Choosing between Gaussian and MPS simulation: the role of data information content—a case study using uncertain interpretation data points

Rasmus Bødker Madsen¹ · Ingelise Møller¹ · Thomas Mejer Hansen²

Accepted: 21 November 2020 / Published online: 7 January 2021
© The Author(s) 2021

Abstract

Integration of geophysical data with information such as from boreholes and expert domain knowledge is often performed as cognitive or explicit geological modeling when generating deterministic geological models of the subsurface. However, such cognitive geological models lack the ability to express the uncertainty of layer boundaries. To remedy the shortcomings of this strategy we propose a novel stochastic methodology combining the efforts of probabilistic data integration and cognitive modeling. We treat geological interpretation points from the cognitive model as uncertain “soft” data. These data are then combined with analogous geology in a probabilistic model. We test two ways of combining and sampling from such a probabilistic model. Firstly, a high-entropy setup based on Gaussian distributions simulation. Secondly, lower entropy (and conceivable more realistic) geological solutions are obtained from multiple-point geostatistics (MPS). We apply both ways of solving the problem at a study site near Horsens, Denmark, where airborne transient electromagnetic measurements, seismic data, and borehole information are available and interpreted in cognitive modeling. We explain the complete framework for integrating the uncertain interpretations in geostatistical simulation. Results show that both the Gaussian simulation and multiple-point geostatistical approach allows satisfactory simulations of uncertain geological interpretations and are consistent with prior geological knowledge. Our results suggest that the number of uncertain data points and their information content play a pivotal role in selecting the most appropriate simulation method for the given framework. MPS simulations allow connectivity in scenarios with few data points due to the low entropy of the model. When the number of soft data increases, Gaussian simulation is less prone to produce simulation artifacts, faster to apply, and could be considered preferential to using MPS simulation.

Keywords Geostatistics · Gaussian simulation · MPS simulation · Uncertain continuous data · Simulation method selection

1 Introduction

Geoscience as any other scientific discipline relies on gathering of information to describe and enable a deeper understanding of processes that form the subsurface. Sources of information usually vary from survey to survey. In some cases, direct geological observations and measurements (e.g. from boreholes or exposed geological outcrops) are available. Direct observations are typically far between due to physical constraints. Some interpolation between direct observations is therefore needed to form a bigger picture of the geology. Thus, geophysical data are often collected to increase spatial information of geology and aid the interpolation. These geophysical data are commonly processed to produce geophysical images that

✉ Rasmus Bødker Madsen
rbm@geus.dk

Ingelise Møller
ilm@geus.dk

Thomas Mejer Hansen
tmeha@geo.au.dk

¹ Groundwater and Quaternary Geology Mapping, Geological Survey of Denmark and Greenland (GEUS), C.F. Møllers Alle 8, 8000 Aarhus C, Denmark

² Department of Geoscience, University of Aarhus, Høgh-Guldbergs Gade 2, 8000 Aarhus C, Denmark

are easy to interpret and form the basis for conceptual geological understanding (Wellmann and Caumon 2018). The goal of cognitive or explicit structural geological modeling is then to integrate/interpret these multiple sources of subsurface information with pre-existing geological knowledge into a combined model of interfaces between geological units (Royse 2010). Examples of such cognitive structural modeling to form 3D structural geological models can e.g. be found in Bosch et al. (2009), Jørgensen et al. (2013, 2015), Oldenborger et al. (2014) and Høyer et al. (2015, 2019). The popularity of the cognitive approach may be attributed to the easy and intuitive integration of geological knowledge (Wycisk et al. 2009; Jørgensen et al. 2013).

Because only a single model is produced, it is difficult to account for sources of uncertainty within the framework of cognitive modeling. Uncertainties and biases are introduced at all steps of the model creation. From uncertainties on the measurements during the acquisition, through the way the data is processed and finally in the combination with other sources of information. Uncertainty also arise due to spatial variability at unsampled locations. The interpreter might also induce some subjective information into the model (Wycisk et al. 2009; Bond 2015). Even direct geological observations can have non-quantifiable properties that instead need personal interpretation (Wellmann and Caumon 2018). Accepting the presence of uncertainty in the input information inevitably allows a multitude of models to be reasonable representations of the subsurface.

Within the field of geostatistics, different methods allow the incorporation of such uncertainties. The results are typically not analytically obtainable but can instead be simulated as realizations through sampling procedures (Journel and Huijbregts 1978; Journel 2002; Lantuéjoul 2002; Remy et al. 2009; Mariethoz et al. 2010; Chilès and Delfiner 2012; Mariethoz and Lefebvre 2014; Laloy et al. 2017). There exist two main types of geostatistical simulation methods: those based on two-point statistics, where the statistical models are assumed as some variant of a multivariate Gaussian model quantified by a mean and a covariance, and those based on multiple-point statistics in which higher-order statistics are inferred from an example model, such as a training image (TI).

Gaussian simulation methods (Journel and Huijbregts 1978; Goovaerts 1997) are based on covariance functions and semivariograms with correlations depending on the separation between pairs of points in space. Gaussian models can hence be described by an analytical expression and are attributed to high-entropy, i.e. relatively low information content (Shannon 1948; Journel and Zhang 2006; Hansen 2020). The more recently developed multiple-point geostatistics (MPS) (Strebelle 2002; Mariethoz

et al. 2010; Rezaee et al. 2013; Mariethoz and Caers 2015; Hansen et al. 2016; Hoffmann et al. 2017; Gravey and Mariethoz 2020) makes connectivity of features possible, which lowers information entropy. This has a great beneficial impact on subsequent flow modeling and tends to produce more geological realism (Mariethoz and Caers 2015). Kessler et al. (2013) for example showed better reproduction of sand lens geometry using MPS than a method based on high-entropy transition probabilities. The process of acquiring such a TI is not always straightforward (Tahmasebi 2018) and the finite size plays a pivotal role in how well the statistical model is reproduced in simulations (Emery and Lantuéjoul 2014). A common criticism of MPS is that TIs carry too much subjective information, although choosing a Gaussian model is a subjective choice in itself (Journel and Zhang 2006). Furthermore, one also has to address issues of how well simulations from the TI are a reasonable representation of the geology in question (Høyer et al. 2017). The issue of quantifying representative geology is also present in the Gaussian case (Li et al. 2015). Recent research involves the inclusion of non-stationarity (Sabeti et al. 2017; Madsen et al. 2020a) and multi-modality (Grana et al. 2017; De Figueiredo et al. 2019) while maintaining a computationally feasible problem (Zunino and Mosegaard 2019). MPS simulations are usually more computationally expensive than Gaussian simulation. However, with computational power increasingly available over time, so is the utility of MPS.

TI versus semivariogram based simulation has previously been discussed in the geoscientific community (see e.g. Journel and Zhang 2006; Kessler et al. 2013; Li et al. 2015). With the separate advantages and disadvantages offered by Gaussian and MPS simulation, it is currently suggested that they should be viewed as complimentary simulation strategies (Mariethoz 2018). These comparison studies between Gaussian and MPS simulation are typically based on synthetic data sets to make comparison of performance easier. However, to our knowledge, few examples (e.g. Kessler et al. 2013) exist of solving a real-world applied problem with both methods.

In the current study, we apply both Gaussian and MPS simulation techniques to solve a specific problem of generating realizations of the depth to one interface separating two geological units based on uncertain interpretation points from a cognitive structural model. The structure of the paper is as follows. First, in Sect. 2, we introduce the data available and the modeling setup. In Sect. 3, we introduce the two modeling techniques in detail. Then, in Sect. 4, we create a base case with all available data. The sensitivity of each method to data availability is evaluated by decreasing the number of data used to condition the results. Here two cases are used in which the conditioning

data is chosen as a (1) random subset of available data, and (2) a subset of the data with a local high gradient, representing potentially a more informed data subset. This provides insights to the nature of picking interpretation points within cognitive modeling and whether these points can be chosen cleverly to better guide and inform geostatistical simulations. When using Gaussian simulation we propose to infer the properties of the Gaussian covariance directly from the interpretation points, thus eliminating the need for a TI or other ways of quantifying prior geological information (Sect. 5). The case study (as discussed in Sect. 6), therefore, presents, for the first time, a complete framework for the steps needed to integrate uncertain interpretation points within different geostatistical simulation methods. Furthermore, it reveals insights into the advantages and disadvantages of each simulation method from an applied perspective.

2 Interpretation data

Cognitive geological modeling is the primary method used for 2D and 3D interpretations of the subsurface in Denmark (Jørgensen and Sandersen 2006; Sandersen et al. 2009; Høyer et al. 2015). Recently, efforts in mapping the hydrogeology of the substratum in Denmark also include geostatistical simulation (Høyer et al. 2017; Barfod et al. 2018; Vilhelmsen et al. 2019; Madsen et al. 2020b). Here, we focus on a study area situated just outside the town of Horsens, Denmark (Fig. 1).

