



OPEN Geological model calibration based on gradual deformation and connectivity function

Junhao Jin¹, Shaohua Li¹, Jun Li² & Fang Ding³

The connectivity is an important feature of the reservoir geological heterogeneity that effects fluid flow responses. In geostatistical modeling, random realizations are generated to describe reservoir heterogeneities. But these realizations do not necessarily honor connectivity data. Especially in multiple-point geostatistical modeling, ensuring the connectivity data of these realizations consistent with the training image is a challenge. The connectivity data of the training image is an important factor in assessing the quality of simulation results in modeling a fluvial hydrocarbon reservoir with meandering channels. The calibration of geological/geostatistical model realizations by measured data is generally performed through history matching, which is an inversion process. This method requires a parameterization of the geostatistical model to allow the updating of an initial model realization. The gradual deformation method has been used to parameterize geostatistical realizations. This method uses a perturbation mechanism to smoothly modify model realizations generated by sequential (not necessarily Gaussian) simulations while preserving its spatial variability. In this paper, a workflow is presented to calibrate the model realizations to the connectivity data. This workflow ensures that the connectivity of model realizations generated by SNESIM is consistent with the training image, and it is applicable to both multiple-point and two-point geostatistical modeling methods based on sequential simulation. This workflow incorporates the computation of the model connectivity function and its calibration to connectivity data using the gradual deformation method. This involves defining and minimizing an objective function, which quantifies the mismatch between the model connectivity function and the connectivity data. In the case study, multiple initial realizations are utilized to construct a final realization of the reservoir model that honors the connectivity data.

Keywords Connectivity function, Optimization, Gradual deformation, SNESIM

Connectivity is one of the fundamental properties of reservoir models, directly impacting the hydrocarbon recovery^{1,3,16}. Analyzing the connectivity can draw attention to geological uncertainties that may affect fluid flow response¹¹. In many studies, sand connectivity is a key issue. For example, in calculating the connected hydrocarbon volume, the connected sandbodies intersected by wells must be considered, then the length of the connected path from each cell to the nearest production well must be calculated²¹. The word connectivity is often used as a broad concept with various definitions. Renard et al.²⁸ provides a comprehensive overview of connectivity metrics, distinguishing between static and dynamic metrics, and global versus localized metrics. The connectivity function is an effective tool to analyze the connectivity of reservoir models²⁵. Pirot et al.²⁶ introduces a Python package called loopUI-0.1 to compute the connectivity function. Pardo-Igúzquiza²⁵ provides a computer program called CONNEC3D for connectivity analysis, which consists of computing connectivity function for different directions and a number of connectivity statistics. In this paper, CONNEC3D is used to compute the connectivity function of the geological model.

In geostatistical modeling, multiple model realizations are generated to describe the reservoir heterogeneities. However, these realizations may not honor the connectivity data. In multiple-point geostatistical modeling, the training image is used to infer the relationships between multiple spatial points to facilitate the prediction of reservoirs³³. The training image serves as a proxy for the variogram in traditional geostatistical methods, capturing the spatial relationships and distribution patterns among multiple points within a geological context³². The connectivity of the training image can assess the quality of simulation results²⁹. If there is a significant discrepancy between the connectivity of the simulation results and that of the training image, then the simulation

¹School of Geoscience, Yangtze University, Wuhan 430100, China. ²Exploration and Development Research Institute, PetroChina Tarim Oilfield Company, 841000 Korla, China. ³Research Institute of CNOOC Shanghai Branch, Shanghai 200030, China. ✉email: lish@yangtzeu.edu.cn

results are considered to be poor. It's important to ensure the connectivity data of random realizations consistent with the training image. To calibrate a geostatistical realization to connectivity data, we can apply a history matching method that involves parameterizing the geostatistical model, then minimizing the mismatch between the connectivity data and the connectivity function of the model realization.

Various methods have been introduced for parameterizing geostatistical models. Initially, Oliver et al.²⁴ proposed the Markov chain Monte Carlo method which involves building model realizations conditioned to all observations. But this method involves using a reliable but computation intensive sampling method. His work emphasizes the importance of choosing an appropriate prior model and the impact it can have on the ability to assimilate data. Subsequently, Marsily et al.²² introduced the pilot point method which employs a limited number of parameters to efficiently calibrate model realizations to the observed data. This method combines a gradient search algorithm, ensuring efficient inversion while preserving the spatial variability. Following that, the gradual deformation method was developed by Hu and Blanc¹³. This approach involves iteratively optimizing combinations of independent realizations of a stochastic model until honoring the observed data. Unlike the pilot point method applicable to continuous models, this approach can be applied to any type of stochastic models^{12,15}. Moreover, the probability perturbation method was proposed by Hoffman and Caers¹⁰. It explicitly deals with the problem of combining prior probabilities with pre-posterior probabilities derived from the data. Ding et al.⁷ introduced the domain deformation method, which allows local modifications to model realizations by altering the geometric domain. Its geometric shape and size can be easily controlled and parameterized. All these methods allow to modify the model realizations while preserving spatial variability.

In recent years, the gradual deformation method has been increasingly used. This approach utilizes the perturbation mechanism to search for solutions matching nonlinear data in the model space. The perturbation mechanism ensures that the spatial variability of the model is preserved⁵. Commencing with an initial reservoir model that does not match the production history, the method allows for the gradual deformation of the initial model while preserving its geological continuity, until a history matching is achieved³⁰. As in the Gaussian case, the gradual deformation of sequential simulation preserves the spatial variability of the stochastic model¹². This method not only applies to Gaussian-related models but also to models generated by sequential (not necessarily Gaussian) simulation¹⁵. Hu and Le Ravalec-Dupin¹⁴ proposed a method to improve the convergence speed of the gradual deformation using gradient information. Ding⁶ proposed an approximate derivative computation to improve the computational efficiency of gradual deformation using gradient information. Le Ravalec-Dupin and Hu²⁰ combined the pilot point method with the gradual deformation method to enhance the flexibility and specificity of model deformation. Marteau et al.²³ combined local gradual deformation and domain deformation methods, dynamically modifying region partitioning during the iterative optimization process, reducing the need for interdependent regions, and ultimately achieving better results than the original gradual deformation method. Gradual deformation method has been successfully applied to several synthetic and real datasets for history matching, characterizing a reservoir by honoring static (e.g., porosity, permeability) and dynamic data (e.g. pressure, oil rate) from wells^{8,17–19,27}. Additionally, the gradual deformation method has been applied to the stochastic seismic inversion^{35–37}.

In this work, we propose a workflow involving gradual deformation and the connectivity function. The connectivity function is utilized to define an objective function to be minimized. The single normal equation simulation (SNESIM) method is utilized to generate model realizations and multiple realizations are utilized to construct a final realization that honors a connectivity function of training image. Renard et al.²⁹ proposed a method to condition stochastic simulations of lithofacies to connectivity data. This method consists of using a training image to build a set of replicates of connected paths that are consistent with the prior model. This method is particularly efficient in computing local connectivity. Compared to his work, the efficiency of the proposed workflow depends greatly on the way the stochastic model is deformed.

This workflow not only enhances the accuracy of reservoir modeling by ensuring that the model realizations honor the connectivity data derived from training images, but also addresses the challenges associated with geological uncertainties that can impact fluid flow. In the case study, multiple model realizations honor the connectivity data of training image are generated by this workflow. By using the multiple dimensional scaling (MDS) method on these realizations, we can select the model that is more similar to the training image, thus achieving the purpose of uncertainty analysis. This workflow is not only suitable for SNESIM, but also for other sequential simulation methods.

The paper is structured as follows: Sect. 1 provides a review of connectivity analysis, the revised version of SNESIM and the gradual deformation method. Section 2 introduces the workflow of the gradual deformation method, integrating it with SNESIM and connectivity function. In Sect. 3, the computational experiments of the workflow are presented. Finally, Sect. 4 presents the conclusions of this work and our future research.

Methodology

Connectivity analysis

In this work, a computer program called CONNEC3D²⁵ is utilized for connectivity analysis of 3D stochastic models. Given an indicator map on a regular 2D or 3D grid, CONNEC3D performs a connectivity analysis of a given phase, a lithofacies or a category, of interest. This phase and the complementary phase can be represented by an indicator valued 0 or 1. Such a dichotomous indicator map can be derived from continuous variables or categorical variables. In the case of continuous variables (e.g., permeability), this program employs one or more thresholds to define different categories for connectivity analysis. For multiple phases, this program will pre-generate a grid of the phase of interest with its indicator assigned 1, while the indicators of all other phases assigned 0. Connectivity analysis involves estimating the connectivity function $\tau(h)$ for different spatial directions. The connectivity function, which is defined as a function of distance h , provides the probability of any two cells separated by h belonging to a given phase are connected.

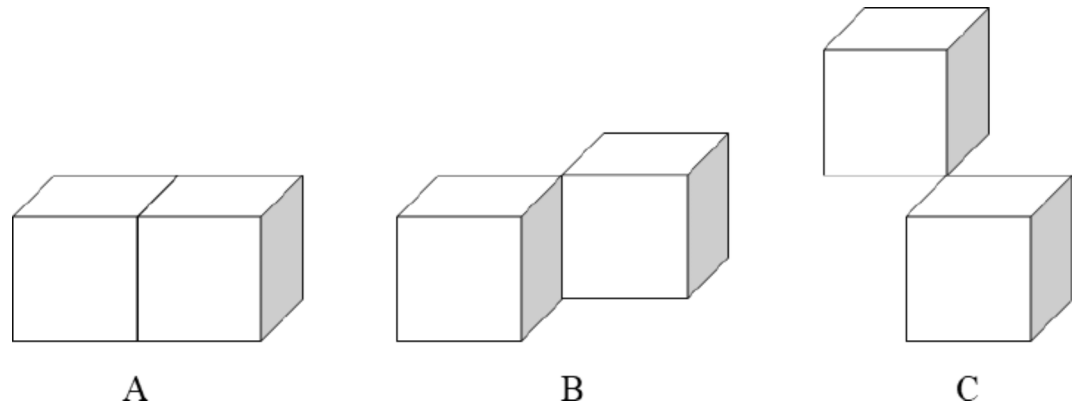


Fig. 1. **A,** Face connectivity. **B,** Edge connectivity. **C,** Vertex connectivity.

6-connectivity analysis	18-connectivity analysis	26-connectivity analysis
 6 faces	 6 faces + 12 edges	 6 faces + 12 edges + 8 vertices

Fig. 2. Six neighbors of 6-connectivity analysis (red), 18 neighbors of 18-connectivity analysis (red plus orange) and 26 neighbors of 26-connectivity analysis (red plus orange plus blue).

Figure 1 shows different ways that two cells are connected in 3D. The program offers three connectivity analysis options: (1) 6-connectivity, where two cells are connected if they have a shared face; (2) 18-connectivity, where two cells are connected if they have a shared face or a shared edge; and (3) 26-connectivity, where two cells are connected if they have a shared face, a shared edge, or a shared vertex. For each cell, there are six neighbors defining 6-connectivity analysis, eighteen neighbors defining 18-connectivity analysis, or twenty-six neighbors defining 26-connectivity analysis, as illustrated in Fig. 2.

