

Article

Accounting for Local Geological Variability in Sequential Simulations—Concept and Application

Adrian Linsel ^{1,*} , Sebastian Wiesler ¹, Joshua Haas ¹, Kristian Bär ²  and Matthias Hinderer ¹

¹ Department of Applied Sedimentary Geology, Institute of Applied Geosciences, Technische Universität Darmstadt, 64287 Darmstadt, Germany; Sebastian.Wiesler@hotmail.de (S.W.); joshuaahaas@gmx.de (J.H.); hinderer@geo.tu-darmstadt.de (M.H.)

² Department of Applied Geothermal Science and Technology, Institute of Applied Geosciences, Technische Universität Darmstadt, 64287 Darmstadt, Germany; baer@geo.tu-darmstadt.de

* Correspondence: linsel@geo.tu-darmstadt.de

Received: 11 May 2020; Accepted: 23 June 2020; Published: 26 June 2020



Abstract: Heterogeneity-preserving property models of subsurface regions are commonly constructed by means of sequential simulations. