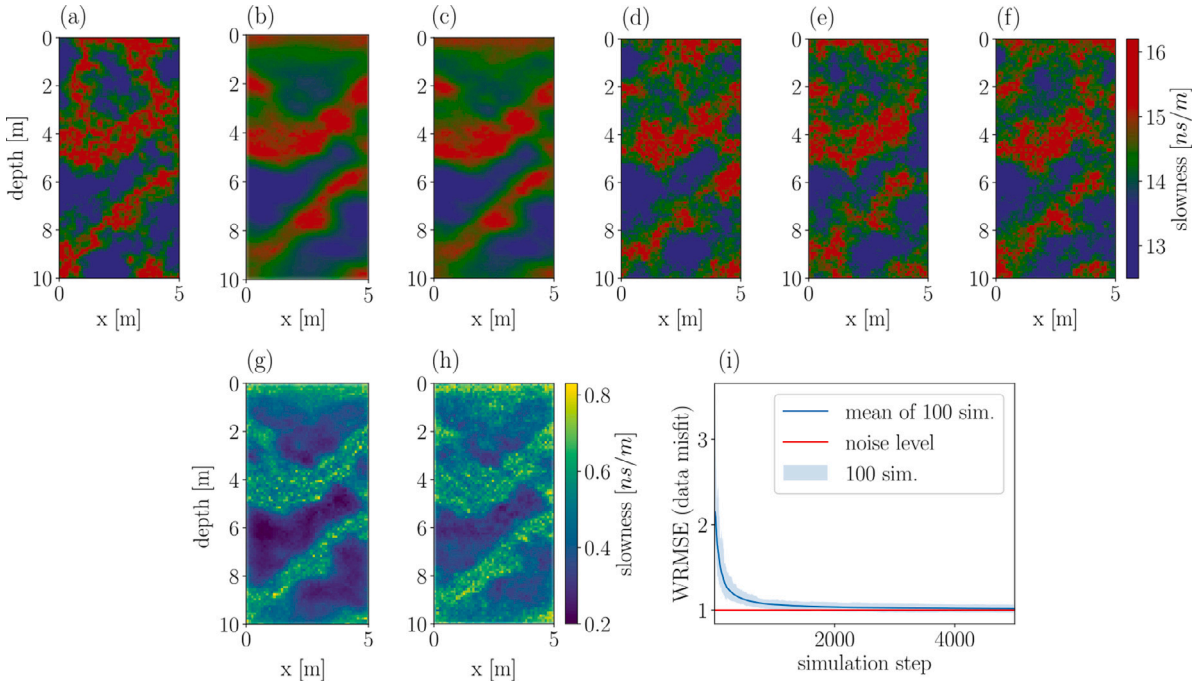

Fig. 4. IDCS results for the isotropic field with connected high values and a linear forward solver. (a) Reference model, posterior pixel-wise mean computed on (b) MCMC samples and (c) 100 IDCS realisations. IDCS realisations with the (d) lowest and (e) highest data WRMSE, while (f) is the IDCS realisation with the lowest model RMSE. Pixel-wise standard deviations of the posterior approximated by (g) MCMC samples and (h) IDCS realisations. (i) The WRMSE curves of 100 IDCS runs as well as their mean.

and the $\hat{R}$-values range from 1.27 to 8.42 with the median value being 4.54.

The reference model contains connected high-value features with different orientations and those aligned vertically present challenges in terms of identifiability in a crosshole setting. This is seen in the posterior mean of both MCMC samples (Fig. 4b) and the IDCS realisations (Fig. 4c). While the features that are horizontally-oriented are present in the estimated posterior mean obtained from the IDCS, the vertically-oriented ones are unresolved leading to a slightly lower SSIM than in the previous test case (0.48 versus 0.50, see Table 2). This is not a method-specific problem as similar SSIM values are observed for the MCMC posterior samples (on average 0.48). The IDCS realisations exhibit higher standard deviation (Fig. 4h) than those observed in the MCMC samples (Fig. 4g). In both methods, the highest standard deviation is observed where high-slowness features are present or at the locations where they are poorly resolved. This is also seen in the notable variability observed at these locations in the independent IDCS realisations displayed in Fig. 4d–f. The median WRMSE among the IDCS realisations is 1.02, indicating an overall appropriate data fit and suggesting that the vertical features are not well constrained by the data.

4.1.3. Binary channels

The last test case with linear physics is a binary reference model with channel-like structures (Fig. 5a). No analytical solution is available and, therefore, the IDCS results are again compared against those obtained from eight MCMC chains. While the computation cost of the IDCS runs is the same as in the previous example, it took 72 hours to perform 20 000 MCMC steps (per chain). The longer computational times are a result of a larger re-simulated sub-domain with a maximum size of 17 × 17 cells ($\delta = 8$) leading to an average acceptance rate of 24%. As in the previous test case, the MCMC chains did not converge after 20 000 iterations and the $\hat{R}$-values range from 1.00 to 28.92 with the median value being 1.36.

The posterior mean of the MCMC inversion (Fig. 5b) and the approximate posterior mean of the IDCS (Fig. 5c) reconstruct the channel features well. In accordance with Zahner et al. (2016), the highest uncertainty is concentrated around the boundaries of the channels (Fig. 5g and h). While most IDCS realisations correctly reproduce the channels in the reference model (e.g. Fig. 5d and f), seven IDCS realisations do not locate channel material around 6–8 m (e.g. Fig. 5e) and one misses it around 3.5–4.5 m. These eight realisations have significantly higher WRMSE values (2.70–3.37) than the remaining realisations (average WRMSE of 1.02, see Fig. 5j). These artefacts originate at an early stage when matrix material is incorrectly placed where a channel should be, thus forcing subsequent simulation steps to put additional matrix material in the vicinity such that the resulting data misfit is high. Nonetheless, the aberrant simulations are easily distinguishable from the other simulations and can be regarded as outliers. One approach is to calculate the Inter Quartile-Range statistic (IQR; Ter Braak and Diks, 2009) during the simulation and to discard those simulations that are identified as outliers. This can be done either during the simulation (to avoid redundant computation) or as a post-processing step. After simulating approximately 23% of the grid (corresponding to 1138 grid cells), the IQR statistics identify the eight aberrant simulations as outliers. If these simulations are excluded, the uncertainty estimates become comparable to those obtained from the MCMC posterior samples (Fig. 5i).