Given an indicator grid M , each point or cell on M belongs to a random subset S or to its complementary set S^c with $M = S \cup S^c$. Each cell x of M is coded by an indicator function $I(x)$ taking value of 1 and 0:

$$I(x) = \begin{cases} 1, & \text{if } x \in S \\ 0, & \text{if } x \in S^c \end{cases}$$

where x is an arbitrary location, point or cell of M . Two cells separated by a fixed distance (lag) h are denoted as x and $x + h$. We introduce a notation ' \iff ' to indicate that two cells are connected; for example, $x \iff x + h$ signifies that these two cells are connected. The connectivity function $\tau(h)$ is defined as the probability that two cells belonging to phase S are connected:

$$\tau(h) = P(x \iff x + h | x, x + h \in S)$$

The connectivity function for distance h is estimated as follows:

$$\hat{\tau}(h) = \frac{\#N(x \iff x + h | x, x + h \in S)}{\#N(x, x + h \in S)} \tag{1}$$

where $\#N(x, x + h \in S)$ for the number of pairs of cells separated by distance h and belonging to phase S , $\#N(x \iff x + h | x, x + h \in S)$ for the number of connected pairs of cells separated by distance h and belonging to phase S .

Figure 3 presents an example of calculating the connectivity function for distance 3 along X direction. Each cell belongs to a subset S (sand) or its complementary set S^c (shale). There are three possibilities for x and $x + h$ ($h = 3$): ① If one or both cells belong to the shale, such as x_1 and $x_1 + 3$ or x_5 and $x_5 + 3$, this pair does not contribute either to $\#N(x, x + h \in S)$ or $\#N(x \iff x + h | x, x + h \in S)$. ② If both cells belong to sand but not to the same connected component, such as x_2 and $x_2 + 3$, then the value of $\#N(x, x + h \in S)$ is increased by 1. ③ If both cells belong to the sand and are part of the same connected component, such as x_3 and $x_3 + 3$ or x_4 and $x_4 + 3$, then both $\#N(x \iff x + h | x, x + h \in S)$ and $\#N(x, x + h \in S)$ values are increased by 1.

The steps for estimating the connectivity function using 6-connectivity analysis in Fig. 3 are as follows. First, count the number of pairs of cells that, for the given distance, belong to the phase of interest. There are 3 such pairs in this example. Second, count the number of pairs of cells that are connected for the given distance. There are 2 connected pairs according to the 6-connectivity analysis. Finally, calculate their ratio. The result in this example is two-thirds.

The connectivity function depends on the discretization scale: what is connected at one scale may be disconnected at finer scales. If grid M is discretized into a finer grid, the result of the connectivity function using 6-connectivity analysis in Fig. 3 may change accordingly. However, for two cells to be connected, as long as both cells belong to the same geological body at finer scales, they are connected.

The model in Fig. 4A includes two phase types named fluvial and mudstone, and it is discretized on a regular grid of $100 \times 100 \times 50$ cells of size $1 \times 1 \times 1$ m. Figure 4B is an example of using CONNEC3D to calculate connectivity functions along different directions represented in blue, green and red curves. This model shows a good connectivity of fluvial along the Y direction. It is reasonable to have good connectivity along the river alignment (Fig. 5). Curves in Fig. 1 Fig. 4B are comparable, when the number of lags (distance) increases, the red curve decreases slower than the blue curve. This means that fluvial is connected better in the X direction compared to the Z direction. When the number of lags increases, the number of connected pairs of cells decreases, so the connectivity functions of fluvial along Y and Z directions show a decreasing trend in general. For a lag of 50 units, the connectivity function along Z direction is 0 (Fig. 4B). Because this model only contains 50 cells in Z direction. While in the case of a lag of 1 unit, the connectivity function is evidently 1.0, as two units that are adjacent (i.e., separated by a lag of 1 unit) and belong to the same phase are necessarily connected. A phase with a small proportion will provide a curve suggesting lower connectivity (*ceteris paribus*) in comparison with a phase with a higher proportion. The model shows that the mudstone is generally well connected compared to the river.

The nomenclature used in this paper follows that of Allard et al.¹ and Allard²: a regular grid of $NX \times NY \times NZ$ of size $DX \times DY \times DZ$, where NX is the number of grid cells in the X direction, NY is the number of grid cells in the Y direction and NZ is the number of cells in the Z direction, $DX \times DY \times DZ$ is the size of each grid cell. Figure

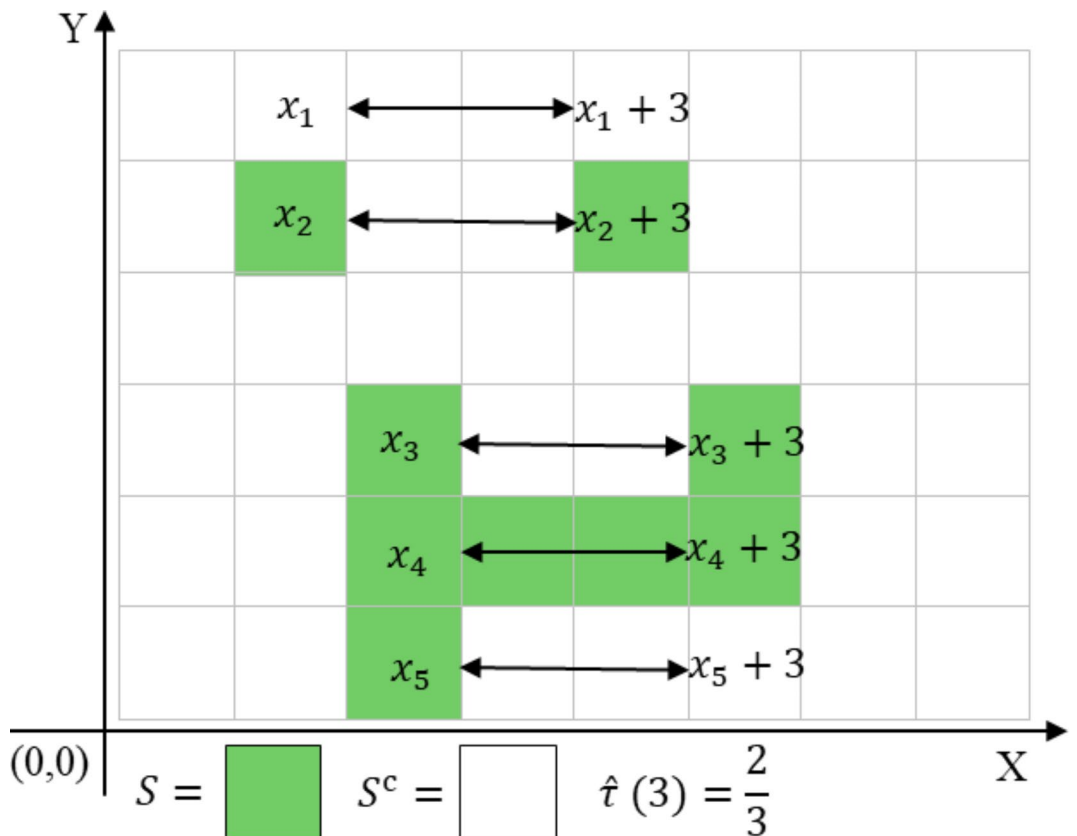


Fig. 3. An example of calculating connectivity function for distance 3 along X direction.

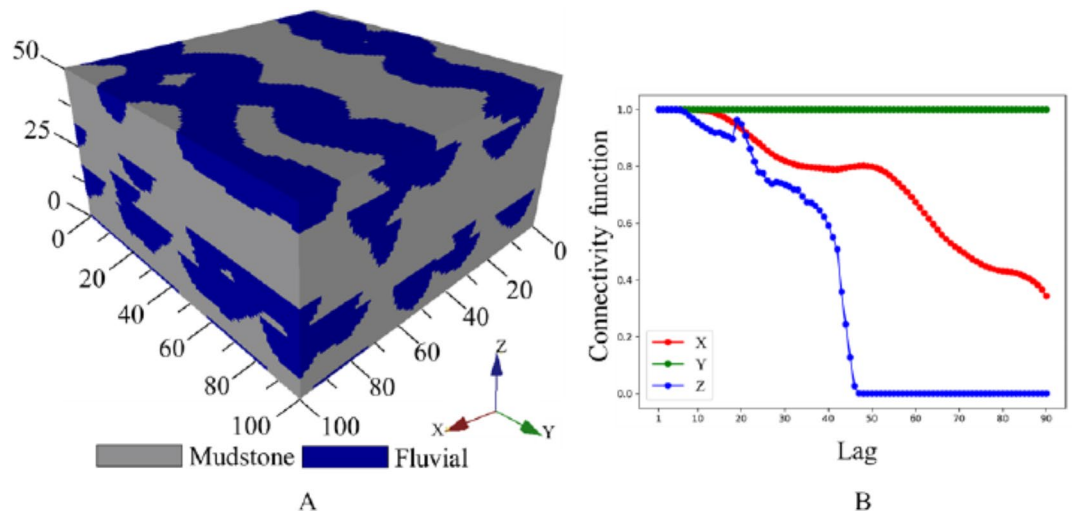


Fig. 4. A, A 3D stationary categorical fluvial reservoir model. B, Connectivity functions of fluvial along X, Y and Z direction for 3D stationary fluvial reservoir model.

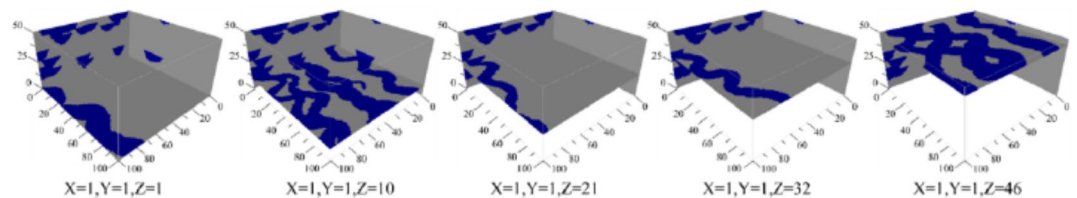


Fig. 5. The 3D stationary fluvial reservoir model with different locations along Z directions.

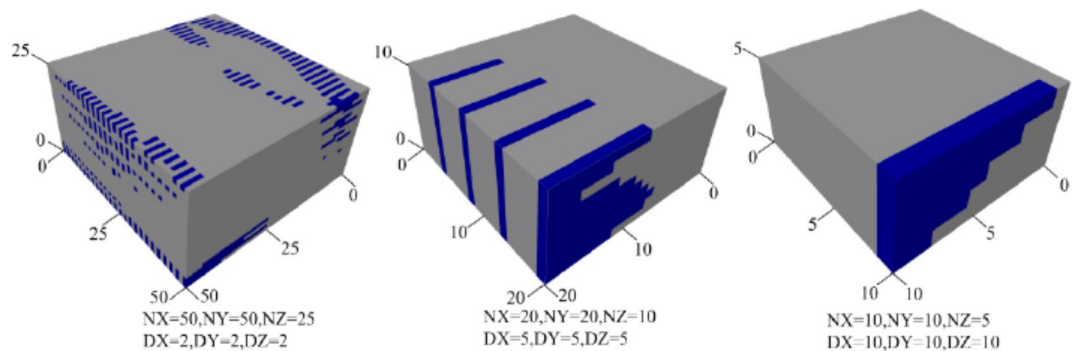


Fig. 6. The 3D stationary fluvial reservoir model with different discretization scales.