An interpreted cognitive geological 3D model from geophysical (airborne electromagnetic and seismic reflection) data, borehole logs, and digital elevation maps exists for the study area (Jørgensen et al. 2010; Møller and Jørgensen 2011). The data used in the cognitive modeling study of Jørgensen et al. (2010) are available through the geophysical database GERDA (Møller et al. 2009) and the shallow borehole database JUPITER (Hansen and Pjetursson 2011). Glacial processes and their products, including buried tunnel valleys, influence the subsurface of the Horsens area. Buried tunnel valleys are elongated structures formed by meltwater erosion underneath the margins of ice-sheets during glaciations (e.g. Sandersen and Jørgensen 2017). Such buried valleys are common features in the Danish substratum both onshore and offshore (Huuse and Lykke-Andersen 2000; Stewart 2009; Sandersen and Jørgensen 2017). In Denmark, the buried valleys are incised into both Quaternary and Pre-Quaternary sediments and are filled with Quaternary deposits (Sandersen and Jørgensen 2017). Once these buried valleys are filled with younger material, the hydraulic properties are in contrast to the underlying sediments (Sandersen and Jørgensen 2012). Detailed mapping of buried valleys is

therefore of importance to determine flow paths in groundwater aquifer systems (Refsgaard et al. 2012; Andersen et al. 2013).

We focus here on the interpretation of the interface between Quaternary sediments and the Paleogene clay sediments underneath. In the following, we refer to these interpretation points as our conditional data or simply data. Data used in this study ($N_d = 379$) are shown at their geographical location in Fig. 1 and in more detail on an arbitrary grid in Fig. 2a. Each data point displays the elevation of the interface in meters above sea level (m a.s.l.). The interface spans roughly an interval between 0 m and -250 m and covers an area of approximately 40 km^2 ($5.6 \text{ km} \times 7 \text{ km}$). By examining the data, three and possibly four buried valleys are visually identified. The buried valleys mainly align in NW–SE trending elongated incisions of lower elevation compared to the plateau. One valley runs perpendicular (NE–SW) in the upper right corner of Fig. 2a.

Each data point is attributed to a qualitative uncertainty measure during the interpretation. The qualitative uncertainty is expressed through three categories and is displayed in Fig. 2e. These qualitative assessments do not discern between uncertainties on the coordinate, measurement data, the existence of geological elements, etc. Instead, they are based on the combined judgment of all these factors by the interpreter. Uncertainties with category 1 are the most certain interpretations. Boreholes and electromagnetic data guide these points because airborne electromagnetic data are very sensitive to the highly conducting clays (Danielsen et al. 2003; Schamper et al. 2014). This makes electromagnetic data well suited for mapping buried valley structures in the Paleogene clay. The interface is therefore well resolved at these points; hence, the uncertainty is low in the upper right corner of the Horsens area (see Fig. 2e). Interpretation points based on wells are typically also attributed a low uncertainty (e.g. the sole black dot at $X = 3$ and $Y = 28$ in Fig. 2e). Interpretations based on seismic images and some places with shallow depth of investigation of the airborne methods comprise the second uncertainty category (pink dots in Fig. 2e). The most uncertain data (yellow dots in Fig. 2e) are based primarily on geological interpretation that are far from any geophysical data or boreholes.

A normal distribution with zero mean and a specific standard deviation for each category is assumed to represent interpretation uncertainty based on the interpreter's judgment. Figure 2f shows a graphical representation of the attributed uncertainty for each category. The most certain category consists mainly of electromagnetic measurement points. Even though the method is very effective at resolving the interface at the Paleogene clay, some uncertainty is expected at these points due to the vertical

Fig. 1 Overview of the study area displaying the location of data and training image

grid cell spacing of 5 m. Including potential modeling/inversion/measurement errors, a standard deviation of 2.5 m is used to describe the uncertainty for category 1. The seismic data available are of poor quality (low signal to noise ratio). We assign category 2 with a standard deviation of 8 m, due to uncertainties related to data quality and depth conversion. Interpretation points of category 3 are described as weakly documented with little data evidence, weak knowledge on model formation, and big heterogeneity compared with the scale (Møller and Jørgensen 2011). A standard deviation of 12 m is chosen for these points.

During the interpretation phase of cognitive modeling, the geologist generally proceeds by considering nearby data points that have been already interpreted. Some degree of correlation between the uncertainties must therefore be expected. For example, if the elevation of the interface is set higher than its real value by the interpreter, adjacent points are also likely to be set higher, which causes correlation in the data uncertainty. We assume that these

correlations follow the structures of the data. Using a spherical type semivariogram model, we compute the likelihood for a set of ranges given the data and uncertainties available (see Sect. 3.2 for details). The normalized likelihoods are presented in Fig. 2d. For the full dataset, the range of the correlation is likely around 3500 m. Because we know the spatial dependency between the data points and the semivariogram model, a covariance model is easily computed. We construct a covariance matrix for all interpretation points $C_{d,shape}$ with sill = 1 m using the inferred range. This covariance matrix is then scaled such that it has the correct standard deviation for each interpretation point as described above:

$$C_d = C_{d,std} C_{d,shape} C_{d,std} \quad (1)$$

where the $C_{d,std}$ ($N_d \times N_d$) is a covariance matrix with the standard deviations of each interpretation point along the diagonal. The final Gaussian model for the elevation uncertainty is described by the covariance matrix C_d (Fig. 2g) calculated from Eq. 1 and a mean vector $\mathbf{0}$

Fig. 2 **a** Overview of available interpretation points of elevation (conditional data). **b** Reduced data set of 60 points based on gradients. **c** Reduced data set of 60 randomly selected points. **d** Inference of range from full data set (blue) and the reduced data sets (yellow and red). **e** Attributed qualitative uncertainty category for each

interpretation point. Highest uncertainty (yellow) to lowest uncertainty (black). **f** 1D representation of quantitative uncertainties in elevation depending on the uncertainty category. **g** The covariance matrix of data uncertainty

($N_d \times 1$) consisting of zeros. The values on the diagonal of C_d refers to the variance of each individual interpretation point whereas the off-diagonal values describe the correlation between these variances.

3 Modeling strategies

The problem at hand is to simulate the location of the continuous interface of the pre-Quaternary based on the uncertain interpretation points in Fig. 2a. The simulation grid, $\mathbf{m}$, covers all data points in a grid of size $[N_x, N_y] = [56, 70]$ with a grid cell size of $100 \text{ m} \times 100 \text{ m}$. We intend to describe the uncertainty of the interface. Therefore, a simple interpolation does not suffice, as only a single model is produced and data points themselves are typically assumed certain. Instead, simulation is used to generate multiple realizations of the depth to the interface.

3.1 Gaussian modeling

In Gaussian modeling, the correlations between model parameters are described using semivariograms and covariance functions. The underlying assumption is that the model parameters $\mathbf{m}$ are distributed following a prior Gaussian distribution $\mathcal{N}(\mathbf{m}_{\text{prior}}, \mathbf{C}_M)$. In addition, information about some of the model parameters is available through the Gaussian data model $\mathcal{N}(\mathbf{d}_{\text{obs}}, \mathbf{C}_d)$ (see above), representing interpreted data. The conjunction of these two Gaussian distributions is a ‘posterior’ Gaussian distribution (Tarantola and Valette 1982). Realizations of the posterior Gaussian distribution can be generated in a variety of ways.

One approach is to use linear least-squares inversion to analytically compute the posterior Gaussian distribution $\mathcal{N}(\tilde{\mathbf{m}}, \tilde{\mathbf{C}}_M)$ as (Tarantola and Valette 1982; Omre 1987):

$$\tilde{\mathbf{C}}_M = (\mathbf{G}^T \mathbf{C}_d^{-1} \mathbf{G} + \mathbf{C}_M)^{-1} \quad (2)$$

$$\tilde{\mathbf{m}} = \mathbf{m}_{\text{prior}} + \mathbf{C}_M \mathbf{G}^T (\mathbf{G} \mathbf{C}_M \mathbf{G}^T + \mathbf{C}_d)^{-1} (\mathbf{d}_{\text{obs}} - \mathbf{G} \mathbf{m}_{\text{prior}}) \quad (3)$$

where $\mathbf{G}$ is the linear forward operator. Because both model parameters and data are elevations at certain locations, the forward operator $\mathbf{G}$ is a binary matrix ($N_d \times N$) with 1 where data coincides with the simulation grid. Realizations of the final Gaussian distribution can be drawn using e.g. Cholesky decomposition (Higham 2009). In the outlined scenario of simulating the pre-Quaternary interface the data and model parameters are both given in elevation.

For our current problem with only $N = N_x \times N_y = 3920$ model parameters, Cholesky realizations of the Gaussian solution takes only a few milliseconds to compute and are stable. However, larger system Cholesky decomposition becomes too computationally demanding. For such problems, one can make use of sequential Gaussian simulation (SGS) to generate Gaussian realizations (Alabert 1987; Gómez-Hernández and Cassiraga 1994; Remy et al. 2009). SGS involves solving a number of smaller least-squares/kriging systems many times (one time per model parameter), as opposed to using Eqs. 2 and 3 in which one big least-squares system is solved once. Hence, SGS can potentially be significantly faster for larger problems. If simple kriging is utilized as part of the sequential Gaussian simulation, one should obtain equivalent simulation results as those obtained using linear least-squares followed by Cholesky decompositions, as long as the used data neighborhood is large enough (Gómez-Hernández and Cassiraga 1994; Emery and Peláez 2011; Nussbaumer et al. 2018).

3.2 Gaussian likelihood model

Using Gaussian simulation methods, one must choose the mean $\mathbf{m}_{\text{prior}}$ and covariance $\mathbf{C}_M$. Traditionally semivariogram analysis can be used to infer the properties of the covariance model from observed point data. This, however, is only viable if the observed data are noise-free. In the present case observed data are associated with uncertainty, as described above. In this case one can use a Bayesian approach to the inference of the parameters of the covariance model. The following likelihood function describes the (unnormalized) likelihood of a specific choice of covariance model, $\mathbf{C}_M$, given the observed data, $\mathbf{d}_{\text{obs}}$, for a system with linear forward operator $\mathbf{G}$ and a Gaussian distributed prior $\mathcal{N}(\mathbf{m}_{\text{prior}}, \mathbf{C}_M)$ and data uncertainty $\mathcal{N}(\mathbf{0}, \mathbf{C}_d)$ (Hansen et al. 2015):

$$L = \left((2\pi)^N |\mathbf{G}\mathbf{C}_M\mathbf{G}^T + \mathbf{C}_d| \right)^{-0.5} \exp \left(-\frac{1}{2} (\mathbf{d}_{\text{obs}} - (\mathbf{G}\mathbf{m}_{\text{prior}} + \mathbf{0}))^T (\mathbf{G}\mathbf{C}_M\mathbf{G}^T + \mathbf{C}_d)^{-1} (\mathbf{d}_{\text{obs}} - (\mathbf{G}\mathbf{m}_{\text{prior}} + \mathbf{0})) \right) \quad (4)$$

where $|\cdot|$ is the determinant. We choose to model $\mathbf{C}_M$ as an isotropic spherical type covariance model and thus the only free parameters in $\mathbf{C}_M$ are the range and the sill. We utilize the data to infer the values and let the prior be defined in a grid consisting only of the data points (not the entire simulation domain). Thus, the linear forward operator $\mathbf{G} = \mathbf{I}$ is then a diagonal matrix consisting of ones. The expression $\mathbf{G}\mathbf{m}_{\text{prior}} + \mathbf{0}$ simplifies to our a priori expectation value for the data $\mathbf{d}_{\text{prior}}$ and we can rewrite Eq. 4 as:

$$L = \left((2\pi)^N |\mathbf{C}_M + \mathbf{C}_d| \right)^{-0.5} \exp \left(-\frac{1}{2} (\mathbf{d}_{\text{obs}} - \mathbf{d}_{\text{prior}})^T (\mathbf{C}_M + \mathbf{C}_d)^{-1} (\mathbf{d}_{\text{obs}} - \mathbf{d}_{\text{prior}}) \right) \quad (5)$$

where the full $\log(L)$ is then given by:

$$\log(L) = -\frac{N}{2} - \frac{N}{2} \log(2\pi) - \log |\mathbf{C}_M + \mathbf{C}_d| - \frac{1}{2} (\mathbf{d}_{\text{obs}} - \mathbf{d}_{\text{prior}})^T (\mathbf{C}_M + \mathbf{C}_d)^{-1} (\mathbf{d}_{\text{obs}} - \mathbf{d}_{\text{prior}}) \quad (6)$$

Minimizing Eq. 6 for, in this case, the range and sill, will provide the optimal (most likely) choice of $\mathbf{C}_M$, given the observed data with uncertainty.

3.3 MPS modeling

As for Gaussian based modeling, multiple algorithms exist that allow sampling from MPS models, such as ENESIM (Guardiano and Srivastava 1993), GENESIM (Hansen et al. 2016), Direct Sampling (DS) (Mariethoz et al. 2010), SNESIM (Strebelle 2002), IMPALA (Straubhaar et al. 2011), CCSIM (Tahmasebi et al. 2012), QS (Gravey and Mariethoz 2020). For an extended review of MPS sampling strategies and their drawbacks and benefits, the reader is referred to Mariethoz and Caers (2015). All these algorithms allow using a discrete training image, but only a few, such as DS, CCSIM and QS, allow using a continuous training image, as is needed here. In the following, we make use of the Direct Sampling MPS algorithm (Mariethoz et al. 2010) as implemented in DEESSE (Straubhaar 2019). The main reason for using DS is that it allows simulation of continuous parameters and that it is widely used.

3.4 Prior information

Regardless of which method is used for simulation, a way of describing our expectations about the system (prior knowledge) needs to be quantified. For MPS this is done via a TI. Here, we extract the pre-Quaternary Paleogene clay surface from the common public hydrological model, FOHM, in a region situated north and northwest of Aarhus roughly 40 km from the study area (see Fig. 1). The FOHM model is compiled from numerous cognitive models constructed within the Danish national groundwater

mapping program (Thomsen et al. 2004) and covers the entire Western Denmark (Arvidsen et al. 2020). The area north of the city of Aarhus is, as the study site and the rest of eastern Jutland, characterized by incised buried valleys formed by Quaternary meltwater (Høyer et al. 2015). We adopt this surface as the TI because the areas are geographically close, and we can expect that similar geological processes shaped the landscape at both locations. Moreover, the analog region is interpreted through cognitive modeling like the Horsens data.

Fig. 3 **a** TI displaying buried valley structures in the Paleogene clay deposits from an area north and northwest of Aarhus. **b** The TI from **a** scale corrected to match the elevation heights of data in Horsens. **c** Experimental semivariogram of both unscaled and final TI along with a fitted spherical type semivariogram. **d** Histograms of elevations

for the Horsens data and the two TIs plus 1D Gaussian distributions with sample mean and standard deviation. The histograms are normalized to probability. **e** Quantile-quantile plot of TI and data values

Figure 3a displays the extracted surface of the Paleogene clay from the area near Aarhus (Fig. 1). TIs are in essence an abstraction that should be infinitely large. For practical purposes, TIs are understandably finite and their size plays a role in the number of possible patterns to be simulated. Although workarounds do exist to enhance the number of possible patterns from a small training image, a general rule of thumb is the larger the TI the better. The TI is approximately 21 times larger than the simulation grid from Horsens (black rectangle in Fig. 3a). Even though the two areas display the same subsurface structures affected by the same geological processes, the elevation of the Paleogene clay is found much deeper in Horsens than north and northwest of Aarhus as seen in Fig. 3a. To remedy this, we scale the TI to match the elevations expected in Horsens. The incised valleys with low elevation comprise about 20% of the TI whereas the remaining 80% lie as a smooth “plateau” at higher elevations. We compute the mean of the highest 80% percent of elevations for the interpretation points in the Horsens data. Then this mean is subtracted from the elevations in the TI ensuring that the elevations of the plateau in the TI match the values in Horsens. The valleys in the TI are not incised as deep as in the data. We further scale the values of the TI by multiplying with a heuristic factor of 1.25 to remedy this problem. The final TI used for the simulations is shown in Fig. 3b. Besides the elevation difference, another downside is the non-stationarity of the proposed TI. This is unfortunately a recurring theme in MPS modeling since stratigraphic and structural events lead to significant spatial variation (Strebelle 2012). To enhance the statistical foundation for finding patterns in the TI, we make use of scaling and rotation (Mariethoz et al. 2010). By allowing rotation and scaling, the orientation of the buried valleys in the TI becomes less important as they can be recreated in any direction. This improves vastly on the non-stationarity issue. Rotation and scaling should also have a positive impact on verbose-copying, wherein parts of the TI is recreated entirely within the simulation domain, as the number of possible matches increases (Mariethoz and Caers 2015).

For the Gaussian method, the prior information should be provided by a Gaussian covariance model $\mathcal{N}(\mathbf{m}_{\text{prior}}, \mathbf{C}_M)$. We infer such a covariance matrix from the TI using an experimental semivariogram (red curve in Fig. 3c). A spherical semivariogram model is fitted based on linear decreasing weighting with distance (black curve). The fitted spherical model has a range of 2660 m and a sill of 2080 m² and shows a reasonable fit with the experimental semivariogram. The fitted model shows a slight tendency to overestimate the variance at lower ranges while underestimating the larger ranges. Although larger ranges are based on fewer samples, this misestimation

might indicate that Gaussian simulation based on this semivariogram will tend to fail in reproducing longer connected features.

A semivariogram analysis is also applied to the unscaled TI for comparison (blue curve). The two models differ in sill but show an identical range. This range is somewhat smaller but comparable to the one inferred for the data in Fig. 2d. We define a stationary mean vector $\mathbf{m}_{\text{prior}}$ ($N \times 1$) equal to the mean of the TI and construct a covariance matrix $\mathbf{C}_M$ from the semivariogram model of the TI. We choose the TI for creating the covariance model instead of the interpretation points in Horsens for two reasons. Firstly, the TI represents independent information. Secondly, it makes comparisons between Gaussian and MPS fairer as both are then based on the TI information. The use of a Gaussian covariance matrix to describe the spatial variability can partly be justified by the histogram of elevations in the data and TI matching a 1D Gaussian distribution based on their respective mean and standard deviation (Fig. 3d). The Gaussian assumption mainly breaks down for higher elevations as seen in the Q-Q plot (Fig. 3e). This problem is less significant for the TI than the data. This implies that the Paleogene clay surface could be described by a Gaussian distribution and it can be misleading to base normality tests solely on the data.

3.5 Unconditional (prior) realizations

From the prior information outlined in the previous section, we simulate a set of 50 unconditional (prior) realizations. Four resulting realizations are shown in Fig. 4 along with the E-type and Std-type (mean and standard deviation for each grid node respectively) for each prior/simulation technique.

For MPS we test two sampling strategies with different search radius to the visited node. The first scenario has a small neighborhood (search radius = 5 nodes, 500 m). The lower elevation in the unconditional realizations of the MPS simulation with small neighborhood (Fig. 4g–j) does not portray the buried valley structures outlined in the TI very well. This is in accordance with what is expected for a setup that only accounts for short range spatial correlations. The reasoning for using a small neighborhood is that it can be preferred over a larger neighborhood due to either computational efficiency or its ability to loosen up inconsistencies between prior and conditional data (Vilhelmsen et al. 2019). A large neighborhood (search radius = 30 nodes, 3000 m) that is able to represent the buried valleys in the TI is also considered. The search radius covers more or less the entire simulation and the unconditional realizations clearly rank superior to all other in terms of geological realism (Fig. 4m–p). The two neighborhood sizes

Fig. 4 Unconditional realizations from different prior distributions. Mean (E-type) and standard deviation (Std-type) of 50 realizations shown in the right. Top row: Gaussian simulations. Second row:

simulations from MPS with small neighborhood. Third row: simulations from MPS with large neighborhood. The cyan circles represent outliers

hence represent end-member cases of MPS simulation. All other algorithm parameters are identical between the two strategies. A list of the relevant parameters in the configuration of DEESSE can be found in the Appendix Table 1.

3.6 A measure for delineating outliers in the simulation results

The use of a limited neighborhood in the simulated fields is often associated with the presence of artifacts in sequential simulation methods (Meyer 2004; Mariethoz and Caers 2015; Nussbaumer et al. 2018). Considering the above, some simulation artifacts are bound to arise even though our TI is deemed a representative analog and we added rotation and scaling. In terms of geology, we expect a relatively smooth plateau in the Paleogene clay with sudden changes in elevation due to the buried valley. As long as sudden elevation changes are consistent with parts of the surrounding area, we consider the simulation in accordance with our expectations. We define an outlier as a sole point with a high contrast in elevation to the surrounding nodes. In practice, an outlier is identified by considering the eight adjacent nodes and calculating the elevation difference to the center node. If the absolute value of the contrast in elevation is higher than a given threshold t for all adjacent points, the center node is considered an outlier. An example of an outlier and the adjacent nodes is shown in Fig. 5a.

Gaussian simulation is not supposed to have simulation artifacts. Based on 50 Gaussian realizations we select the maximum t value where no outliers are identified. Based on this experiment, $t = 32$ m. Because this value is exceeded in the absolute contrast between the center node and the adjacent nodes, as shown in Fig. 5b, this simulation point is classified as an outlier. In fact any $t < 40$ would result in the node being an outlier as seen in Fig. 5b. A more sophisticated approach could also be implemented. For instance, the current approach cannot take into account two adjacent grid nodes that have a high contrast to the remaining nodes. Those two points could potentially be considered an artifact pair.

For all 50 unconditional realizations, a mean number of outliers for the MPS simulations is calculated and shown in Fig. 6. In the unconditional case, MPS simulation with the small neighborhood leads to the most outliers. Due to the lack of correlations over longer ranges, areas simulated close to each other can have vastly different elevations. When stitched together large gradients need to be obeyed in the final iterations and outliers might occur close to larger elevation contours as seen in e.g. Figure 4j. In contrast, the MPS simulation with the large neighborhood show very few (< 0.2) outliers on average. This confirms that not only is MPS with a large neighborhood able to produce the geological realism needed but also simulate the most coherent when no data is available.

Fig. 5 Example of an outlier as defined in the current work. **a** Zoom in on an outlier found in Fig. 4m. The white rectangle marks the outlier node and the yellow rectangle marks the adjacent nodes. **b** The absolute difference between the center node and the adjacent nodes from **a**

Fig. 6 Bar chart showing average number of outliers in 50 MPS realizations with a $t = 32$. Blue: small neighborhood, yellow: large neighborhood. The results are grouped by experiment. Each experiment consists of different data amount and selection strategy (see Fig. 2b, c). Ranging from using all available data points (left) to the unconditional simulations (right)

4 Conditioning to data

4.1 Handling correlated data uncertainty

The correlated data uncertainty is handled differently within each simulation method. For Gaussian simulation we solve the analytical expression with C_d as described in Eqs. 2 and 3. Handling uncertain data in MPS simulations is usually done straightforwardly by aggregating the probability of getting certain values with the conditional value found in the TI (Mariethoz and Caers 2015). In practice, for MPS simulation with discrete categories, each data point is provided a normalized probability for each category that should sum to one over all categories. Similarly, for continuous simulation, the array of possible values is discretized into a set of ranges with an attributed probability. This strategy of handling uncertain data is implemented in DEESSE. In the current setup, we would need to discretize the elevations of the interface into probability classes using a 1D Gaussian distribution for all

data points. However, it is computationally impractical, if not impossible, to discretize the spectrum of possible elevations sufficiently small enough to portray the 1D Gaussian distribution at different depths. In addition, simulation using a random path, tends to ignore uncertain information in MPS when it is sparsely available as here (Hansen et al. 2018). Finally, DEESSE, as many other MPS simulations, does not handle non-collocated uncertain data points. The latter two issues usually lead to the uncertain information being downplayed or sometimes neglected in the final realizations (Høyer et al. 2017; Hansen et al. 2018). To remedy this, we therefore suggest to draw realizations from the Gaussian data model $\mathcal{N}(\mathbf{d}_{\text{obs}}, C_d)$ using Cholesky decomposition. Subsequently, these realizations can act as already simulated (hard) data before the start of each realization. This strategy is computationally inexpensive as a realization only requires Cholesky decomposition of a $N_d \times N_d$ matrix. Furthermore, it forces data uncertainties to be Gaussian even for the MPS case. We strongly suggest the implementation of better approaches and a further

investigation of handling correlated uncertainties in MPS simulation of continuous fields.

We establish a base case using all available uncertain data as seen in Fig. 2a. The simulation results for the base case are displayed in Fig. 7 with the same configuration as in Fig. 4. The buried valley structures outlined in the data are well represented as coherent geological structures in all realizations. Small-scale variability is the major difference between the Gaussian and MPS results. The Gaussian method will naturally show variability on smaller scales, especially for a spherical type semivariogram model compared with e.g. a Gaussian type (Journel and Huijbregts 1978). The MPS simulations meanwhile are representations from the statistical properties of the TI, which show minuscule small-scale variability. Figure 8 show two profiles, one trending east-west and the other north-south. The small-scale variability is clearly seen in both profiles of the Gaussian simulations (yellow) as opposed to the MPS simulations (red and blue).

The E-type (Fig. 7) of all three strategies reveal a strikingly similar response due to the excess amount of data available. As expected, the Gaussian simulation show increasing standard deviation with distance to the data points. This trend is also clearly present in the variability between the individual realizations in Fig. 8. The pattern of lower variance for category 1 uncertainties is honored (compare Fig. 7f with Fig. 2e). The prior variance of the

uncertain data points is honored as seen by the realizations staying within two times the standard deviation from the interpretation point in Fig. 8, which is unsurprising for MPS as these points are presimulated and placed as hard data. In the MPS cases, the larger standard deviations are found along the edges of the buried valley tunnels. Because the uncertainty in MPS simulations follows structures, it probably represents the uncertainty of the geologist who provided the interpretation points better than the uncertainty estimates obtained using the Gaussian approach. Since MPS is superior to Gaussian methods at creating elongated coherent structures, this is in accordance with expectations. In contrast to the Gaussian methods, standard deviations of points near data can still be very high using the MPS methods. These data points are difficult to incorporate in the realizations. In general, we consider the STD-type more informative in the MPS case than for Gaussian. For the MPS case, the uncertainty is related to the edge of the buried valley and not whether it exists or not.

From these results it is difficult to claim one simulation strategy better than the other. If the TI is deemed our best estimate on how the geological system should behave, then the MPS rank superior due to the lack of small-scale variability. However, the TI is generated with a smooth interpolation scheme and sought-after small-scale variability is therefore not present. In terms of producing

Fig. 7 Realizations from different posterior distributions based on the *full data set* (base case). Mean (E-type) and standard deviation (Std-type) of 50 realizations shown in the right. Top row: Gaussian

simulations. Second row: simulations from MPS with small neighborhood. Third row: simulations from MPS with large neighborhood. The cyan circles represent outliers

Fig. 8 Profiles of five conditional realizations using the *full data set* (*base case*). **a** East-west profile at $Y = 20$. **b** South-north profile at $X = 21$. The thick lines represent the mean value of 50 realizations from

the base case. The black dots represent the data (large dot) and two times the standard deviation (small dots) for each point respectively

outliers (Fig. 6), the biggest change between the base case and the unconditional realizations is the reversal of the trend in outliers found in the MPS realizations. The small neighborhood now produces very few outliers because of all the large-scale patterns are already in place from the data. The lower ranges of the small neighborhood make it a very consistent prior with the data. Conversely, the prior information from the large neighborhoods are prone to produce more outliers than the small neighborhood because

of inconsistencies at longer correlation length and the big amount of data. In other words, the low entropy prior information provided in MPS with large neighborhoods is potentially difficult to match if the information is slightly inconsistent. In the current MPS setup a maximum of 15 conditional points is used at each iteration within each simulation (see “Appendix”). For a higher number, possible inconsistencies between the prior and data would become even more noticeable.

4.2 Simulations with reduced data sets

The data coverage from geological interpretation is dense as shown in Fig. 2 and the choice of geostatistical simulation technique is less important as all methods produce reasonable results in the base case (Fig. 7). Placing interpretation points in cognitive geological modeling is time-consuming. Limiting the number of data points necessary in the interpretation phase would make geological cognitive modeling more feasible for larger systems. An impact of reduced data is to be expected, but how much does limiting the number of data points affect the results? Usually, the spatial coverage of interpretation points is homogeneous as is also the case for the Horsens data. Geologists likely tend towards evenly spreading interpretation points to make sure that all areas are informed. Traditionally, the structural information between interpretation points are kriged or otherwise smoothly interpolated. Some data points may also be included to support the interpolation routine and not be directly geologically relevant per se. In the end, some interpretation points may carry more information for the final realizations than others may. In the following, we assume that only fewer data points are available and test whether developing a strategy for placing interpretation points is worthwhile. The full number of $N_d = 379$ data set is reduced to both $N_d = 60$ (Fig. 2b, c) data points and an extreme case of only $N_d = 12$ points.

Since the main geological features in Horsens are the buried valleys, an informative method of selecting data points would involve describing the placement of such features. Here, we adopt a strategy based on calculating the elevation gradient between all points. Once calculated we take out points that either have a high-gradient or as close to zero as possible. The latter is done to also select points that lie on the plateau. Because some channels are incised deeper into the Paleogene clay, some areas may be over-represented from just using the maximal gradient criteria alone. We compensate this by subdividing the simulation area into smaller sections and selecting three data points within each. Two points with the maximum positive and negative gradient and a final point closest to zero gradient. This gives us a first-order approximation for selecting important points based on the geological background knowledge and a spatial constraint. The two maximum gradient points are usually next to each other as seen in Fig. 2b. This happens because the gradients are equal in size but with opposite signs. For comparison, a random selection strategy is also adopted (Fig. 2c).

We run the same setup as in the base case for all four reduced data sets. The realizations from the 60 point data sets are shown in Figs. 9 and 10. In general, having fewer

data points introduces more variability in all solutions. Indications of two buried valleys trending E–W are seen in most realizations in the random subset of 60 data points (Fig. 9). A faint hint of a third valley is also seen in the upper right corner trending NE–SW. The most reasonable results in terms of reproducing these features come from the Gaussian method and the MPS with the large neighborhood, with the latter ranking superior in geological realism as expressed by the connectivity of the valleys. Due to the low range coverage in the MPS with a small neighborhood, the realizations look fragmented. Another useful aspect of the MPS with a large neighborhood is the increased standard deviation in areas where possible buried valleys might occur (Fig. 9r). However, from the base case, we know that there should be two buried valleys in the lower parts of the simulation grid. The results in Fig. 9q indicate the presence of only one valley and with a low standard deviation (Fig. 9r). This type of certainty of a wrong feature might occur because of small inconsistencies between prior and data in the low entropy case of MPS, differing from Gaussian simulation where the standard deviation is only low close to data points. The standard deviation of the MPS simulations with the small neighborhoods (Fig. 9l) is also the lowest around data. In contrast to the Gaussian result, the standard deviation changes more dramatically between low and high values closer to data. The smoothness in the correlations in the Gaussian solutions explains this behavior. The results from MPS are closer to the results from the unconditional realizations than the base case in terms of outliers (Fig. 6). The structures seen in the simulation and the number of outliers tend toward the results from the unconditional simulation as the data are reduced. This is another indication that perhaps the prior and data is difficult to match during simulation with fewer data. The large neighborhood performs considerably better than the small neighborhood, which is also the situation for the extremely reduced data sets not shown here.

For the experiment with the gradient-based selection of data points (Fig. 10), we see that the realizations from the MPS with the small neighborhood are not very useful since they are very fragmented and not very plausible from a geological point of view. These realizations are also extremely prone to produce outliers (Fig. 6). The mean does not provide any clear indication of longer buried valleys and the variance is low around the data points. The MPS with longer correlations again excels in producing the most geologically realistic features and the Gaussian simulation produces useful results with the caveat of small-scale variability (Fig. 11). The mean value from the base case generally fits within the uncertainties outlined by the realizations in Fig. 11, thus indicating that the base case could still be a realization of the reduced data set. The deep parts of the buried valley (around $Y = 40$ in Fig. 11b) are

Fig. 9 Realizations from different posterior distributions based on the reduced data set of 60 points based on random selection (Fig. 2c). Mean (E-type) and standard deviation (Std-type) of 50 realizations shown in the right. Top row: Gaussian simulations. Second row:

simulations from MPS with small neighborhood. Third row: simulations from MPS with large neighborhood. The cyan circles represent outliers

Fig. 10 Realizations from different points posterior distributions based on the reduced data set of 60 points based on gradient selection (Fig. 2b). Mean (E-type) and standard deviation (Std-type) of 50 realizations shown in the right. Top row: Gaussian simulations.

Second row: simulations from MPS with small neighborhood. Third row: simulations from MPS with large neighborhood. The cyan circles represent outliers

however difficult to obtain for all methods, especially the MPS simulation with a large neighborhood (blue).

Comparing the E-types in Fig. 10 to the random based point selection in Fig. 9, we see that the buried valleys now more clearly trends NE–SW and NW–SE as we know from the base case to be the preferred orientation of the buried valley in the data. The spatial distribution of the standard deviation of the MPS solution with a large neighborhood is also more in accordance with our expectations from the base case. This indicates that choosing points in the simulation grid for interpretation does matter if time only

allows a reduced number of interpretations to be made. For MPS, the number of outliers, in this case, are considerably higher than for any other experiment (Fig. 6).

When reducing the number of data even further (12 points) the aforementioned trends continue and the importance of having high information data becomes amplified. For a random strategy, having only 12 points to choose from, makes it completely plausible that all selected points are plateau values. In such a case, no information of buried valleys are present in the interpretation data and must come from prior information alone.

Fig. 11 Profiles of five conditional realizations using the 60 points based on gradient selection (Fig. 2b). **a** East-west profile at Y = 20. **b** South-north profile at X = 21. The thick lines represent the mean

value of 50 realizations from the base case. The black dots represent the data (large dot) and two times the standard deviation (small dot) for each point respectively

The number of outliers for MPS is still relatively large for both reduced data sets with 12 points (Fig. 6). As opposed to the reduced data set with 60 points, the gradient-based method does not provide more outliers than the random selection for the MPS with a small neighborhood. A possible explanation could be the emphasis on the information from the prior when the number of data is limited. For the large neighborhood, the average number of outliers is considerably higher (0.4) than the unconditional realizations (0.14). Once again, a reminder that the way we simulate from the training image is prone to produce artifact/outliers when combined with data even when allowing for rotation and scaling of the values.

5 Inferring the optimal prior from the data

In the following, we test the possibility of inferring the Gaussian prior (C_M) from the interpretation data as described in Sect. 3.2. C_d is already obtained using Eq. 1 in the initial phase of setting up the Gaussian problem with the base case (Fig. 2g). We sweep through different ranges (2000 to 4500 m) and sills (1000 to 10,000 m²) in C_M and calculate the corresponding likelihood from Eq. 6 (Fig. 12a). We scale the likelihood $\log(L)$ with the maximum likelihood found to normalize the results. A value of 1 hence represent the most likely configuration of sill and range given the spherical model in the prior model. For the case of using the C_d from the base case, we find an optimal range $r_0 = 3290$ m and an optimal sill $s_0 = 3000$ m². The precision of these results will vary slightly depending on the sweep step length. Both of these values are larger than the inferred values using semivariogram modeling on the TI in Fig. 3c. The most interesting aspect of Fig. 12a is that the maximum is isolated with a steep gradient in the likelihood away from the maximum. This provides some assurance of robustness in inferring the properties of the prior directly from the data without the need for a TI.

Finally, we also assess the sensitivity of the inferred model to the number of data. We infer the range and sill of the prior with the two reduced data sets from random selection and their corresponding C_d . The results are displayed in Fig. 12b, c. As the number of data is reduced the span of likely configurations of sill and range expand. With only 12 data available (Fig. 12c) almost all ranges and sills are possible. Despite this, the optimal value is surprisingly stable across the three experiments. As a final remark, note that in all three described scenarios, the optimal sill is estimated below the variance of the data (compare the cyan dot to the cyan line in Fig. 12) in accordance with the result of Barnes (1991). This is probably because the combined variance is the sum of structural variance and the variance related to the uncertainty on the measurement points.

Overall, the findings in this section suggest that it is possible to infer a suitable prior model if it is assumed Gaussian. This is especially true when the number of data is high as is the case with the interpretation points. Ideally, one could make use of an anisotropic covariance. This would require the estimation of two additional parameters, namely angle and anisotropy length.

6 Discussion

The initial objective of the study was to suggest a complete framework for integrating uncertain interpretation points within MPS and Gaussian simulation and evaluate the steps needed in each case. In the presented experiments, MPS realizations showed greater connectivity of buried valleys than the Gaussian simulations. MPS also showed more informative uncertainty estimates and better handling of the reduced data sets in accordance with the results of Mariethoz (2018).

The current study suggests that if enough data points are available, the choice of modeling strategy becomes secondary, which is also in accordance with the results of Mariethoz (2018) for a moderately informed case. All realizations of the base case are adequate representations of the subsurface independent of the geostatistical simulation technique. Using MPS with a large neighborhood might be difficult to match with a large data set due to the high risk of inconsistency between prior and data. Roughly speaking, when the number of data increases, a smaller neighborhood becomes more attractive. The main issue is to establish the best way to represent the prior information so that it fulfills two criteria: (1) that it should be a fair representation that covers longer correlation ranges in the training image and (2) that the two types of information are consistent in case of large data amounts. We encourage that this continues to be an active source of research for years to come. Some remediation is already suggested in the literature by post-processing realizations to clear up inconsistencies and artifacts in the simulation (e.g. Meerschman et al. 2013).

Up until this point in the discussion, it may seem that the Gaussian approach is only as good as or in most cases a worse option than MPS modeling depending on the amount of data available. Is it conceivable that a Gaussian solution could be preferential to MPS? In the current base case setup, generating a conditional MPS simulation takes approximately 3 s on a normal laptop. In comparison, a realization of the Cholesky decomposition approach takes a few milliseconds. This speed gain would however drastically reduce for a larger system if SGS is needed instead of Cholesky decomposition. The size of the simulation neighborhood would play a significant role in both the accuracy and efficiency of both MPS and SGS (Hansen and

Fig. 12 Inference of range and sill in a Gaussian covariance model directly from data. **a** Using C_d from the base case, **b** using the randomly reduced data set of 60 data, and **c** using the randomly reduced data set of 12 data

Mosegaard 2008). For example, Journel and Zhang (2006) provided a case where MPS produced similar results as SGS at advantageous speed. However, we argue that in most cases SGS would still be faster than MPS as computation times depends highly on the size of the TI in MPS.

As stated in the introduction, one major disadvantage (and strength) of MPS is the need for prior information to be quantified through a TI. Finding a suitable candidate TI is non-trivial. This could be based on existing nearby data (as in our case and Kessler et al. (2013)), a more conceptual idea from expert knowledge (e.g. Zunino et al. 2015) or something perhaps originating from a bank of training images (e.g. Pyrcz et al. 2008). There is no clear-cut strategy in choosing such a TI as it involves subjective decisions. However, if multiple candidate TIs are available, recent work focused on objective selection strategies has

shown promising results (Park et al. 2013; Hermans et al. 2015; Juda et al. 2020).

Conditional simulation showing signs of inconsistency between prior and data is perhaps the strongest indication of how well the TI is suited. In the study, we see some indication that a large neighborhood and hence longer correlations in the MPS prior leads to more inconsistencies when the amount of data is increased. This suggest that the low-entropy MPS model may be ‘too informative’ and in practice inconsistent with actual observations. One way to remedy this is to use much larger training images that capture more variability. But, as discussed, it is most often not feasible to obtain such a large and representative TI.

The main strength of using MPS with TIs, the low-entropy properties allowing simulation of complex spatial structures, is then also a potential drawback, leading to

inconsistencies when combined with other data as considered here. For Gaussian models, the case is just the opposite. The Gaussian models may not be able to represent as realistic spatial structures due to its high-entropy nature but rarely leads to inconsistencies when combined with other information.

In addition, while it is non-trivial to establish a representative TI of a useful size, our results indicate that at least for the proposed setup it is possible to reasonably infer the properties of the Gaussian prior model directly from the data. This is a great advantage of using Gaussian simulation as a priori expert information to some extent become superfluous. An implication of this is the possibility to automate the prior generation. Cognitive geological models consist of multiple (typically between 10–15) layer boundaries. Consider that all interfaces in the cognitive structural model needed simulation. In MPS this would require coming up with a suitable TI and parameterization for each interface. Even with the use of cross-validation, this is a daunting task. Currently, the best option would probably be to use an objective selection strategy to an ensemble of possible TIs each representing interpreted interfaces from nearby geology. Using the strategy of inferring the optimal Gaussian prior from the data on the contrary could be set up to automatically process interpretation points from each interface in a relatively efficient manner and because cognitive models typically consist of many interpretation points for each layer, the high-entropy Gaussian model might suffice.

In principle, a variety of different parameters could be included as free parameters in determining the best possible prior and data distribution, such as the range of the data, the standard deviation of the data, the semivariogram type, etc. A likely consequence is additional computation time from expanding the number of values to sweep. To remedy this, an extended Metropolis-Hastings algorithm could instead be used to infer the free parameters (Mosegaard and Tarantola 1995). Because the absolute likelihood in Eq. 6 depends on the number of model parameters (N), we normalized the likelihood for comparability between three cases presented in Sect. 5. In general, we suggest normalizing the results for easier comparisons between covariance types even if the number of model parameters stays fixed. As stressed earlier, even with the Gaussian method having some advantages over MPS, a relatively large amount of data is still needed for these advantages to be present.

It is important to bear in mind that this is just one scenario with uncertain interpretation points in a continuous field, in which the responses from Gaussian simulation and MPS are tested. It is also perfectly reasonable to believe that a better training image more suitable for the simulation problem could potentially be found. Although, we argue that given the proportions of the proposed TI to the

simulation area and that the TI describes the exact geology we interested in, makes the job of finding a better-suited TI difficult. At the very least not a real-world TI.

Finally, the improved results from using a gradient-based selection method suggest that not only is the number of data points available, but also the information content, important for selecting the optimal simulation strategy. The gradient-based selection also showed that inconsistencies might be enhanced even though the final solutions are ultimately closer to what we intend to see from a geological point of view. We, therefore, encourage future work on how geologists pick points that maximize the information level in the system. An interpretation strategy that maximizes information content while maintaining consistency with prior information would lead to a more efficient interpretation process as fewer data would be necessary. Furthermore, it might help to bridge the gap between geologists and geostatisticians as consideration regarding the final geostatistical model becomes part of the interpretation routine. We suspect that the lowered number of data points would likely be mostly of benefit for the low-entropy MPS simulations as was the case in our results.

7 Conclusion

The present research aimed to examine the effect of choosing different simulation strategies for a case with uncertain data points from a cognitive geological model. Specifically, we investigated Gaussian simulation vs. MPS simulation. Our results suggest that MPS simulation with many conditional data and a large neighborhood might not be reasonable due to consistency problems, whereas a higher entropy solution such as Gaussian or MPS with a small neighborhood is favorable. With fewer data points available, MPS simulation rank superior as long the training image provides enough variability such that the uncertainties are not underestimated. The opportunity of inferring a Gaussian prior directly from the data exists in the Gaussian setup. The Gaussian methods do not produce outliers as defined in the current study. Based on these features, we tend to favor Gaussian simulations for the present case, as setting up a semi-automatic procedure for all interfaces is less time-consuming than for MPS. Results would be adequate due to the high number of data available from cognitive modeling. Further research might explore how to make cognitive geological modeling more efficient based on the information content of interpretation points. This can assist in developing interpretation-targeted geostatistical simulation tools that guide the interpreter towards regions that need information.

In conclusion, this study suggests that using a specific simulation strategy for all geostatistical problems might be

at best redundant and at worst inferior to other methods. One should make use of the geostatistical method providing an information level consistent with the data. The number of data and its corresponding information content plays a pivotal role in determining this level.

Acknowledgements The authors are grateful to GEUS for making the interpretation data available and covering open access production costs. We also thank the Independent Research Fund Denmark (Grant Number 7017-00160B) for supporting the current research financially. We thank two anonymous reviewers for thorough and constructive feedback during the revision of the manuscript.

Funding This research is made possible by the Independent Research Fund Denmark (Grant Number 7017-00160B).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

Appendix

See Table 1.

Table 1 Relevant parameters in the DEESSE setup

Parameter	Value
Rotation	360°
Consistency of conditioning data (scaling)	0.1
Normalization type	Linear
Maximal number of neighboring nodes	15
Maximal density of neighboring nodes	1
Distance type	L2-norm
Distance threshold	0.01
Maximal scan fraction of TI	1
Starting seed + (seed increment)	1 (1)
Simulation path	Random
Post-processing	No

References

- Alabert F (1987) The practice of fast conditional simulations through the LU decomposition of the covariance matrix. *Math Geol* 19:369–386. <https://doi.org/10.1007/BF00897191>
- Andersen TR, Poulsen SE, Christensen S, Jørgensen F (2013) A synthetic study of geophysics-based modelling of groundwater flow in catchments with a buried valley. *Hydrogeol J* 21:491–503. <https://doi.org/10.1007/s10040-012-0924-5>
- Arvidsen AG, Andersen TB, Nielsen OF et al (2020) Samling af geologiske modeller i Jylland: FOHM - Fælles Offentlig Hydrologisk Model (In Danish). MST-Miljøstyrelsen, Miljøministeriet
- Barfod AS, Møller I, Christiansen AV et al (2018) Hydrostratigraphic modeling using multiple-point statistics and airborne transient electromagnetic methods. *Hydrol Earth Syst Sci* 22:3351–3373. <https://doi.org/10.5194/hess-22-3351-2018>
- Barnes RJ (1991) The variogram sill and the sample variance. *Math Geol* 23:673–678. <https://doi.org/10.1007/BF02065813>
- Bond CE (2015) Uncertainty in structural interpretation: lessons to be learnt. *J Struct Geol* 74:185–200. <https://doi.org/10.1016/j.jsg.2015.03.003>
- Bosch JHA, Bakker MAJ, Gunnink JL, Paap BF (2009) Airborne electromagnetic measurements as basis for a 3D geological model of an Elsterian incision. *Zeitschrift der Dtsch Gesellschaft für Geowissenschaften* 160:249–258. <https://doi.org/10.1127/1860-1804/2009/0160-0258>
- Chilès J-P, Delfiner P (2012) *Geostatistics*, 2nd edn. Wiley, Hoboken
- Danielsen JE, Auken E, Jørgensen F et al (2003) The application of the transient electromagnetic method in hydrogeophysical surveys. *J Appl Geophys* 53:181–198. <https://doi.org/10.1016/j.jappgeo.2003.08.004>
- De Figueiredo LP, Grana D, Roisenberg M, Rodrigues BB (2019) Gaussian mixture Markov chain Monte Carlo method for linear seismic inversion. *Geophysics* 84:R463–R476. <https://doi.org/10.1190/geo2018-0529.1>
- Emery X, Lantuéjoul C (2014) Can a training image be a substitute for a random field model? *Math Geosci* 46:133–147. <https://doi.org/10.1007/s11004-013-9492-z>
- Emery X, Peláez M (2011) Assessing the accuracy of sequential Gaussian simulation and cosimulation. *Comput Geosci* 15:673–689. <https://doi.org/10.1007/s10596-011-9235-5>
- Gómez-Hernández JJ, Cassiraga EF (1994) Theory and practice of sequential simulation. In: Armstrong M, Dowd PA (eds) *Geostatistical simulations. Quantitative geology and geostatistics*, vol 7. Springer, Dordrecht, pp 111–124
- Goovaerts P (1997) *Geostatistics for natural resources evaluation*, applied geostatistics series. Oxford University Press, Oxford
- Grana D, Fjeldstad T, Omre H (2017) Bayesian Gaussian mixture linear inversion for geophysical inverse problems. *Math Geosci* 49:1–23. <https://doi.org/10.1007/s11004-016-9671-9>
- Gravey M, Mariethoz G (2020) QuickSampling v1.0: a robust and simplified pixel-based multiple-point simulation approach. *Geosci Model Dev* 13:2611–2630. <https://doi.org/10.5194/gmd-13-2611-2020>
- Guardiano FB, Srivastava RM (1993) Multivariate geostatistics: beyond bivariate moments. In: Soares A (ed) *Geostatistics Tróia '92. Quantitative geology and geostatistics*, 5th edn. Springer, Dordrecht, pp 133–144
- Hansen TM (2020) Entropy and information content of geostatistical models. *Math Geosci*. <https://doi.org/10.1007/s11004-020-09876-z>
- Hansen TM, Mosegaard K (2008) VISIM: sequential simulation for linear inverse problems. *Comput Geosci* 34:53–76. <https://doi.org/10.1016/j.cageo.2007.02.003>

- Hansen M, Pjetursson B (2011) Free, online Danish shallow geological data. *Geol Surv Denmark Greenl Bull* 23:53–56. <https://doi.org/10.34194/geusb.v23.4842>
- Hansen TM, Cordua KS, Mosegaard K (2015) A general probabilistic approach for inference of Gaussian model parameters from noisy data of point and volume support. *Math Geosci* 47:843–865. <https://doi.org/10.1007/s11004-014-9567-5>
- Hansen TM, Vu LT, Bach T (2016) 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>
- Hansen TM, Vu LT, Mosegaard K, Cordua KS (2018) Multiple point statistical simulation using uncertain (soft) conditional data. *Comput Geosci* 114:1–10. <https://doi.org/10.1016/j.cageo.2018.01.017>
- Hermans T, Nguyen F, Caers J (2015) Uncertainty in training image-based inversion of hydraulic head data constrained to ERT data: workflow and case study. *Water Resour Res* 51:5332–5352. <https://doi.org/10.1002/2014WR016460>
- Higham NJ (2009) Cholesky factorization. *Wiley Interdiscip Rev Comput Stat* 1:251–254. <https://doi.org/10.1002/wics.018>
- Hoffmann J, Scheidt C, Barfod AS, Caers J (2017) Stochastic simulation by image quilting of process-based geological models. *Comput Geosci* 106:18–32. <https://doi.org/10.1016/j.cageo.2017.05.012>
- Høyer A-S, Jørgensen F, Sandersen PBE et al (2015) 3D geological modelling of a complex buried-valley network delineated from borehole and AEM data. *J Appl Geophys* 122:94–102. <https://doi.org/10.1016/j.jappgeo.2015.09.004>
- Høyer A-S, Vignoli G, Hansen TM et al (2017) Multiple-point statistical simulation for hydrogeological models: 3-D training image development and conditioning strategies. *Hydrol Earth Syst Sci* 21:6069–6089. <https://doi.org/10.5194/hess-21-6069-2017>
- Høyer A-S, Klint KES, Fiandaca G et al (2019) Development of a high-resolution 3D geological model for landfill leachate risk assessment. *Eng Geol* 249:45–59. <https://doi.org/10.1016/j.enggeo.2018.12.015>
- Huuse M, Lykke-Andersen H (2000) Overdeepened quaternary valleys in the eastern Danish North Sea: morphology and origin. *Quat Sci Rev* 19:1233–1253. [https://doi.org/10.1016/S0277-3791\(99\)00103-1](https://doi.org/10.1016/S0277-3791(99)00103-1)
- Jørgensen F, Sandersen PBE (2006) Buried and open tunnel valleys in Denmark-erosion beneath multiple ice sheets. *Quat Sci Rev* 25:1339–1363. <https://doi.org/10.1016/j.quascirev.2005.11.006>
- Jørgensen F, Møller RR, Sandersen PBE, Nebel L (2010) 3-D geological modelling of the Egebjerg Area, Denmark, based on hydrogeophysical data. *Geol Surv Denmark Greenl Bull* 20:27–30. <https://doi.org/10.34194/geusb.v20.4892>
- Jørgensen F, Møller RR, Nebel L et al (2013) A method for cognitive 3D geological voxel modelling of AEM data. *Bull Eng Geol Environ* 72:421–432. <https://doi.org/10.1007/s10064-013-0487-2>
- Jørgensen F, Høyer A-S, Sandersen PBE et al (2015) Combining 3D geological modelling techniques to address variations in geology, data type and density—an example from Southern Denmark. *Comput Geosci* 81:53–63. <https://doi.org/10.1016/j.cageo.2015.04.010>
- Journel AG (2002) Combining knowledge from diverse sources: an alternative to traditional data independence hypotheses. *Math Geol* 34:573–596. <https://doi.org/10.1023/A:1016047012594>
- Journel AG, Huijbregts CJ (1978) Mining geostatistics, 1st edn. Academic Press Inc., London
- Journel A, Zhang T (2006) The necessity of a multiple-point prior model. *Math Geol* 38:591–610. <https://doi.org/10.1007/s11004-006-9031-2>
- Juda P, Renard P, Straubhaar J (2020) A framework for the cross-validation of categorical geostatistical simulations. *Earth Sp Sci* 7:1–17. <https://doi.org/10.1029/2020EA001152>
- Kessler TC, Comunian A, Oriani F et al (2013) Modeling fine-scale geological heterogeneity-examples of sand lenses in tills. *Groundwater* 51:692–705. <https://doi.org/10.1111/j.1745-6584.2012.01015.x>
- Laloy E, Héroult R, Lee J et al (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. <https://doi.org/10.1016/j.advwatres.2017.09.029>
- Lantuéjoul C (2002) Geostatistical simulation. Springer, Berlin
- Li L, Romary T, Caers J (2015) Universal kriging with training images. *Spat Stat* 14:240–268. <https://doi.org/10.1016/j.spasta.2015.04.004>
- Madsen RB, Hansen TM, Omre H (2020a) Estimation of a non-stationary prior covariance from seismic data. *Geophys Prospect* 68:393–410. <https://doi.org/10.1111/1365-2478.12848>
- Madsen RB, Kim H, Kallesøe AJ et al (2020b) 3D multiple point geostatistical simulation of joint subsurface redox and geological architectures. *Hydrol Earth Syst Sci Discuss.* <https://doi.org/10.5194/hess-2020-444> (in review)
- Mariethoz G (2018) When should we use multiple-point geostatistics? In: Daya Sagar B, Cheng Q, Agterberg F (eds) Handbook of mathematical geosciences: fifty years of IAMG. Springer, Cham, pp 645–653
- Mariethoz G, Caers J (2015) Multiple-point geostatistics: stochastic modeling with training images, 1st edn. Wiley, New York
- Mariethoz G, Lefebvre S (2014) Bridges between multiple-point geostatistics and texture synthesis: review and guidelines for future research. *Comput Geosci* 66:66–80. <https://doi.org/10.1016/j.cageo.2014.01.001>
- Mariethoz G, Renard P, Straubhaar J (2010) The direct sampling method to perform multiple-point geostatistical simulations. *Water Resour Res* 46:1–14. <https://doi.org/10.1029/2008WR007621>
- Meerschman E, Pirot G, Mariethoz G et al (2013) A practical guide to performing multiple-point statistical simulations with the direct sampling algorithm. *Comput Geosci* 52:307–324. <https://doi.org/10.1016/j.cageo.2012.09.019>
- Meyer TH (2004) The discontinuous nature of kriging interpolation for digital terrain modeling. *Cartogr Geogr Inf Sci* 31:209–216. <https://doi.org/10.1559/1523040042742385>
- Møller RR, Jørgensen F (2011) GEUS Rapport 2011/37: 3D geologisk model for Egebjerg. Danmarks og Grønlands Geologiske Undersøgelse, Copenhagen
- Møller I, Søndergaard VH, Jørgensen F (2009) Geophysical methods and data administration in Danish groundwater mapping. *Geol Surv Denmark Greenl Bull* 17:41–44. <https://doi.org/10.34194/geusb.v17.5010>
- Mosegaard K, Tarantola A (1995) Monte Carlo sampling of solutions to inverse problems. *J Geophys Res* 100:431–447. <https://doi.org/10.1029/94JB03097>
- Nussbaumer R, Mariethoz G, Gloaguen E, Holliger K (2018) Which path to choose in sequential Gaussian simulation. *Math Geosci* 50:97–120. <https://doi.org/10.1007/s11004-017-9699-5>
- Oldenborger GA, Logan CE, Hinton MJ et al (2014) 3D hydrogeological model building using airborne electromagnetic data. In: Near surface geoscience 2014—20th European meeting of environmental and engineering geophysics, pp 14–18. <https://doi.org/10.3997/2214-4609.20142011>
- Omre H (1987) Bayesian kriging—merging observations and qualified guesses in kriging. *Math Geol* 19:25–39. <https://doi.org/10.1007/BF01275432>
- Park H, Scheidt C, Fenwick D et al (2013) History matching and uncertainty quantification of facies models with multiple