4.2. Non-linear physics

In this section, results obtained when running IDCS with a non-linear forward response (see Section 2.3.2) are presented. As the calculation of the Jacobian is generally expensive, instead of calculating $\mathbf{J}$ for each MPS candidate it is computed based on the kriging mean, given the informed grid cells at step t: $\boldsymbol{\vartheta}_I^{(t)}$. Accordingly, the number of Jacobian updates reduces to the number of grid cells to be simulated. To calculate the forward response, the pyGIMLi geophysical modelling library (Rücker et al., 2017) is used to calculate the shortest path between a source and receiver pair given a slowness model. The accuracy of the forward response depends on the number of secondary nodes on the edges of the grid cell, allowing for more ray angles. Here the

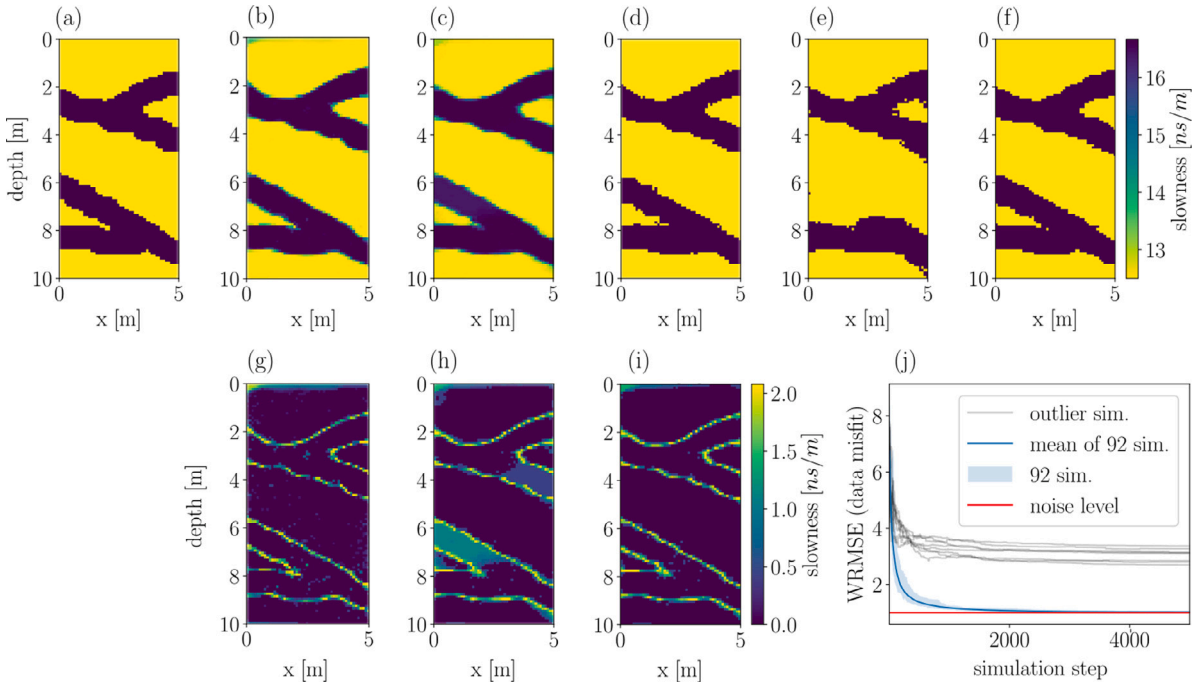

Fig. 5. IDCS results for the binary channelised field and a linear forward solver. (a) Reference model, posterior pixel-wise mean computed on (b) MCMC samples and (c) 100 IDCS realisations. IDCS realisations with the (d) lowest and (e) highest data WRMSE, while (f) is the IDCS realisation with the lowest model RMSE. Pixel-wise standard deviations of the posterior approximated by (g) MCMC samples, (h) IDCS realisations and (i) IDCS realisations excluding outliers. (j) The WRMSE curves of 92 good IDCS runs (light blue) as well as their mean (dark blue line) and eight outlier simulations identified by the IQR statistic.

number of secondary nodes used to compute the Jacobian is limited to two nodes in order to avoid too long computation times.

Running the IDCS with the aforementioned non-linear forward response takes 21 hours on average. The mean of the approximate posterior for the multi-Gaussian case (Fig. 6b) is similar to the reference model (Fig. 6a) and the standard deviations of the 100 realisations (Fig. 6f) are similar to the linear-physics case. The isotropic field with connected high values results in similar posterior mean and standard deviations (Fig. 6h and l, respectively) as with linear physics, however, structures are less connected and are more patchy (Fig. 6i–k). In both types of subsurface models, the SSIM metric was slightly reduced from 0.50 and 0.48 to 0.46 and 0.44 for the multi-Gaussian and the field with connected high values, respectively (see Table 2 and Fig. 7). Although the WRMSE is close to 1 in both cases, it increased by 0.03 compared to the WRMSE reached with linear physics.

In contrast to the continuous test cases with its satisfactory results, the application of non-linear physics to the binary channels model yields a substantial decrease in the quality of the posterior approximation in comparison with the case of linear physics. The SSIM metric is reduced from 0.86 to 0.55 and the WRMSE increased from 1.02 to 2.38. While the mean of the IDCS posterior (Fig. 8b) captures the channel structure, it is excessively smooth. Additionally, the individual realisations are of lower quality compared to those obtained with linear physics. This can also be observed in the large uncertainty on the boundaries as well as inside the channels (Fig. 8f). Correspondingly, the WRMSE curves in Fig. 9a are scattered and none of the IDCS realisations fit the data to the noise level. The underlying approximations (Gaussianity, continuity, single Jacobian update for all candidates) together with a higher level of non-linearity intensifies the aberrant simulations problem already observed for the linear physics case in Fig. 5 when considering this binary training image.

In an attempt to improve the results, the IDCS is run for the same observed data, but with a constant Jacobian that is linearised around the realisation corresponding to the lowest WRMSE (Fig. 8c). This additional run adds 100 min of computation to the total computation time (see Section 4.1.2) as the forward operator remains constant during the simulation and no update is performed. The posterior approximation is overall improved as characterised by an increase in SSIM to an average of 0.60. Moreover, the mean of the posterior (Fig. 8g) becomes better defined and the uncertainty within the channels is reduced (Fig. 8k). The channel feature within the 6–8 m range still presents a significant degree of uncertainty. Nevertheless, this specific feature posed a challenge even in the linear case, as seen in Fig. 5h. The WRMSE curves in Fig. 9b are calculated using the Jacobian linearised around the realisation in Fig. 8c. While these exhibit an overall reduction in the WRMSE, the real WRMSE of the final realisations (with the Jacobian being computed for each individual realisation) has increased from 2.44 in the first run to 2.56 in the second run (Table 2).

5. Discussion

The proposed IDCS method successfully approximates the posterior distribution when considering an ensemble of simulations (e.g., 100), given linear physics, for both Gaussian and non-Gaussian reference models. The IDCS runs are able to provide posterior approximations that are comparable in quality to those obtained with MCMC, but at a much lower computational cost. This applies also to the binary test case, which poses a greater challenge as the values being discrete, while the likelihood approximation assumes continuity and a Gaussian distribution. As a consequence of this discrepancy between the nature of the model and the estimation method, a small fraction of the IDCS runs introduce artefacts in the early stages, as matrix material is erroneously placed in locations where channel material exists in the reference model. Fortunately, these outlier simulations are easily distinguishable and can be removed at an early stage or in a post-processing step using statistical metrics for dispersion, such as the IQR.

In comparison with the block data method of Straubhaar et al. (2016), our method estimates the influence of unknown parameters and selects MPS candidates according to an approximate likelihood. When considering non-linear physical responses and provided that the forward response is differentiable and can be linearised, the Jacobian

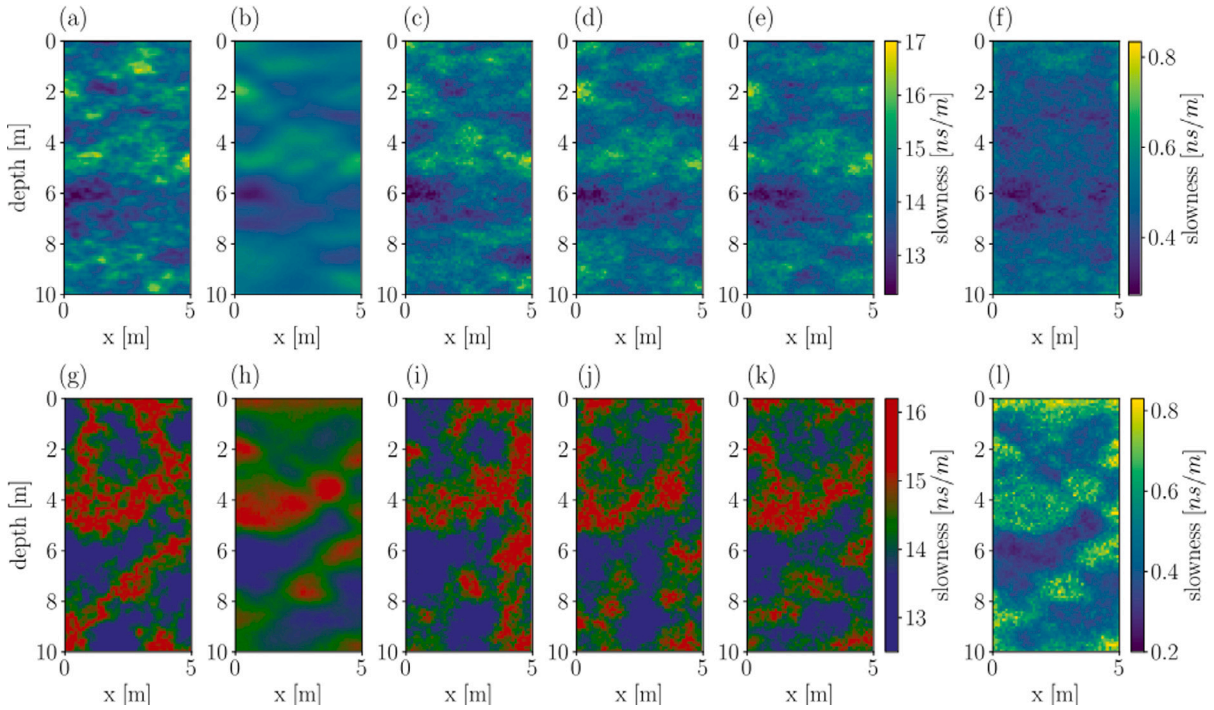

Fig. 6. IDCS results given reference models (a) and (g) and a non-linear forward solver. (b) and (h) Means of 100 IDCS realisations, (c) and (i) realisations with the lowest WRMSE, (d) and (j) realisations with the highest WRMSE, (e) and (k) realisations with the lowest model RMSE and (f) and (l) the standard deviations of 100 IDCS realisations.

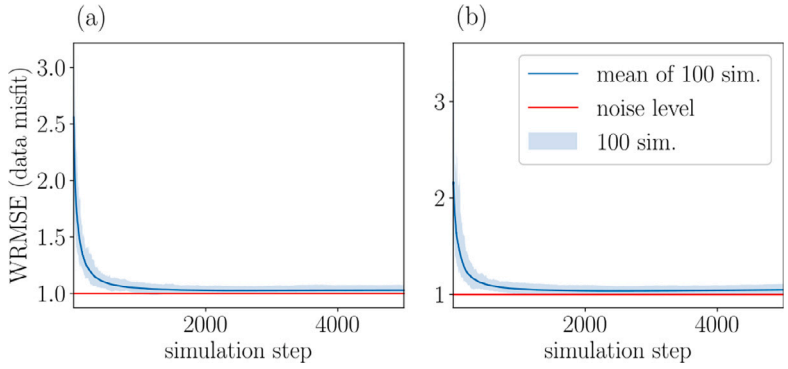

Fig. 7. Data WRMSE curves during the IDCS run given the reference model in (a) Fig. 6a and (b) Fig. 6g. The WRMSE is calculated at each simulation step using a non-linear forward solver.

can be used to obtain the kriging error and a first-order approximation of the forward response. Results for both continuous reference models suggest that the IDCS is able to approximate the posterior well. This is even the case when the forward response is non-linear, as there is only slight deterioration in SSIM and WRMSE metrics compared to the linear case. For the binary subsurface example, the issue with aberrant simulations seen in the linear case is worsened in the non-linear case likely due to a combination of the Gaussian approximation and a poor ray coverage within the channels (see Appendix C). Possible improvements could be gained by using a different likelihood approximation for binary or categorical model parameters as well as using a multigrid procedure (Caers, 2001) to improve the continuity of large scale features. The results from conducting a second run of IDCS simulations, where the Jacobian is linearised around the best data-fitting realisation from the initial run, show improvement of the posterior approximation.