6 illustrates that the model shown in Fig. 4A has been discretized into three scales. A smaller scale means that the grid cells in the model are finer, thus capturing geological features and variations with greater precision. Conversely, a larger scale implies larger grid cells, which may prevent the model from capturing small-scale geological details. Therefore, as the scale changes, the details of the model become coarser and the result of the connectivity function of the model in Fig. 4A will change accordingly.

Figure 7 shows a realization of 3 facies model in volume and in slices from three different orientations. The model is discretized on a regular grid of $100 \times 60 \times 20$ cells of size $1 \times 1 \times 1$ m. The local conditional probability distribution (LCPD) for each grid cell is estimated using the exponential variogram model. The main azimuth of variogram function is 10° , with dip and plunge angles of 0° . The ranges of the variogram function are 60, 50, and 2 m, respectively. The proportions of facies 0, 1, and 2 are respectively 30, 50 and 20%. Apparently, this model shows a good connectivity of facies 1 (green).

In Fig. 8, the connectivity function along the X, Y and Z directions is calculated using the reference model. Subsequently, the average connectivity function is obtained by averaging the connectivity functions of the X, Y

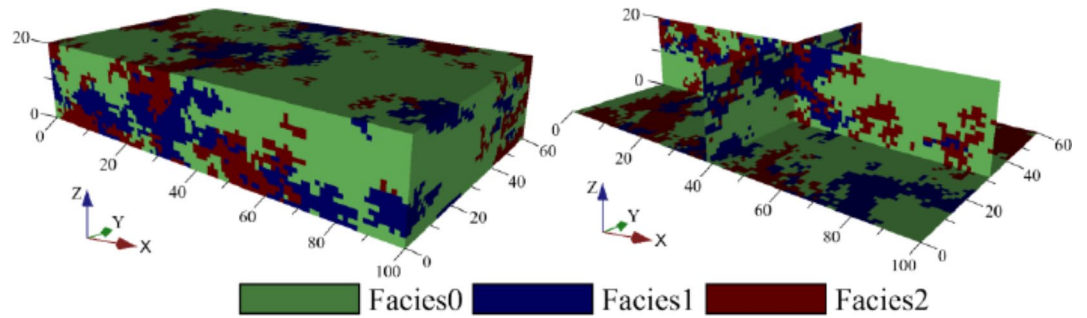


Fig. 7. From left to right: A reference model generated by sequential indicator simulation and its slices (Slices are respectively positioned at X=50, Y=30 and Z=0).

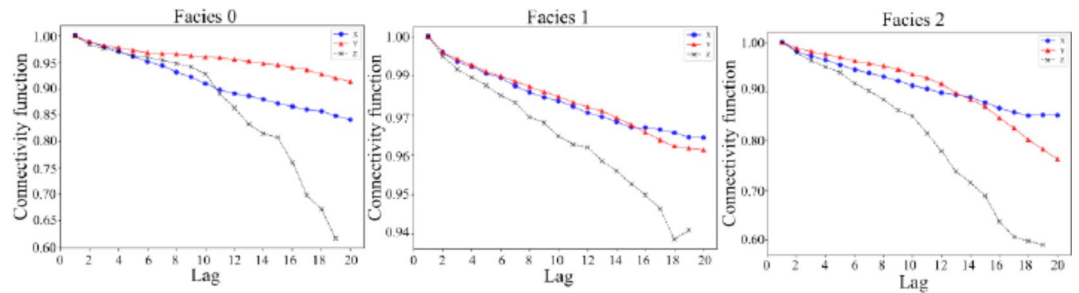


Fig. 8. Connectivity functions of the reference model along different directions.

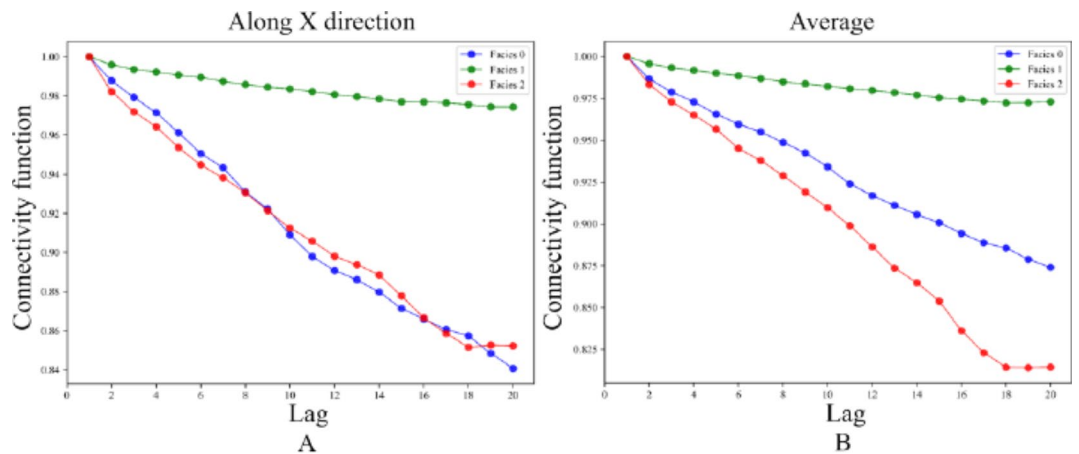


Fig. 9. A, Connectivity functions of different facies along X direction of the reference model. B, Average connectivity functions of Fig. 8. Connectivity functions of the reference model along different directions.

and Z directions (Fig. 9B). Figure 9A is an example of using CONNEC3D to calculate connectivity functions of three facies represented in blue, green and red curves. Comparing these three curves, when the number of lags (distance) increases, the green curve decreases slower than the red and blue curves. It means facies 1 is connected better than the two others. When the number of lags increases, the number of connected pairs of cells decreases, so the connectivity functions of these three facies show a decreasing trend in general. Because the reference model only contains 20 cells in Z direction (Fig. 7), the connectivity function along Z direction is 0 for a lag of 20 units (Fig. 8). And the connectivity function is also 1.0 in the case of a lag of 1 unit.

The SNESIM and its revised version

The SNESIM is the earliest probabilistic sequential simulation algorithm³¹. For each cell, it computes the local conditional probability distribution (LCPD) along a random path based on the likelihood of potential events and their frequencies observed in training images. Let $Z = (Z_1, Z_2, \dots, Z_N)$ stand for the random vector of interest. The distribution of Z has not to be multivariate Gaussian or associated with a stationary random

function. In practice, the N elements of Z will be arranged according to a random path defined in advance and the value of them will be calculated one after another from Z_1 to Z_N . For each element Z_i ($i = 1, 2, 3, \dots, N$), the value will be calculated according to the following steps:

1. Build the distribution function F of Z_i conditioned to $(Z_1, Z_2, \dots, Z_{i-1})$:

$$F(z_i) = P(Z_i < z_i | Z_1, Z_2, \dots, Z_{i-1}) \quad (2)$$

2. Generate a random number u_i from a uniform distribution within the interval $[0, 1]$.
3. Based on the value of u_i , draw a value for Z_i from the distribution $F(z_i)$.

Different types of sequential simulation define different distribution functions $F(z_i)$ ($i = 1, 2, 3, \dots, N$) to generate realizations of multi-Gaussian vectors or non-Gaussian indicator vectors. Hansen et al.⁹ provide the computer codes of multiple-point statistics for estimating these distribution functions. There exist many algorithms for drawing a value of Z_i from the distribution $F(z_i)$. We consider the inverse distribution method by which a realization of Z_i is obtained by $z_i = F^{-1}(u_i)$ where u_i is obtained from the uniform distribution between 0 and 1. Thus, the random vector Z corresponds to a random vector U whose elements U_1, U_2, \dots, U_N are mutually independent and follow the uniform distribution between 0 and 1. From the computing point of view, any sequential simulation procedure can be regarded as an operation which transforms the random vector U (generated by a pseudo random number generator) to the random vector Z that are distributed according to the underlying stochastic model.

The gradual deformation of sequential simulation achieves gradual deformation by perturbing the realization of U . This paper focuses on the challenge of calibrating the vector Z to connectivity data by consistent fitting of realizations of the uniform vector $U = (U_1, U_2, \dots, U_N)$ using gradual deformation of sequential simulation.

Due to the inability to directly apply the original SNESIM to gradual deformation of sequential simulation method, certain modifications to the algorithm are necessary. Algorithm is modified as follows:

- Fix the random seed used for generating the random path.
- Improve step 3 of the algorithm by adding the ability to record and fix the random number.

Gradual deformation of sequential simulation

The gradual deformation method is a geostatistical parametrization method for solving inverse problems under the given model constraint. It is based on the fact that any linear combination of Gaussian realizations remains Gaussian. By optimizing the combination coefficients, one can calibrate realizations of stochastic models to nonlinear data such as connectivity function. Gradual deformation of sequential simulation (GDSS) is a variant of the gradual deformation method, both of which utilize perturbation mechanisms for deformation. In the GDSS, to gradually change a realization z of Z while preserving the spatial variability, a realization u of the uniform vector U is perturbed gradually¹⁵.

The random path needs to be fixed for all realizations in GDSS, otherwise the perturbations would not be gradual. Generating all realizations using one same random path does not affect the reproduction of the geological structure. Note also that within the sequential drawings from the conditional distribution functions, the uniform random numbers are independent from each other, and the intended geological structure is reproduced⁴.

The core idea of the gradual deformation method is that a new Gaussian vector can be generated by linearly combining two independent Gaussian vectors. Consider two independent Gaussian vectors y_1 and y_2 , a new Gaussian vector $y(r)$ is defined as a function of parameter r :

$$y(r) = y_1 \cdot \cos(r) + y_2 \cdot \sin(r) \quad r \in [-\pi, \pi] \quad (3)$$

From $y(r)$, the vector $u(r)$ whose components are uniformly distributed is defined as follows:

$$u(r) = G(y(r)) = G(y_1 \cdot \cos(r) + y_2 \cdot \sin(r)) \quad (4)$$

where G denotes the standard Gaussian cumulative distribution function. Therefore, when $u(r)$ are utilized in sequential simulation instead of u_1 , one can generate a series of realizations $z^{(r)}$ which constitute a gradual deformation of the initial realization $z = z^{(0)}$. As $u(r)$ can be transformed to $y(r)$, therefore, the gradual deformation method of Gaussian models can be applied to any type of stochastic models⁴. Sequential indicator simulation has been successfully applied to GDSS¹⁵. Figure 10 is an example of 2D grid containing 100 cells in X direction and 60 cells in Y direction, and it illustrates the evolution of the model under the gradual deformation method with different deformation parameters. It is set on a 100×60 grid, with the remaining setting following the reference model in Fig. 7. The parameter r is discretized into ten values equally spaced between $[0, \pi/20]$. For each value r , a new gradually deformed realization is generated, resulting in a total of ten realizations. With r increasing, the initial realization $z^{(r=0)}$ is gradually deformed to the realization $z^{(r=\pi/20)}$ corresponding to $r = \pi/20$. We observe that the model deviates substantially from the initial model when $r > \pi/36$. Because of the nature of the sequential simulation algorithm, realization of a categorical variable cannot be perturbed in a gradual way. Figure 11A shows the random numbers used for modeling $z^{(r=0)}$, which represents the initial state without perturbation. Figure 11B shows the distribution of these random numbers. By applying the gradual deformation method with a deformation parameter of $r = \pi/20$, we obtain the perturbed random numbers used for modeling $z^{(r=\pi/20)}$, as shown in Fig. 11C. Figure 11D then shows the distribution of these perturbed random numbers, highlighting the changes in distribution due to the deformation. In the context of

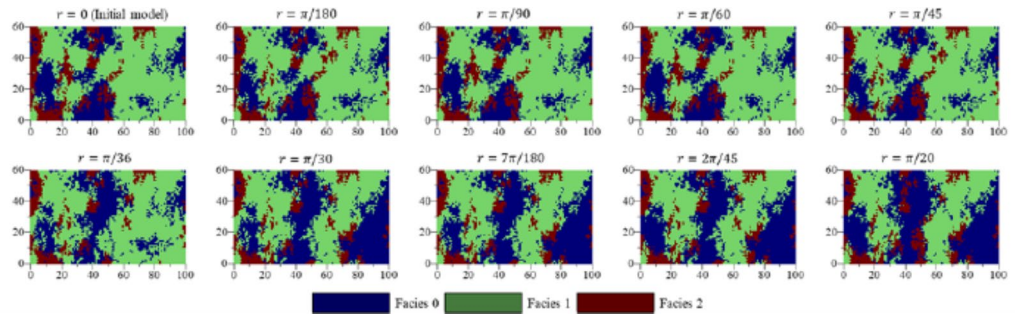


Fig. 10. Ten realizations of the 2D model result from Eq. 4 (with different values of the deformation parameter r).

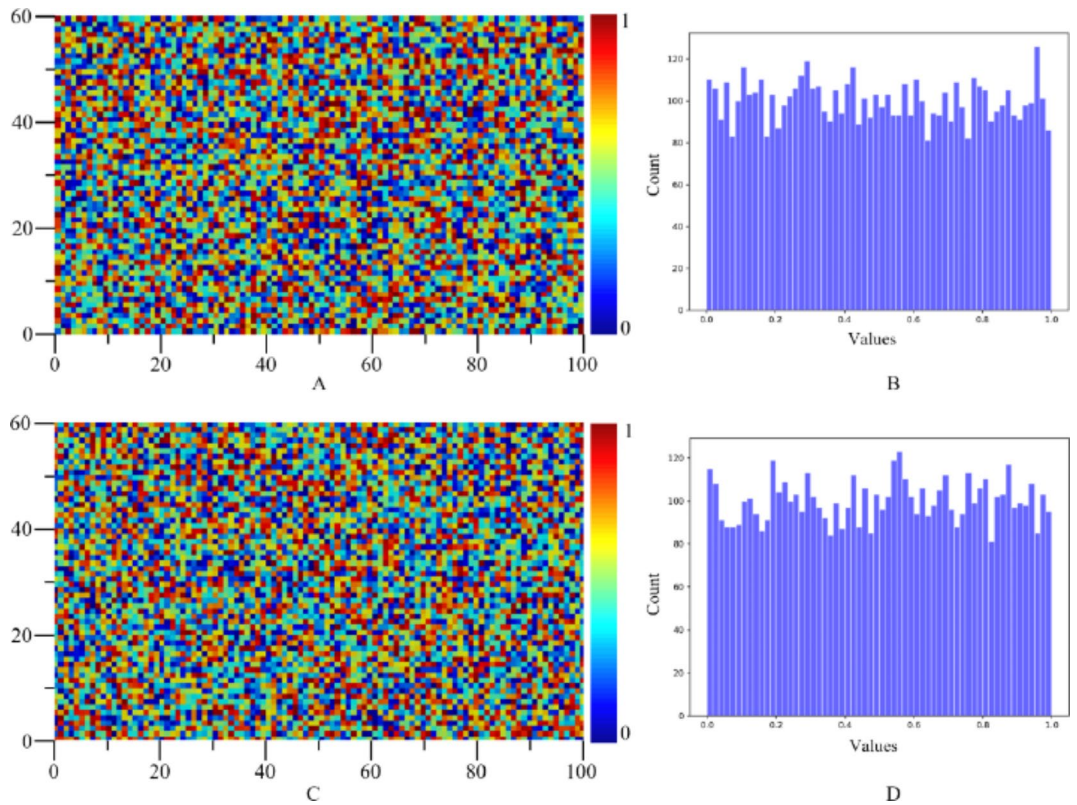


Fig. 11. A, The random numbers used for modeling the initial model realization in Fig. 10. B, The distribution of Fig. 11A. C, The perturbed random numbers used for modeling $z^{(r=\pi/20)}$. D, The distribution of Fig. 11C.

this algorithm, a small change in the random numbers utilized to simulate the initial values along the random path could cascade into large changes along the way⁵.

Calibrating the realization of a stochastic model to nonlinear data can be formulated as an optimization problem. Consider $f^{(obs)} = (f_1^{obs}, f_2^{obs}, \dots, f_p^{obs})$ is the nonlinear data we want to calibrate and $f^{(r)} = (f_1^r, f_2^r, \dots, f_p^r)$ is the corresponding vector of the responses of the realization $z^{(r)}$. To construct a realization of the stochastic model that best matches the nonlinear data using GDSS, the following objective function needs to be minimized with respect to r :

$$O(r) = \frac{1}{2} \sum_{i=1}^p w_i^r (f_i^r - f_i^{obs})^2 \tag{5}$$

where w_i^r denotes the weight attributed to response f_i^r . This is a one-dimensional optimization problem which can be solved with an efficient optimization algorithm.

The weight w_i^r is designed to reflect the relative importance associated with response f_i^r . If the realization $z^{(r)}$ has been shown to predict certain responses more reliably, those responses can be given higher weights. In some cases, expert knowledge or additional information not captured by the data can be used to assign weights.

For example, if certain responses are known to be more critical for the application at hand, they can be given more weight. The weights should be chosen to ensure that the objective function $O(r)$ accurately represents the goal of calibrating the model realization to the observed data.

In GDSS, the deformed model realization preserves its original spatial variability (reproduction of variogram) and yields in general a regular objective function $O(r)$ is amenable to minimization through an efficient optimization algorithm (e.g., a golden section search method)¹⁵.

Proposed workflow

In multiple-point geostatistical modeling, the training image is derived from a comprehensive integration of geological knowledge, outcrop analogs, well logs, seismic data, and any other available geological and geophysical information. Once the training image is derived, it serves as a reference model for generating stochastic models and guiding the simulation process to ensure that the results reflect the complex spatial relationships and geological features. The quality of simulation results greatly depends on whether their connectivity honors the training image. This workflow focuses on controlling the connectivity of stochastic model generated by sequential simulation method.

Given an initial stochastic model generated with any sequential simulation method, the CONNEC3D involves the estimation of connectivity functions for different spatial directions. In this experiment, the GDSS is proposed for finding a realization that honors given connectivity functions. The objective function $O(r)$ is defined as follows:

$$O(r) = \sum_{lag=1}^n |CF_{lag}^{target} - CF_{lag}^r| \quad (6)$$

The target connectivity function CF_{lag}^{target} is derived from the training image and CF_{lag}^r is the corresponding connectivity function derived from the proposed realization $z^{(r)}$. The objective function needs to be minimized during each iteration until finding a realization $z^{(r_{opt})}$. The complete workflow in Fig. 12 is utilized in this experiment.

This workflow is not only suitable for SNESIM, but also for other sequential simulation methods. To show its adaptability to a variety of sequential modeling methods, we applied this workflow to sequential indicator simulation (SISIM). As we integrate SISIM into workflow, similar modifications to those made for SNESIM are

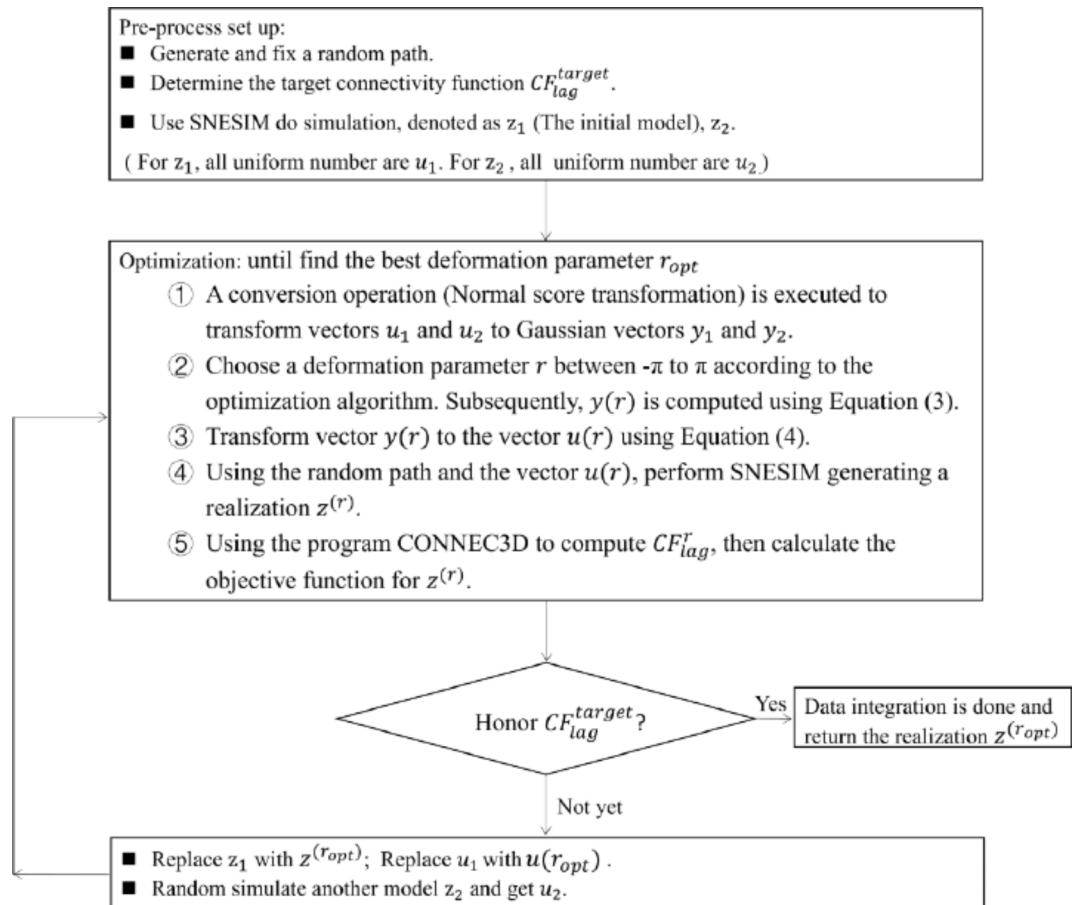


Fig. 12. Visualization of the workflow.

required. This is consistent with the approach discussed in Sect. 1.2, as both SISIM and SNESIM are based on sequential simulation. Some certain modifications to the workflow are necessary:

- The target connectivity function CF_{lag}^{target} is derived from a more reliable model that reflects the desired connectivity characteristics, as there is no TI.
- The SISIM is used to do simulation instead of SNESIM.

Results

To illustrate the application of the proposed workflow, we considered the calibration of 2 phases models to the connectivity data. This case study is similar to¹⁵, except that we utilize the connectivity data instead of the well-test pressure data.

SNESIM

Fluvial reservoirs are characterized by the presence of sinuous sand-filled channels within a background of mudstone. For this example, we consider one single type of sand: channel sand which corresponds to the best reservoir rock. A realistic modeling of the curvilinear sand channel patterns is critical for reliable connectivity assessment and flow simulation of such reservoirs. The training image in Fig. 13A, derived from the real-world data⁹, carries information about the mean width, the major direction, and the sinuosity of the channels, as well as their spatial grouping. This training image, which could have been drawn by a geologist and subsequently digitized, was considered for our model. The TI pattern reflects the natural characteristics of river distribution within geological structures, where rivers typically extend along their main flow direction rather than crossing frequently. The continuity of this river distribution is reflected in the connectivity function, thereby influencing the connectivity features of the entire reservoir model.

We consider the connectivity function of the training image in Fig. 13A as the target connectivity function. Based on this training image, the revised SNESIM algorithm will be used to generate realizations in a 2D grid $100 \times 100 \times 1$ where each cell has a size of $1 \times 1 \times 1$ and the coordinates of the first pixel are (0, 0). The search template size is set as $7 \times 7 \times 1$ and the number of multiple grids is 3. The initial information to construct realizations is derived from Hansen et al.³⁴ Figure 13B shows a realization of 2 phases model which is considered as the initial model. Figure 13C compares the connectivity functions of the training image and the initial model, and shows a big gap to them. The connectivity function of TI behaves in this particular way due to the less crossing channels across the Y direction. Our objective is to modify the initial model realization until honoring the target connectivity function. Figure 14 shows the model realizations resulting from each iteration of the optimization procedure. These realizations can be investigated successively to identify an optimal realization of the stochastic model. Figure 15 shows their corresponding connectivity functions (green) compared to the target (blue). It can be seen that the connectivity function gets closer to the target as the number of iterations increases.

Figure 16A demonstrates that after the first iteration, the realization of the initial model has been modified to be close to the target connectivity function, the objective function decreased by 78% (from 56.75554 to 12.63937), showing the efficiency of the gradual deformation method. A good match is obtained after tenth iterations, the objective function decreased by 97% (from 56.75554 to 1.7121), although the deformed model is still different from the training image. This is due to the connectivity function in this case do not fully condition the model. To better visualize the decrease in the objective function, Fig. 16B shows the evolution of the objective function versus the number of iterations in logarithmic scale.

In Fig. 17, Multi-Dimensional Scaling (MDS) is utilized to map the training image and the realizations from high-dimensional spaces to low-dimensional spaces (typically two or three dimensions). This statistical technique is employed for data visualization while preserving the distances or similarities between the data points as much as possible²⁶. In MDS, the Sliced Wasserstein distance is calculated between the training image and the realizations, then distance discrepancies can be projected and visualized in 2D space. In 2D space, the x-axis is a dimension that captures the primary variations (e.g., attribute information, spatial variability characteristics, channel connectivity and so on) among the data points. The distance between each point reflects

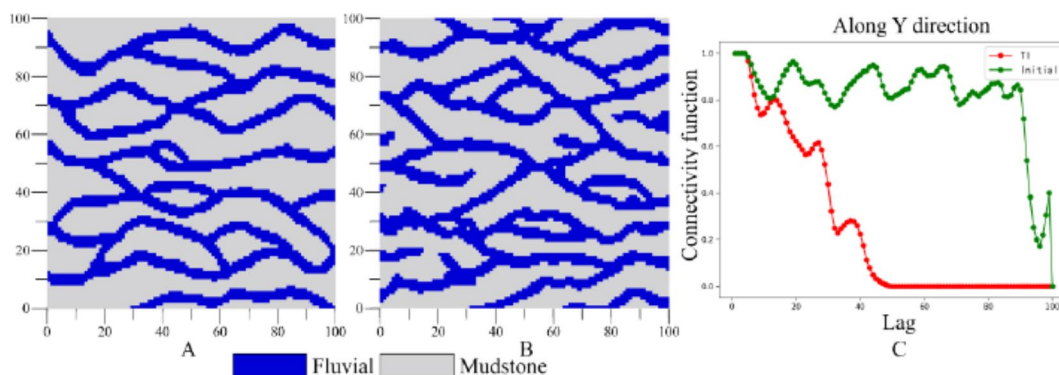


Fig. 13. A, The training image⁹. B, The initial model. C, Gap between the connectivity functions of the training image and the initial model.

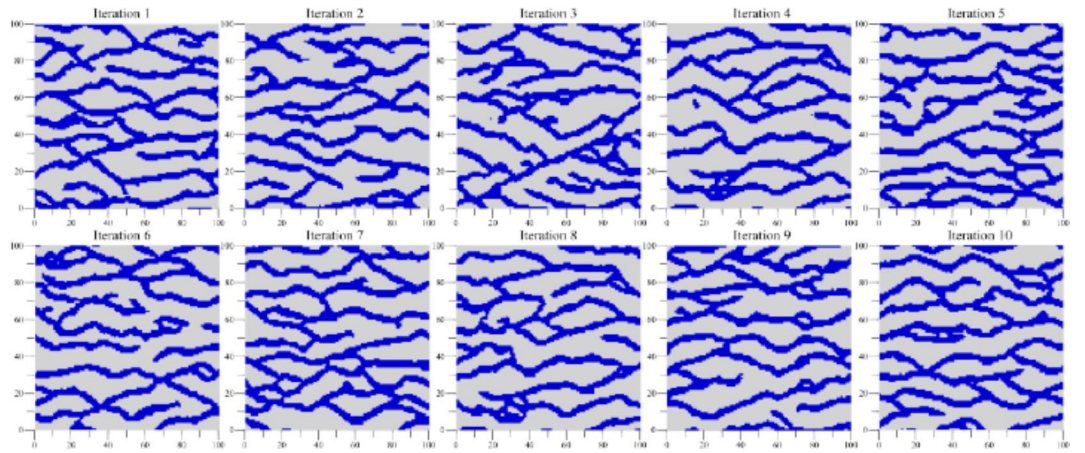


Fig. 14. The model realizations resulting from the optimization procedure.

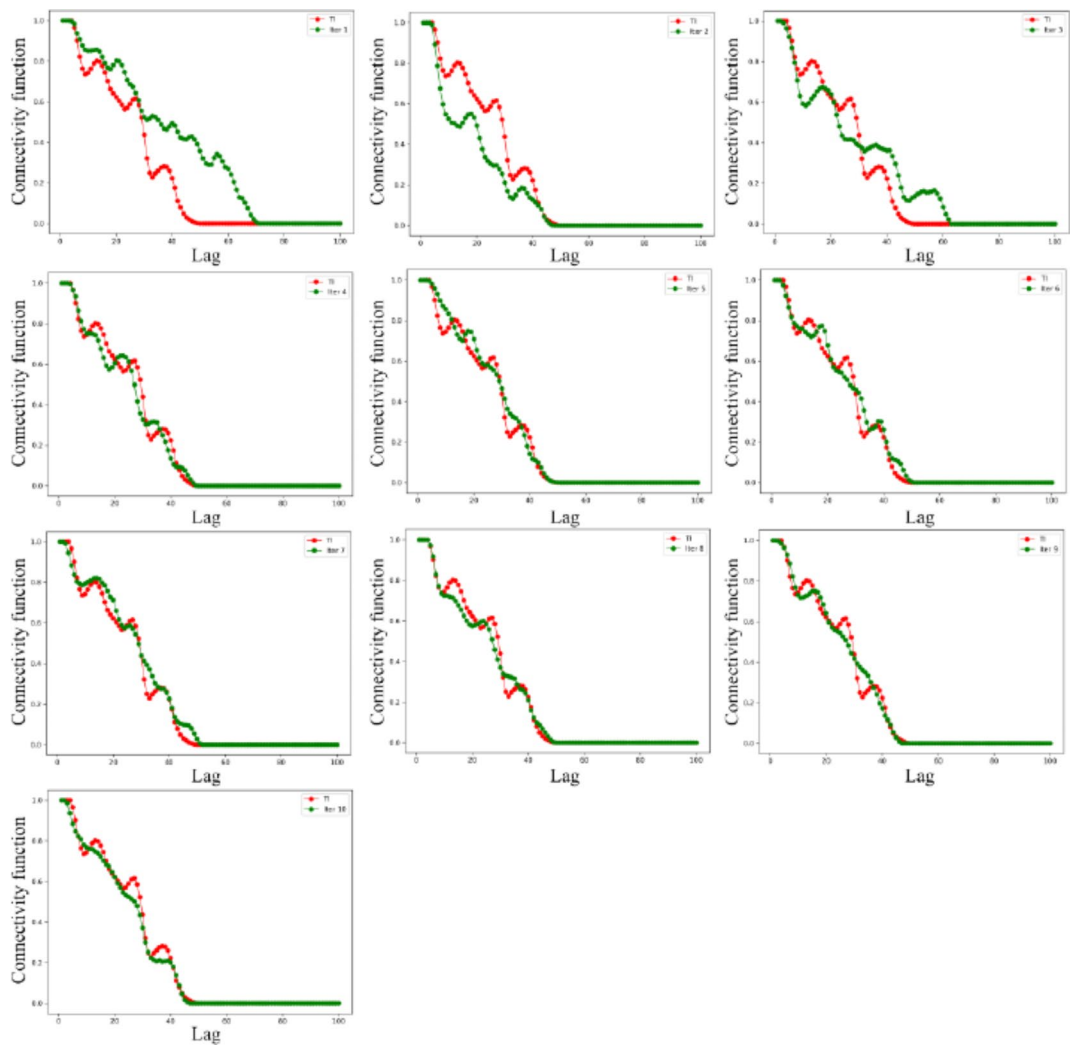


Fig. 15. Comparing the connectivity function after each iteration with the target.