- geological interpretations. *Comput Geosci* 17:609–621. <https://doi.org/10.1007/s10596-013-9343-5>
- Pyrz MJ, Boisvert JB, Deutsch CV (2008) A library of training images for fluvial and deepwater reservoirs and associated code. *Comput Geosci* 34:542–560. <https://doi.org/10.1016/j.cageo.2007.05.015>
- Refsgaard JC, Christensen S, Sonnenborg TO et al (2012) Review of strategies for handling geological uncertainty in groundwater flow and transport modeling. *Adv Water Resour* 36:36–50. <https://doi.org/10.1016/j.advwatres.2011.04.006>
- Remy N, Boucher A, Wu J (2009) *Applied geostatistics with SGeMS*, 1st edn. Cambridge University Press, Cambridge
- Rezaee H, Mariethoz G, Koneshloo M, Asghari O (2013) Multiple-point geostatistical simulation using the bunch-pasting direct sampling method. *Comput Geosci* 54:293–308. <https://doi.org/10.1016/j.cageo.2013.01.020>
- Royse KR (2010) Combining numerical and cognitive 3D modelling approaches in order to determine the structure of the Chalk in the London Basin. *Comput Geosci* 36:500–511. <https://doi.org/10.1016/j.cageo.2009.10.001>
- Sabeti H, Moradzadeh A, Ardejani FD et al (2017) Geostatistical seismic inversion for non-stationary patterns using direct sequential simulation and co-simulation. *Geophys Prospect* 65:25–48. <https://doi.org/10.1111/1365-2478.12502>
- Sandersen PBE, Jørgensen F (2012) Substratum control on tunnel-valley formation in Denmark. *Geol Soc Spec Publ* 368:145–157. <https://doi.org/10.1144/SP368.12>
- Sandersen PBE, Jørgensen F (2017) Buried tunnel valleys in Denmark and their impact on the geological architecture of the subsurface. *Geol Surv Denmark Greenl Bull* 38:13–16. <https://doi.org/10.34194/geusb.v38.4388>
- Sandersen PBE, Jørgensen F, Larsen NK et al (2009) Rapid tunnel-valley formation beneath the receding Late Weichselian ice sheet in Vendsyssel, Denmark. *Boreas* 38:834–851. <https://doi.org/10.1111/j.1502-3885.2009.00105.x>
- Schamper C, Jørgensen F, Auken E, Effersø F (2014) Assessment of near-surface mapping capabilities by airborne transient electromagnetic data—an extensive comparison to conventional borehole data. *Geophysics* 79:B187–B199. <https://doi.org/10.1190/geo2013-0256.1>
- Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27:623–656. <https://doi.org/10.1002/j.1538-7305.1948.tb00917.x>
- Stewart M (2009) 3D seismic analysis of pleistocene tunnel valleys in the Central North Sea. Dissertation, Imperial College London
- Straubhaar J (2019) *DeeSse User 's guide*. the centre for hydrogeology and geothermics (CHYN), edited by University of Neuchâtel, Neuchâtel
- Straubhaar J, Renard P, Mariethoz G et al (2011) An improved parallel multiple-point algorithm using a list approach. *Math Geosci* 43:305–328. <https://doi.org/10.1007/s11004-011-9328-7>
- Strebelle S (2002) Conditional simulation of complex geological structures using multiple-point statistics. *Math Geol* 34:1–21. <https://doi.org/10.1109/CEC.2011.5949612>
- Strebelle S (2012) Multiple-point geostatistics: from theory to practice. In: Expanded abstract collection from ninth international geostatistics congress, pp 1–65
- Tahmasebi P (2018) Multiple point statistics: a review. In: Daya Sagar B, Cheng Q, Agterberg F (eds) *Handbook of mathematical geosciences*. Springer, Cham, pp 613–643
- Tahmasebi P, Hezarkhani A, Sahimi M (2012) Multiple-point geostatistical modeling based on the cross-correlation functions. *Comput Geosci* 16:779–797. <https://doi.org/10.1007/s10596-012-9287-1>
- Tarantola A, Valette B (1982) Inverse problems = quest for information. *J Geophys* 50:159–170
- Thomsen R, Søndergaard VH, Sørensen KI (2004) Hydrogeological mapping as a basis for establishing site-specific groundwater protection zones in Denmark. *Hydrogeol J* 12:550–562. <https://doi.org/10.1007/s10040-004-0345-1>
- Vilhelmsen TN, Auken E, Christiansen AV et al (2019) Combining clustering methods with MPS to estimate structural uncertainty for hydrological models. *Front Earth Sci* 7:1–15. <https://doi.org/10.3389/feart.2019.00181>
- Wellmann F, Caumon G (2018) *3-D Structural geological models: concepts, methods, and uncertainties*, 1st edn. Elsevier Inc., Amsterdam
- Wycisk P, Hubert T, Gossel W, Neumann C (2009) High-resolution 3D spatial modelling of complex geological structures for an environmental risk assessment of abundant mining and industrial megasites. *Comput Geosci* 35:165–182. <https://doi.org/10.1016/j.cageo.2007.09.001>
- Zunino A, Mosegaard K (2019) An efficient method to solve large linearizable inverse problems under Gaussian and separability assumptions. *Comput Geosci* 122:77–86. <https://doi.org/10.1016/j.cageo.2018.09.005>
- Zunino A, Mosegaard K, Lange K et al (2015) Monte Carlo reservoir analysis combining seismic reflection data and informed priors. *Geophysics* 80:R31–R41. <https://doi.org/10.1190/geo2014-0052.1>

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