The results obtained by IDCS are inherently approximate due to the finite training image (prior distribution), the limited number of candidate values considered at each simulation step, and the approximation of the intractable likelihood function. Testing of the influence of the number of candidates (k) and the number of neighbours (n) (see Appendix B) suggests that the quality of the simulation and the data fit is less sensitive to changes in n than in k. For a finite training image with a fixed size, large n (50 and above for a 500 × 500 pixels training image) can potentially lead to pattern degradation and the generation of artefacts. This is due to the limited number of distinct patterns available in the training image. On the other hand, the choice of k represents a trade-off between structure and data fit. For large k, the algorithm is forced to sample more values with decreasing pattern similarity and some of them will be accepted by the algorithm as they might lead to sufficiently low data misfit values. This means that the optimal choice of n and k depends on the size of the training image and the diversity of its patterns. Given that the algorithm is vectorised with respect to k, increasing k does not introduce additional computation time. However, the computation time increases quadratically with the size of the training image (Gravey and Mariethoz, 2020).

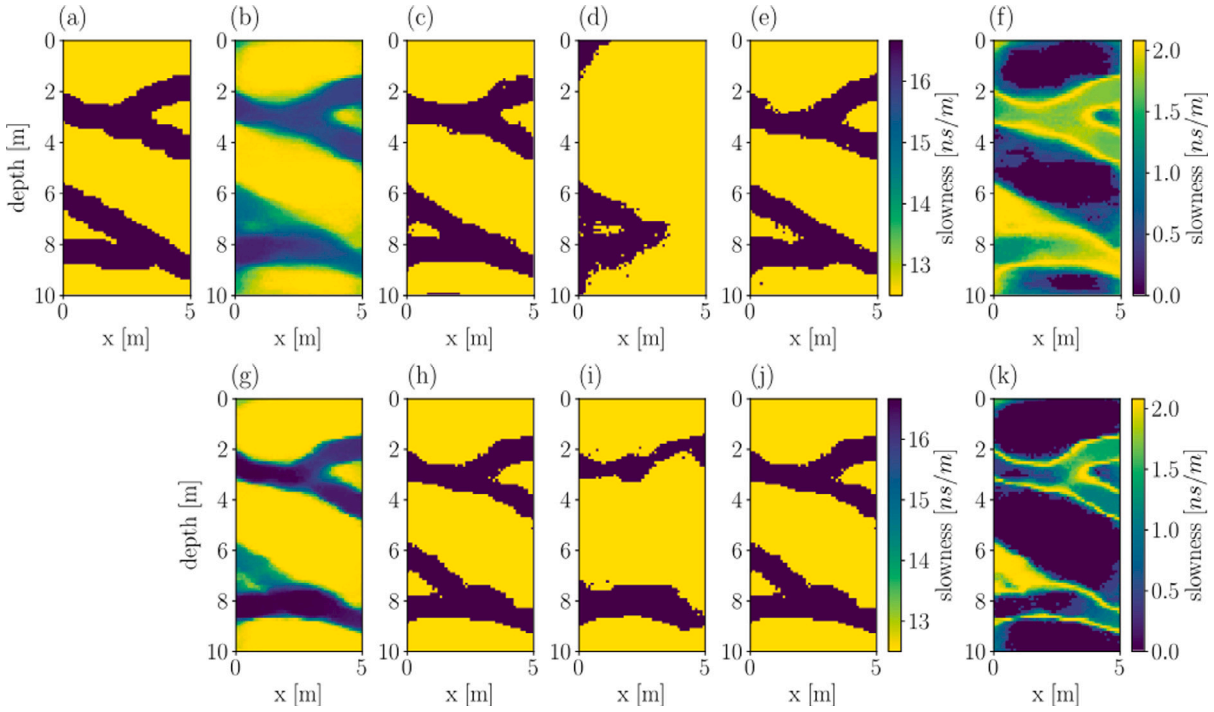

Fig. 8. IDCS results for the (a) binary channelised subsurface reference model and a non-linear forward solver considering (b)–(f) IDCS runs with the Jacobian updated according to the kriging mean and (g)–(k) considering subsequent IDCS runs with a constant Jacobian corresponding to the realisation in (c). (b) and (g) Means of 100 IDCS realisations, (c) and (h) realisations with the lowest WRMSE, (d) and (i) realisations with the highest WRMSE, (e) and (j) realisations with the lowest model RMSE and (f) and (k) the standard deviations of 100 IDCS realisations.

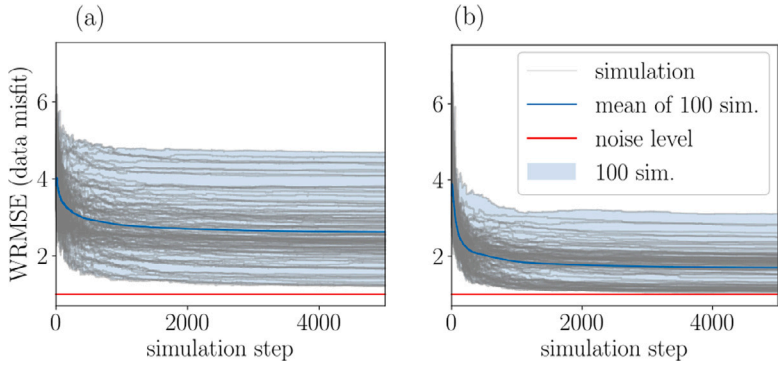

Fig. 9. Data WRMSE curves during the IDCS run given the reference model in Fig. 8a. The WRMSE in (a) is calculated at each simulation step using the linearised Jacobian around the kriging mean while in (b) the WRMSE is calculated using a constant Jacobian corresponding to the realisation in Fig. 8(c).