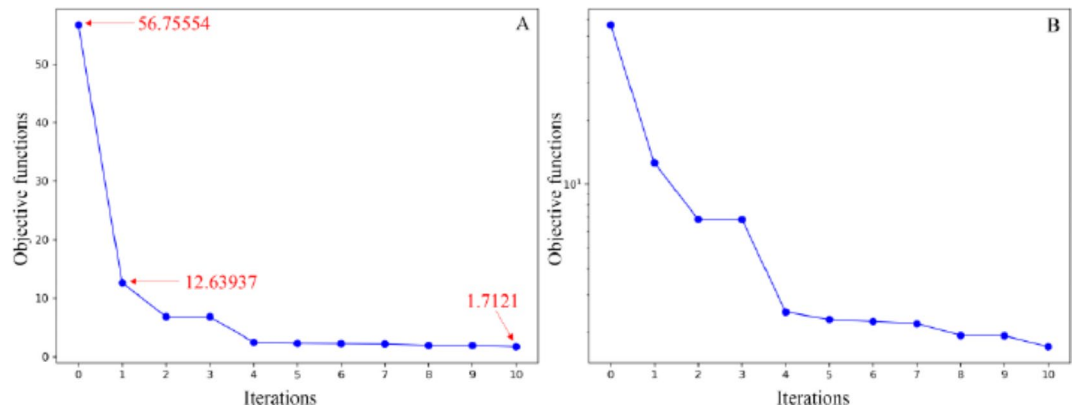


Fig. 16. A, Objective function versus the number of iterations. B, Objective function versus the number of iterations – objective function (ordinate) in logarithmic scale.

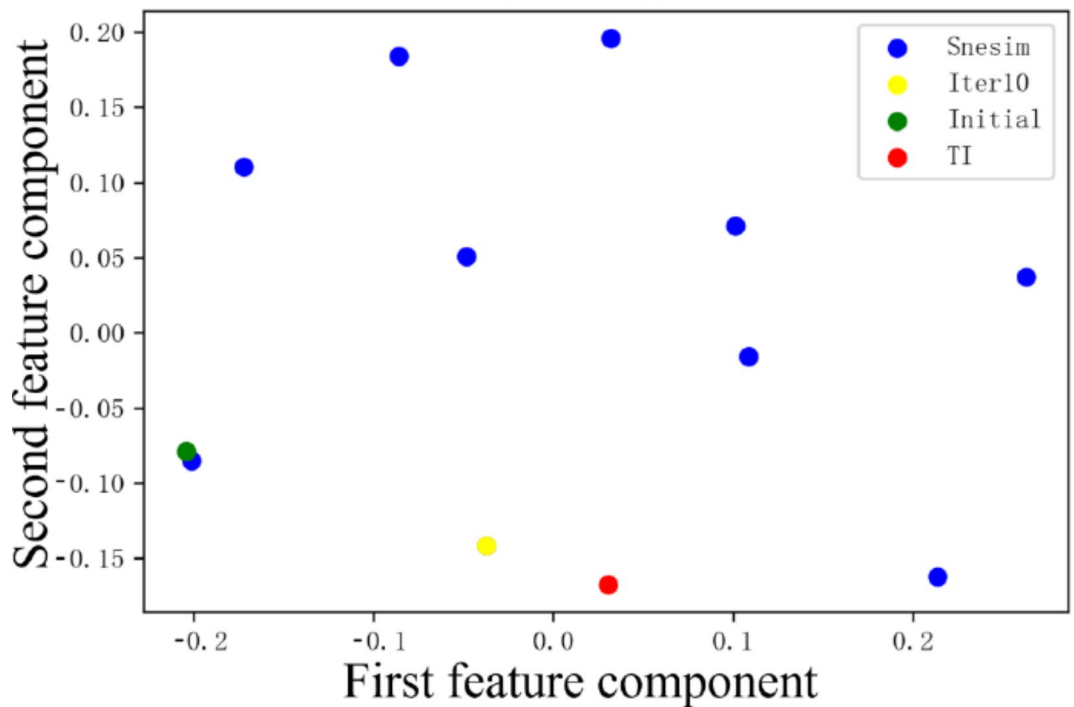


Fig. 17. MDS scatter plot.

the similarity (or dissimilarity) of the models, which in turn reflects the similarity (or dissimilarity) of spatial patterns. In Fig. 17, all points are distributed in the range $[-0.3, 0.3]$, with a relatively small discrepancy between the training image and the optimized realizations at each iteration (Blue point). We observe that the green point is far away from the red point, while the yellow point is close to the red point, showing a relatively minor discrepancy between the model realization generated from iteration 10 and TI. According to the connectivity function for both simulations and TI in the Y direction are similar to each other and the MDS scatter plot, it proves that for the 2d fluvial reservoir model, the simulations resemble the original TI in terms of attribute information reproduction, spatial variability characteristics, channel connectivity and so on, with a greater reconstruction diversity.

SISIM

The proposed workflow can be applied to other sequential simulation methods. In this case, we use the SISIM method to do simulation instead of SNESIM method. Figure 18A, derived from the synthetic data, is a reference model realization of 2 phases in volume. The model is discretized on a regular grid of $100 \times 60 \times 20$ cells of size $1 \times 1 \times 1$ m. The exponential variogram model was used for estimating the conditional distribution. The main anisotropic is horizontal and diagonal with respect to reservoir grid. The main azimuth of variogram function

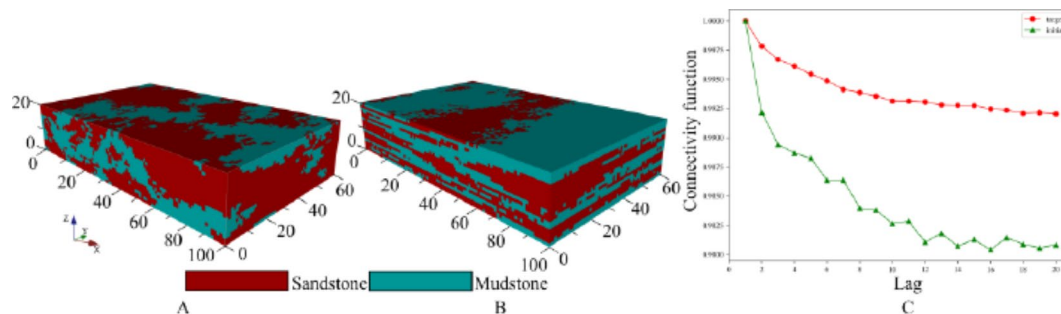


Fig. 18. A, A reference model. B, An initial model. C, Gap between the connectivity functions of the training image and the initial model.

is 10° , with dip and plunge angles of 0° . The ranges of the variogram function are 100, 50, and 1 m, respectively. The proportions of sandstone and mudstone are respectively 60 and 40%.

We consider the average connectivity function of reference model as the target connectivity function. Given that we believe the average connectivity function of this model is particularly smooth and the probability of connectivity for each lag is very high, which is a low-probability event in reservoir models, we want to use the workflow to generate more models that are similar to this high connectivity. Figure 18B show a realization of 2 facies model which is considered as the initial model. Figure 18C compares the connectivity functions of the reference model and the initial model, and shows a big gap to them. Figure 19 shows the model realizations resulting from each iteration of the optimization procedure. These realizations can be investigated successively to identify an optimal realization of the stochastic model. Figure 20 shows their corresponding connectivity functions (green) compared to the target (red). It can be seen that the connectivity function gets closer to the target as the number of iterations increases. Figure 21 demonstrates that after the first iteration, the realization of the initial model has been modified to be close to the target connectivity function, the objective function decreased by 77% (from 0.187632 to 0.044676), showing the efficiency of the gradual deformation method. A good match is obtained after fifteen iterations, the objective function decreased by 97% (from 0.187632 to 0.005815), although the deformed model is still different from the reference model. This is due to the connectivity function in this case do not fully condition the model. We also observe that the model realization appears the same from iteration 7 to 8, from iteration 11 to 12, from iteration 13 to 15, and from iteration 17 to 18, but there are noticeable changes in the realizations generated from iterations 1 to 4, 7 to 9, 11 to 13. This is further confirmed by the evolution of the objective function versus the number of iterations (Fig. 21B). The model realizations from iteration 13 to 18 are considered as the results of this workflow, due to the high similarity of connectivity function.

In Fig. 22, all points are distributed in the range $[-0.4, 0.4]$, with a relatively small discrepancy between the reference model and the optimized realizations at each iteration (Blue point). According to the average connectivity function for both simulations and the reference are similar to each other and the MDS scatter plot, it proves that for the 3d reservoir model, the simulations resemble the reference in terms of attribute information reproduction, spatial variability characteristics, sandbody connectivity and so on, with a greater reconstruction diversity. According to the results, this workflow can calibrate the model realization generated by SISIM to the target connectivity function.

Conclusions and future work

In this paper, we introduce a workflow to calibrate reservoir model realizations until honoring the connectivity data of the training image. This workflow integrates the gradual deformation method with the multiple-point geostatistical modeling approach. The model is constructed by single normal equation simulation (SNESIM). The connectivity criteria are defined by the connectivity function as implemented in the Code CONNEC3D. The gradual deformation of a stochastic model realization will yield a regular objective function which can be minimized by an optimization method. To optimize the realizations at each iteration, we employ the golden section search algorithm, which minimizes the objective function efficiently. The optimized realizations at each iteration all preserve the spatial variability of the initial model. The final outcomes are obtained by perturbing the random numbers of the initial model, with the perturbation aimed at matching the connectivity of the training image. Based on this, Multi-Dimensional Scaling (MDS) is utilized to evaluate the discrepancy between the training image and the optimized realizations at each iteration. Additionally, the uncertainty can be handled by generating multiple model realizations that satisfy the connectivity data of the training image through our workflow.

This workflow is applied to the calibration of a fluvial reservoir model, constructed by SNESIM, to match the connectivity function of the training image. Results from the case study demonstrate the efficiency of the workflow in calibrating fluvial reservoir models. For this case, only a few iterations were required to obtain satisfactory fields. Additionally, this workflow is suitable for other sequential simulation methods, and we have successfully applied the sequential indicator simulation (SISIM) into the proposed workflow. Multiple model realizations that satisfy the target connectivity data are generated to reduce the uncertainty.