The computational cost of the IDCS algorithm does not scale linearly with the size of the model domain. Three factors come into play: the increase in the number of simulation steps (linear effect), the need to multiply a larger covariance matrix (Eq. (10); non-linear effect) and the need for a larger training image (non-linear effect). In our examples, the approximation of the likelihood for the linear case is responsible for 97% of the computation time of a single simulation step. Out of the time it takes to approximate the likelihood and return a simulated value, 76% is spent on computing $\tilde{\boldsymbol{\Sigma}}_L$ (Eq. (10)), 11% on the (linear) forward response (for all candidates using vectorised operations), 2% on the likelihood function (Eq. (11)) and the rest on various small operations. This suggests that advanced matrix multiplication algorithms (e.g. Nowak et al., 2003) could help enhance the efficiency of our approach and mitigate the impact of the matrix multiplication operation in Eq. (10). When compared with MCMC, IDCS is at least an order of magnitude faster for approximations of a similar quality (the MCMC runs did not converge for the considered computational budget). The number of forward simulations required in a single IDCS run depends linearly on the number of cells to simulate and on the k candidates. In contrast, the computation time of MCMC depends on the number of chains and the number of MCMC steps needed to converge, which is unknown before running the inversion. Using Gibbs sampling, the computational time is also influenced by the size of the re-simulated domain. When the physical response is non-linear, the number of Jacobian updates during the IDCS is equal to the number of grid cells to be simulated (as described in Section 4.2). Thus, IDCS provides a more predictable and efficient alternative compared to MCMC inversions provided that the associated approximations are acceptable. Furthermore, as simulations are independent, the number of simulations that are running simultaneously scale with the number of available processing units (either CPUs or GPUs). A further reduction in the computational time can be achieved by conditioning the simulation on indirect data only up to a stage where the data fit curve stabilises and changes in the data misfit are small (e.g. around 2000 steps in Fig. 7). At this stage, all necessary large-scale features are present to which the rest of the simulation is constrained. This is exemplified

in Figure 2 in Laloy et al. (2016), which suggests that when re-simulated parameters are distributed throughout the model domain, a large fraction of the domain has to be re-simulated (50% and above) to obtain significant large differences in the likelihood (and as a result in the data RMSE). The effect of such an approach on the results would need to be tested but the potential reduction in the computational time is substantial.

The IDCS method is suitable when a conceptual model of the subsurface is available in terms of a training image and the physical response is either linear or can be linearised. Examples with linear physics include: tomographic problems where parameters are integrated along straight lines such as muon (Rosas-Carbajal et al., 2017) and X-ray tomography, as well as potential-field applications such as gravity, magnetics (Blakely, 1996) and the self-potential method (Revil and Jardani, 2013). The method can of course also be used, as in this example, when a linear physics assumption might be acceptable as in GPR amplitude inversion (Jensen et al., 2022). Some additional improvement could probably be gained by using a preferential path strategy (Hansen et al., 2018; Jóhannsson and Hansen, 2023). This approach prioritises the simulation of locations that are highly constrained by the available data, such as those traversed by multiple rays and thereby, might decrease the risk of introducing artefacts. In challenging scenarios where conditional MPS simulations struggle to fit the data, particularly in cases involving categorical models, the approximation can be improved by using the MPS conditional realisations as an initial solution for MCMC chains. For instance, one can run multiple conditional MPS simulations and use the realisations that fit the data best to initialise the MCMC chains. By doing so, the burn-in period is effectively shortened and possibly, convergence is enhanced compared to using MCMC only.

6. Conclusions

We have introduced a novel approach for conditioning multiple-point statistics simulations to geophysical data represented as linear averages over the model domain. These linear averages are either constant during the simulation (linear physics) or varies as the simulation is built up (non-linear physics). Our method, named IDCS, is stochastic in nature and offers an efficient framework for approximating the posterior distribution by performing many simulation runs in parallel. The conditioning of the geophysical data is performed, for each simulated grid cell, by drawing k conditional values from the prior and accepting one of them proportionally to a kriging-based approximation of the intractable likelihood. In non-linear problem settings, the forward response has to be linearised, leading to a first-order approximation of the likelihood. Considering crosshole ground-penetrating radar data, the method was found to successfully approximate the posterior distribution for three subsurface models: multi-Gaussian, field with connected high values, and binary channels. Its main practical limitation is that the computational time scales non-linearly with the size of the model domain due to operations involving the covariance matrix. Nonetheless, for the model size tested in this paper, IDCS was found to be one to two orders of magnitude faster than MPS-based MCMC inversion considering a posterior approximation of similar quality. Possible directions for future work include more sophisticated approaches to approximate the intractable likelihood, enhance the efficiency of IDCS by exploring more sophisticated matrix multiplication techniques and use more elaborate simulation strategies (i.e. ending data conditioning when the data misfit is sufficiently low).

Computer code availability

The IDCS code is available at the following GitHub repository: https://github.com/ShiLevy/IDCS.

CRediT authorship contribution statement

Shiran Levy: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft. **Lea Friedli:** Methodology, Writing – review & editing. **Grégoire Mariéthoz:** Software, Writing – review & editing. **Niklas Linde:** Conceptualization, Methodology, Writing – review & editing, Supervision, Funding acquisition.

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.

Data availability

The IDCS code is available at the following GitHub repository: https://github.com/ShiLevy/IDCS.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used GPT-3.5 in order to improve the readability and language of the manuscript. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Acknowledgements

We would like to thank Mathieu Gravey for his expert advice and guidance concerning the QS algorithm throughout the course of this research. His support and availability for addressing our questions is greatly appreciated. We appreciate the valuable comments provided by two anonymous reviewers. This work was supported by the Swiss National Science Foundation, Switzerland (project number: 184574).