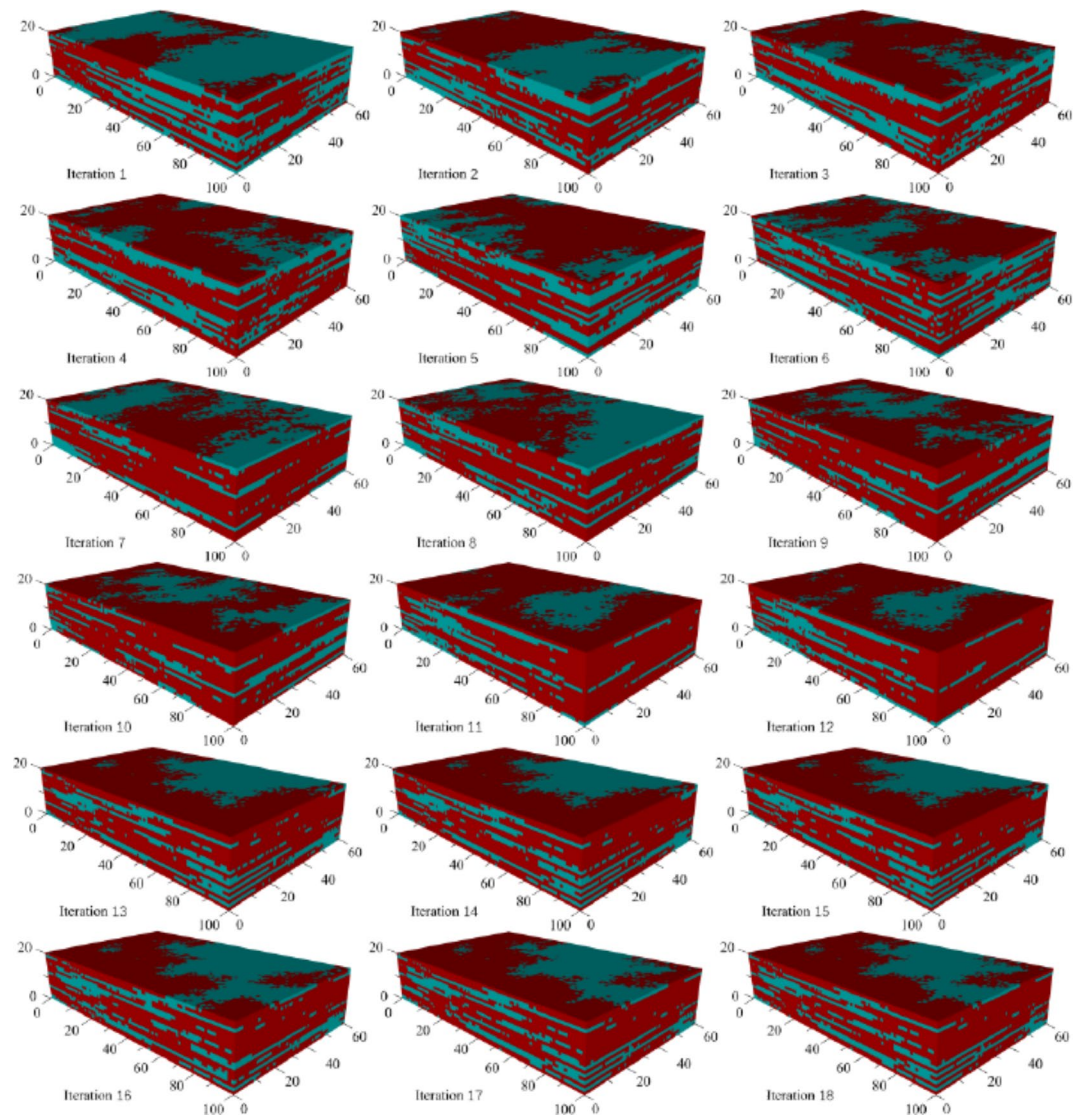


Fig. 19. The realizations of the 3 facies model resulting from each inner loop of the optimization procedure.

This workflow is particularly suitable for reconstructing model realizations that match the connectivity of the training image in multi-geostatistical modeling. However, there are limitations of this workflow. It is only suited for sequential simulation methods. The efficiency of the workflow depends greatly on how the model is deformed. Future enhancements may apply a GPU-based SNESIM method to the workflow. Additionally, there exist alternative model optimization algorithms (e.g., a gradient-based algorithm) which could offer further improvements. Future work will account for the uncertainty of model parameters by implementing multivariate optimization methods.

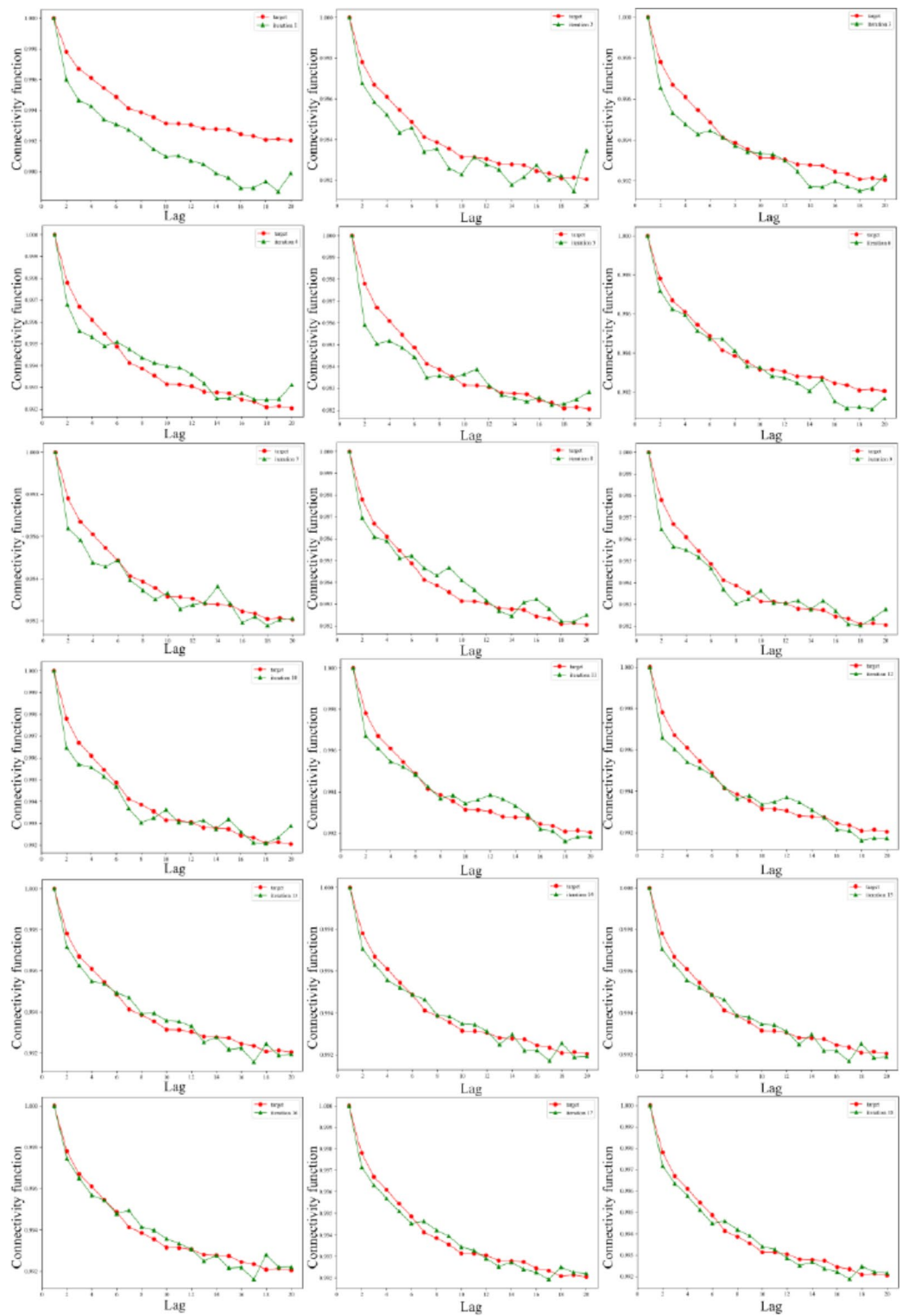


Fig. 20. Comparing the connectivity function after each iteration with the target.

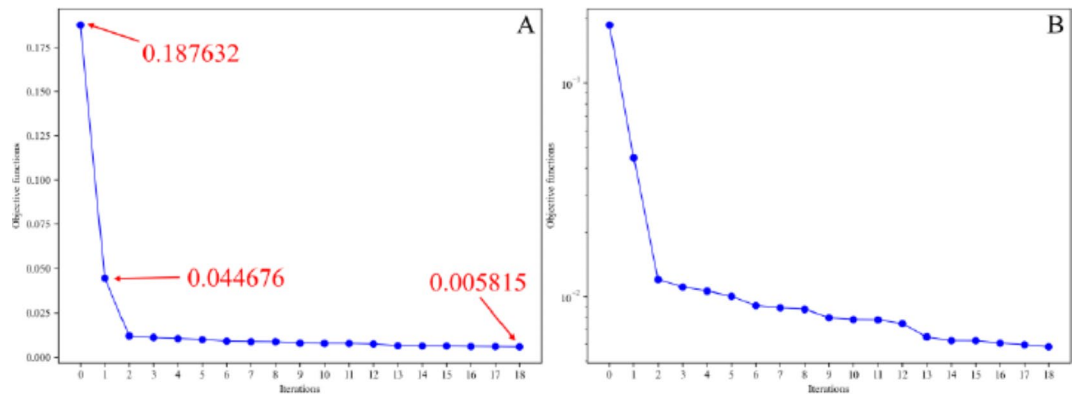


Fig. 21. A, Objective function versus the number of iterations. B, Objective function versus the number of iterations – objective function (ordinate) in logarithmic scale.

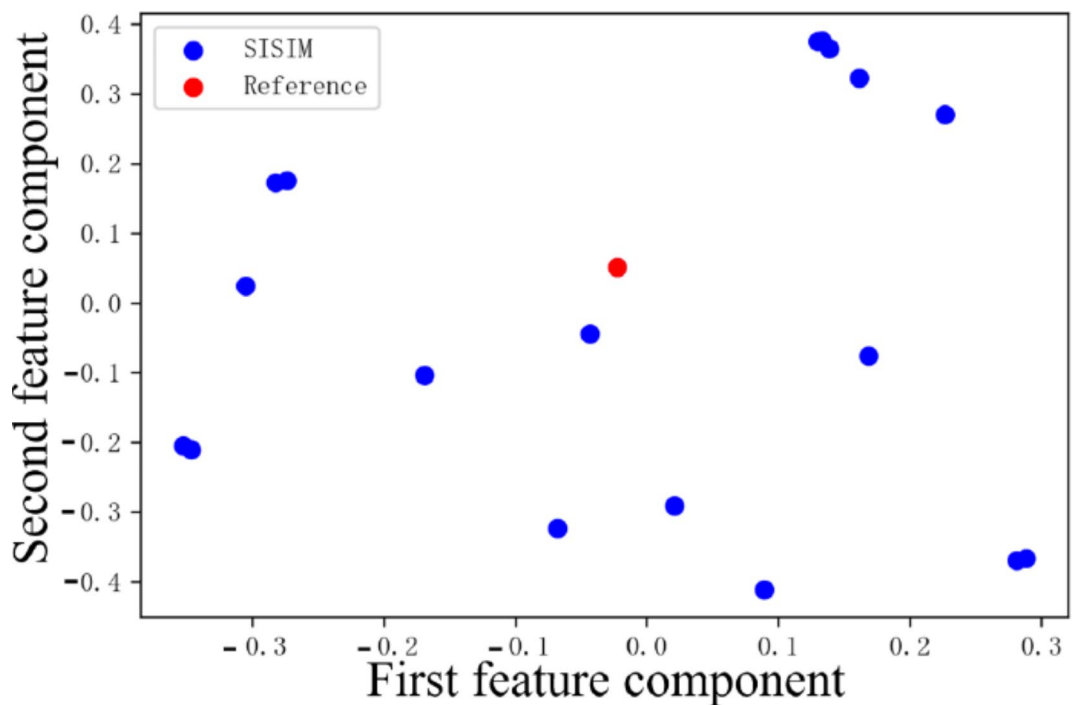


Fig. 22. MDS scatter plot.

Data availability

The datasets generated and/or analyzed during the current study are available from the first author on reasonable request.

Received: 13 August 2024; Accepted: 18 November 2024

Published online: 26 November 2024

References

- Allard, D. The HERESIM group.: on the connectivity of two random set models: the truncated Gaussian and the Boolean. In: (ed Soares, A.) Geostatistics Troia 92. Kluwer Academic, Dordrecht, The Netherlands, 467–478 https://doi.org/10.1007/978-94-011-1739-5_37 (1993).
- Allard, D. Simulating a geological lithofacies with respect to Connectivity Information using the truncated Gaussian Model. In: (eds Armstrong, M. & Dowd, P. A.) Geostatistical Simulations, Quantitative Geology and Geostatistics, vol 7. Springer, Dordrecht https://doi.org/10.1007/978-94-015-8267-4_16 (1994).
- Bukhanov, N., Subbotina, M., Voskresenkiy, A. & Katterbauer, K. Geological and dynamic similarity for reservoir state prediction by well connectivity. *Geoenery Sci. Eng.* **234**, 212667. <https://doi.org/10.1016/j.geoen.2024.212667> (2024).
- Caers, J. Efficient gradual deformation using a streamline-based proxy method. *J. Petrol. Sci. Eng.* **39**, 57–83. [https://doi.org/10.1016/S0920-4105\(03\)00040-8](https://doi.org/10.1016/S0920-4105(03)00040-8) (2003).

5. Caers, J. Comparing the gradual deformation with the probability perturbation method for solving inverse problems. *Math. Geol.* **39**, 27–52. <https://doi.org/10.1007/s11004-006-9064-6> (2007).
6. Ding, D. Y. Approximate derivative computations for the gradient-based optimization methods in the local gradual deformation for history matching. *Math. Geoscience.* **43**, 537–564. <https://doi.org/10.1007/s11004-011-9337-6> (2011).
7. Ding, D. Y. & Roggero, F. History matching Geostatistical Model realizations using a geometrical domain based parameterization technique. *Math. Geosci.* **42**, 413–432. <https://doi.org/10.1007/s11004-010-9273-x> (2010).
8. Gervais, V., Roggero, F., Feraillie, M., Le Ravalec, M. & Seiler, A. Joint History Matching of Production and 4D-Seismic Related Data for a North Sea Field Case. In: SPE Annual Technical Conference and Exhibition, Florence, Italy, paper SPE-135116-MS (2010). <https://doi.org/10.2118/135116-MS>
9. Hansen, T. M. & Bach, T. MPSLIB: a C++ class for sequential simulation of multiple-point statistical models. *SoftwareX* **5**, 127–133. <https://doi.org/10.1016/j.softx.2016.07.001> (2016).
10. Hoffman, B. T. & Caers, J. Regional probability perturbations for history matching. *J. Petrol. Sci. Eng.* **46**, 53–71. <https://doi.org/10.1016/j.petrol.2004.11.001> (2005).
11. Hovadik, J. M. & Larue, D. K. Static characterizations of reservoirs: Refining the concepts of connectivity and continuity. *Pet. Geosci.* **13**, 195–211. <https://doi.org/10.1144/1354-079305-697> (2007).
12. Hu, L. Y. Gradual deformation and iterative calibration of gaussian-related stochastic models. *Math. Geol.* **32**, 87–108. <https://doi.org/10.1023/A:1007506918588> (2000).
13. Hu, L. Y. & Blanc, G. Constraining a Reservoir Facies Model to Dynamic Data Using a Gradual Deformation Method. In: ECMOR VI – 6th European Conference on the Mathematics of Oil Recovery, Peebles, UK, paper B-01 (1998). <https://doi.org/10.3997/2214-4609.201406609>
14. Hu, L. Y. & Le Ravalec-Dupin, M. 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> (2004).
15. Hu, L. Y., Blanc, G. & Noetinger, B. Gradual deformation and iterative calibration of Sequential Stochastic simulations. *Math. Geol.* **33**, 475–489. <https://doi.org/10.1023/A:1011088913233> (2001).
16. Larue, D. K. & Hovadik, J. Connectivity of channelized reservoirs: a modelling approach. *Pet. Geosci.* **12**, 291–308. <https://doi.org/10.1144/1354-079306-699> (2006).
17. Le Gallo, Y., Le Ravalec-Dupin, M. & SPE-62922-MS. History Matching Geostatistical Reservoir Models with Gradual Deformation Method. In: SPE Annual Technical Conference and Exhibition, Dallas, Texas, paper (2000). <https://doi.org/10.2118/62922-MS>
18. Le Ravalec, M. & Mouche, E. Calibrating transmissivities from piezometric heads with the gradual deformation method: an application to the Culebra Dolomite unit at the Waste isolation pilot plant (WIPP), New Mexico, USA. *J. Hydrol.* **472–473**, 1–13. <https://doi.org/10.1016/j.jhydrol.2012.08.053> (2012).
19. Le Ravalec, M., Coureaud, B., Nicolas, L. & Roggero, F. Conditioning an Underground Gas Storage Site to Well pressures. *Oil Gas Sci. Technol.* **59**, 611–624. <https://doi.org/10.2516/ogst:2004044> (2004).
20. Le Ravalec-Dupin, M. L. & Hu, L. Y. Combining the pilot point and gradual deformation methods for calibrating permeability models to Dynamic Data. *Oil Gas Sci. Technol.* **62**, 169–180. <https://doi.org/10.2516/ogst:2007015> (2007).
21. Li, S. H., Zhang, C. M. & Deutsch, C. V. Ranking realizations of stochastic modeling based on reservoir static geological parameters. *Acta Petrolei Sinica.* **31**, 445–448. <https://doi.org/10.7623/syxb201003016> (2010).
22. Marsily, G., de, G., Lavedan, Boucher, M. & Fasanino, G. Interpretation of Interference Tests in a Well Field Using Geostatistical Techniques to Fit the Permeability Distribution in a Reservoir Model. In: Geostatistics for natural resources characterization, South Lake Tahoe, California, USA, 831–849 (1984).
23. Marteau, B., Ding, D. Y. & Dumas, L. A generalization of the local gradual deformation method using domain parameterization. *Comput. Geosci.* **72**, 233–243. <https://doi.org/10.1016/j.cageo.2014.08.004> (2014).
24. Oliver, D. S., Cunha, L. B. & Reynolds, A. C. Markov chain Monte Carlo methods for conditioning a permeability field to pressure data. *Math. Geol.* **29**, 61–91. <https://doi.org/10.1007/BF02769620> (1997).
25. Pardo-Igúzquiza, E. & Dowd, P. A. CONNEC3D: a computer program for connectivity analysis of 3D random set models. *Comput. Geosci.* **29**, 775–785. [https://doi.org/10.1016/S0098-3004\(03\)00028-1](https://doi.org/10.1016/S0098-3004(03)00028-1) (2003).
26. Pirot, G., Joshi, R., Giraud, J., Lindsay, M. D. & Jessell, M. W. loopUI-0.1: indicators to support needs and practices in 3D geological modelling uncertainty quantification. *Geosci. Model Dev.* **15** (12), 4689–4708. <https://doi.org/10.5194/gmd-15-4689-2022> (2022).
27. Pourpak, H., Bourbiaux, B., Roggero, F. & Delay, F. An Integrated Method for Calibrating a Heterogeneous/Fractured Reservoir Model from Wellbore Flow measurements: Case Study. *SPE Reservoir Eval. Eng.* **12**, 433–455. <https://doi.org/10.2118/113528-PA> (2009).
28. Renard, P. & Allard, D. Connectivity metrics for subsurface flow and transport. *Adv. Water Resour.* **51**, 168–196. <https://doi.org/10.1016/j.advwatres.2011.12.001> (2013).
29. Renard, P., Straubhaar, J., Caers, J. & Mariethoz, G. Conditioning facies simulations with connectivity data. *Math. Geosci.* **43**, 879–903. <https://doi.org/10.1007/s11004-011-9363-4> (2011).
30. Roggero, F. & Hu, L. Y. Gradual deformation of continuous geostatistical models for history matching. In: SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, paper SPE-49004-MS (1998). <https://doi.org/10.2118/49004-MS>
31. Strebelle, S. *Sequential Simulation Drawing Structures from Training Images* 39–55 (Stanford University Press, 2000).
32. Strebelle, S. Conditional simulation of complex geological structure using multiple-point statistics. *Math. Geol.* **34**, 1–21. <https://doi.org/10.1023/A:1014009426274> (2002).
33. Strebelle, S. & Journel, A. G. Reservoir modeling using multiple-point statistics. In: SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, 55–66 (2001). <https://doi.org/10.2118/71324-MS>
34. Tan, X., Tahmasebi, P. & Caers, J. Comparing training-image based algorithms using an analysis of Distance. *Math. Geosci.* **46**, 149–169. <https://doi.org/10.1007/s11004-013-9482-1> (2014).
35. Yang, X. W. & Zhu, P. M. A fast stochastic impedance inversion under control of sedimentary facies based on geostatistics and gradual deformation method, In: SEG International Exposition and Annual Meeting, Dallas, Texas, paper SEG-2016-13879993 (2016). <https://doi.org/10.1190/segam2016-13879993.1>
36. Yang, X. W. & Zhu, P. M. Stochastic seismic inversion based on an improved local gradual deformation method. *Comput. Geosci.* **109**, 75–86. <https://doi.org/10.1016/j.cageo.2017.08.010> (2017).
37. Yang, X. W. & Zhu, P. M. Stochastic seismic inversion of nonstationary model based on multigrid gradual deformation method, In: SEG International Exposition and Annual Meeting, Houston, Texas, paper SEG-2017-17725644 (2017). <https://doi.org/10.1190/segam2017-17725644.1>

Acknowledgements

This work is supported by the National Natural Science Foundation of China (No.42172172). Meanwhile, we are grateful to the editors and anonymous reviewers for their constructive comments and efforts with this manuscript.

Author contributions

Initial idea and research plan were provided by S.H. Li. Material preparation, data collection and analysis were performed by J.H. Jin, F. Ding, J. Li. J.H. Jin wrote the main manuscript text. All authors reviewed the manuscript.

Funding

The project was supported by the National Natural Science Foundation of China (No.42172172), and we would like to ensure this information is correctly incorporated into the final publication.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024