Appendix A. Analytical posterior PDF for a multi-Gaussian field

Under the assumptions of linear physics and a Gaussian prior $\boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\theta}; \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta)$, where $\boldsymbol{\mu}_\theta$ and $\boldsymbol{\Sigma}_\theta$ are the mean and covariance of the property field of interest $\boldsymbol{\theta}$, there exists an analytical expression for the posterior PDF $p(\boldsymbol{\vartheta}|\mathbf{d})$. Considering normally distributed observational noise, the likelihood can be expressed as follows:

$$\mathbf{d}|(\boldsymbol{\theta} = \boldsymbol{\vartheta}) \sim \mathcal{N}(\mathbf{d}; \mathbf{G}\boldsymbol{\vartheta} + \mathbf{b}, \boldsymbol{\Sigma}_d). \tag{28}$$

A closed-form expression for the posterior is then obtained by Bishop and Nasrabadi (2006)

$$\mathbf{d} \sim \mathcal{N}(\mathbf{d}; \mathbf{G}\boldsymbol{\mu}_\theta + \mathbf{b}, \boldsymbol{\Sigma}_d + \mathbf{G}\boldsymbol{\Sigma}_\theta\mathbf{G}^T) \tag{29}$$

$$\boldsymbol{\theta}|\mathbf{d} \sim \mathcal{N}(\boldsymbol{\theta}; \boldsymbol{\Sigma}\left\{\mathbf{G}^T\boldsymbol{\Sigma}_d^{-1}(\mathbf{d} - \mathbf{b}) + \boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{\mu}_\theta\right\}, \boldsymbol{\Sigma}), \tag{30}$$

where $\boldsymbol{\Sigma} = (\boldsymbol{\Sigma}_\theta^{-1} + \mathbf{G}^T\boldsymbol{\Sigma}_d^{-1}\mathbf{G})^{-1}$.

Appendix B. Choice of QS parameters

There are two main hyper-parameters in the QS implementation used herein: (1) k_{cand} the number of candidates proposed by QS, sorted in ascending order of mismatch (in the original implementation of Gravey and Mariethoz (2020), k_{rank} is a rank that represents the probability of sampling the sorted candidates) and (2) n the number of informed grid cells around the simulated location on which to calculate the misfit map. To determine the appropriate values for k_{cand} and n, several IDCS simulations are run for different k_{cand} and n values. The different runs are compared with respect to the SSIM and the WRMSE, and the different realisations are visually inspected. During

Table B.1
Average model SSIM and data WRMSE for 50 × 50 pixels simulation given different k values and training image size of 500 × 500; these are calculated on 10 different simulations.

k	n	TI type	SSIM	WRMSE
10	10		0.49	1.02
25	10		0.49	1.01
50	10	Gaussian	0.51	1.00
100	10		0.51	1.00
500	10		0.52	1.00
10	10		0.47	1.04
25	10		0.56	1.02
50	10	Connected high values	0.50	1.02
100	10		0.54	1.02
500	10		0.49	1.01
10	25		0.66	4.14
25	25		0.84	1.81
50	25	Channels	0.89	1.05
100	25		0.89	1.05
500	25		0.89	1.04

Table B.2
Average model SSIM and data WRMSE of 10 simulations given a connected high values model of size 50 × 50 pixels for different α values, $k = 100$ and $n = 10$.

α	SSIM	WRMSE
0	0.54	1.02
0.03	0.53	1.02
0.3	0.48	1.03

initial testing, it was observed that the optimal QS parameters for unconditional QS simulations differ from those deemed optimal for IDCS simulations and, therefore, the determination of optimal MPS parameters should be investigated within the IDCS framework.

Although large values are usually recommended for n in MPS simulations (≥ 30; Gravey and Mariethoz, 2020; Meerschman et al., 2013), it is found that n has little effect on the data fit and that better simulation quality (based on visual appearance and SSIM values) is achieved for small n (10) when the model parameters are continuous. For $n \geq 25$ the quality of the simulation decreases significantly and the realisations become noisy. For the binary channels model, a balance between the quality of the simulation (reflected in better visual appearance with less artefacts) and good model fit is found for $n = 25$. This difference in the optimal n between the continuous and binary models may be the result of the features' different characteristic sizes. It can be seen in Table 1 that the binary training images have larger correlation lengths, thus, the larger the radius within which neighbours contain relevant information.

As the value of k_{cand} increases, the WRMSE decreases and approaches 1.00 (see Table B.1). With a larger k_{cand}, there is a larger chance that one of the proposals have a high likelihood. In most of the tested models, the model's SSIM values generally show improvement when k_{cand} is increased to 100. Further increasing k_{cand} to 500 enhances the SSIM only for the Gaussian model, possibly due to the greater variety of patterns and values present in a continuous Gaussian training image. It is important to note that raising k_{cand} too much can introduce undesired artefacts. This occurs because the algorithm is forced to generate more candidates, which given a finite training image, inevitably leads to a decrease in their quality.

Additionally, Gravey and Mariethoz (2020) indicated that using a weighting kernel can improve the quality of the QS simulation. Tests with the weighting kernel $w = e^{-\alpha||d||_2}$ on the connected high values model were performed, where d is the distance from the simulated pixel, α is the kernel parameter and $||\cdot||_2$ is the Euclidean distance. This kernel gives more weight to closer neighbours as the α increases. Nonetheless, this type of kernel did not lead to improvements for our considered examples (see Table B.2).

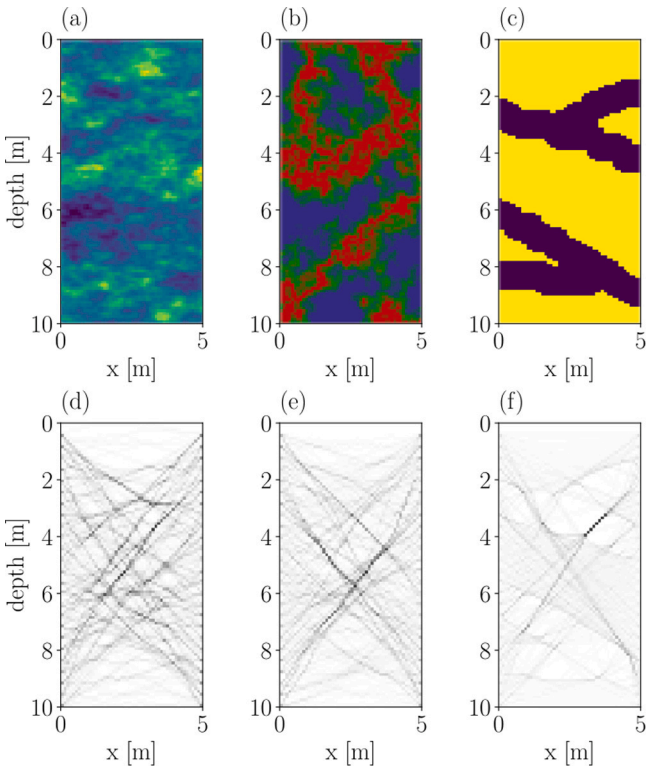

Fig. C.1. True sensitivity associated with the tested subsurface models in Section 4.2. (a)–(c) The reference models and (d)–(f) their associated ray paths.

Appendix C. Impact of the sensitivity matrix

When attempting to approximate the posterior of the binary channels subsurface model (Fig. C.1c) using linear physics, anomalous simulations are observed. This was attributed to the Gaussian approximation involved in computing the kriging mean and covariance to approximate the likelihood. This problem worsens in the presence of non-linear physics as the ray path is now a function of the slowness field. When examining the ray paths (sensitivity) given the different test models (Fig. C.1), it becomes evident that the multi-Gaussian and connected high values models exhibit a more consistent and even coverage (Fig. C.1d and e). Conversely, the binary channel model is distinguished by its elongated features with sharp boundaries having a significant impacts on the ray path and leading to poor ray coverage within the channels. This limited ray coverage (Fig. C.1f) together with the Gaussian approximation, leads to suboptimal results. When inverting the slowness values between the channels and the matrix material (i.e., channels characterised by higher velocity), the results significantly improve. Only a single outlier simulation is identified by the IQR, and after its removal the SSIM and WRMSE are 0.68 and 1.22 (compared to 0.55 and 2.44 in the original test case), respectively.

References

Alcolea, A., Renard, P., 2010. Blocking moving window algorithm: Conditioning multiple-point simulations to hydrogeological data. Water Resour. Res. 46 (8), W08511.

Barfod, A.A., Vilhelmsen, T.N., Jørgensen, F., Christiansen, A.V., Høyer, A.-S., Straubhaar, J., Møller, I., 2018. Contributions to uncertainty related to hydrostratigraphic modeling using multiple-point statistics. Hydrol. Earth Syst. Sci. 22 (10), 5485–5508.

Bishop, C.M., Nasrabadi, N.M., 2006. Pattern Recognition and Machine Learning. Springer.

Blakely, R.J., 1996. Potential Theory in Gravity and Magnetic Applications. Cambridge University Press.

Caers, J., 2001. Geostatistical reservoir modelling using statistical pattern recognition. Journal of Petroleum Science and Engineering 29 (3-4), 177–188.
Chevalier, C., Emery, X., Ginsbourger, D., 2015. Fast update of conditional simulation ensembles. Math. Geosci. 47 (7), 771–789.
Chevalier, C., Ginsbourger, D., Emery, X., 2014. Corrected kriging update formulae for batch-sequential data assimilation. In: Mathematics of Planet Earth. Springer, pp. 119–122.
Chilès, J.-P., Desassis, N., 2018. Fifty years of kriging. In: Handbook of Mathematical Geosciences: Fifty Years of IAMG. Springer International Publishing, pp. 589–612.
Cordua, K.S., Hansen, T.M., Mosegaard, K., 2012. Monte Carlo full-waveform inversion of crosshole GPR data using multiple-point geostatistical a priori information. Geophysics 77 (2), H19–H31.
Emery, X., 2009. The kriging update equations and their application to the selection of neighboring data. Comput. Geosci. 13 (3), 269.
Gelman, A., Roberts, G.O., Gilks, W.R., et al., 1996. Efficient Metropolis jumping rules. Bayesian Statist. 5 (599–608), 42.
Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. Statist. Sci. 7 (4), 457–472.
Gravey, M., Mariethoz, G., 2020. QuickSampling v1.0: a robust and simplified pixel-based multiple-point simulation approach. Geosci. Model Dev. 13 (6), 2611–2630.
Guardiano, F.B., Srivastava, R.M., 1993. Multivariate geostatistics: Beyond bivariate moments. In: Geostatistics Tróia '92: Volume 1. Springer Netherlands, pp. 133–144.
Hansen, T.M., Cordua, K.S., Mosegaard, K., 2012. Inverse problems with non-trivial priors: efficient solution through sequential Gibbs sampling. Comput. Geosci. 16 (3), 593–611.
Hansen, T.M., Journel, A.G., Tarantola, A., Mosegaard, K., 2006. Linear inverse Gaussian theory and geostatistics. Geophysics 71 (6), R101–R111.
Hansen, T.M., Mosegaard, K., 2008. VISIM: Sequential simulation for linear inverse problems. Comput. Geosci. 34 (1), 53–76.
Hansen, T.M., Mosegaard, K., Cordua, K.S., et al., 2018. Multiple point statistical simulation using uncertain (soft) conditional data. Comput. Geosci. 114, 1–10.
Høyer, A.-S., Vignoli, G., Hansen, T.M., Vu, L.T., Keefer, D.A., Jørgensen, F., 2017. Multiple-point statistical simulation for hydrogeological models: 3-D training image development and conditioning strategies. Hydrol. Earth Syst. Sci. 21 (12), 6069–6089.
Jensen, B.B., Hansen, T.M., Nielsen, L., Cordua, K.S., Tuxen, N., Tsitonaki, A., Looms, M.C., 2022. Accounting for modeling errors in linear inversion of crosshole ground-penetrating radar amplitude data: Detecting sand in clayey till. J. Geophys. Res.: Solid Earth 127 (10), e2022JB024666.
Jóhannsson, Ó.D., Hansen, T.M., 2023. Multiple-point statistics and non-colocational soft data integration. Comput. Geosci. 172, 105280.
Laloy, E., Hérault, R., Jacques, D., Linde, N., 2018. Training-image based geostatistical inversion using a spatial generative adversarial neural network. Water Resour. Res. 54 (1), 381–406.
Laloy, E., Hérault, R., Lee, J., Jacques, D., Linde, N., 2017. Inversion using a new low-dimensional representation of complex binary geological media based on a deep neural network. Adv. Water Resour. 110, 387–405.
Laloy, E., Linde, N., Jacques, D., Mariethoz, G., 2016. Merging parallel tempering with sequential geostatistical resampling for improved posterior exploration of high-dimensional subsurface categorical fields. Adv. Water Resour. 90, 57–69.
Le Coz, M., Bodin, J., Renard, P., 2017. On the use of multiple-point statistics to improve groundwater flow modeling in karst aquifers: A case study from the hydrogeological experimental site of poitiers, France. J. Hydrol. 545, 109–119.
Linde, N., Lochbühler, T., Dogan, M., Van Dam, R.L., 2015. Tomogram-based comparison of geostatistical models: Application to the macrodispersion experiment (MADE) site. J. Hydrol. 531, 543–556.
Lochbühler, T., Pirot, G., Straubhaar, J., Linde, N., 2014. Conditioning of multiple-point statistics facies simulations to tomographic images. Math. Geosci. 46 (5), 625–645.
Mariethoz, G., Caers, J., 2014. Multiple-Point Geostatistics: Stochastic Modeling with Training Images. John Wiley & Sons.
Mariethoz, G., Renard, P., Caers, J., 2010a. Bayesian inverse problem and optimization with iterative spatial resampling. Water Resour. Res. 46 (11), W11530.
Mariethoz, G., Renard, P., Straubhaar, J., 2010b. The direct sampling method to perform multiple-point geostatistical simulations. Water Resour. Res. 46 (11), W11536.
Matheron, G., 1963. Principles of geostatistics. Econ. Geol. 58 (8), 1246–1266.
Meerschman, E., Pirot, G., Mariethoz, G., Straubhaar, J., Van Meirvenne, M., Renard, P., 2013. A practical guide to performing multiple-point statistical simulations with the Direct sampling algorithm. Comput. Geosci. 52, 307–324.
Melnikova, Y., Zunino, A., Lange, K., Cordua, K.S., Mosegaard, K., 2015. History matching through a smooth formulation of multiple-point statistics. Math. Geosci. 47, 397–416.
Mosegaard, K., Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems. J. Geophys. Res.: Solid Earth 100 (B7), 12431–12447.
Müller, S., Schüler, L., Zech, A., Heß e, F., 2022. GSTools v1. 3: a toolbox for geostatistical modelling in python. Geosci. Model Dev. 15 (7), 3161–3182.
Nowak, W., Tenkleve, S., Cirpka, O.A., 2003. Efficient computation of linearized cross-covariance and auto-covariance matrices of interdependent quantities. Math. Geol. 35, 53–66.
Oliver, M.A., Webster, R., 1990. Kriging: a method of interpolation for geographical information systems. Int. J. Geogr. Inf. Syst. 4 (3), 313–332.
Prince, S.J., 2012. Computer Vision: Models, Learning, and Inference. Cambridge University Press.
Ravalec, M.L., Noetinger, B., Hu, L.Y., 2000. The FFT moving average (FFT-MA) generator: An efficient numerical method for generating and conditioning Gaussian simulations. Math. Geol. 32 (6), 701–723.
Revil, A., Jardani, A., 2013. The Self-Potential Method: Theory and Applications in Environmental Geosciences. Cambridge University Press.
Rosas-Carbajal, M., Jourde, K., Marteau, J., Deroussi, S., Komorowski, J.-C., Gibert, D., 2017. Three-dimensional density structure of La Soufrière de Guadeloupe lava dome from simultaneous muon radiographies and gravity data. Geophys. Res. Lett. 44 (13), 6743–6751.
Rücker, C., Günther, T., Wagner, F.M., 2017. pyGIMLi: An open-source library for modelling and inversion in geophysics. Comput. Geosci. 109, 106–123.
Ruggeri, P., Irving, J., Holliger, K., 2015. Systematic evaluation of sequential geostatistical resampling within MCMC for posterior sampling of near-surface geophysical inverse problems. Geophys. J. Int. 202 (2), 961–975.
Straubhaar, J., Renard, P., 2021. Conditioning multiple-point statistics simulation to inequality data. Earth Space Sci. 8 (5), e2020EA001515.
Straubhaar, J., Renard, P., Mariethoz, G., 2016. Conditioning multiple-point statistics simulations to block data. Spat. Statist. 16, 53–71.
Straubhaar, J., Renard, P., Mariethoz, G., Froidevaux, R., Besson, O., 2011. An improved parallel multiple-point algorithm using a list approach. Math. Geosci. 43, 305–328.
Strebelle, S., 2002. Conditional simulation of complex geological structures using multiple-point statistics. Math. Geol. 34, 1–21.
Ter Braak, C., Diks, C., 2009. Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. Int. J. Nonlinear Sci. Numer. Simul. 10 (3), 273–290.
Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. 13 (4), 600–612.
Zahner, T., Lochbühler, T., Mariethoz, G., Linde, N., 2016. Image synthesis with graph cuts: a fast model proposal mechanism in probabilistic inversion. Geophys. J. Int. 204 (2), 1179–1190.
Zakeri, F., Mariethoz, G., 2021. A review of geostatistical simulation models applied to satellite remote sensing: Methods and applications. Remote Sens. Environ. 259, 112381.
Zhang, T., 2006. Filter-Based Training Pattern Classification for Spatial Pattern Simulation. Stanford University.
Zhang, T., 2015. MPS-driven digital rock modeling and upscaling. Math. Geosci. 47 (8), 937–954.
Zhang, T., Bombarde, S., Strebelle, S., Oatney, E., 2006. 3D porosity modeling of a carbonate reservoir using continuous multiple-point statistics simulation. SPE J. 11 (03), 375–379.
Zinn, B., Harvey, C.F., 2003. When good statistical models of aquifer heterogeneity go bad: A comparison of flow, dispersion, and mass transfer in connected and multivariate Gaussian hydraulic conductivity fields. Water Resour. Res. 39 (3), 1051.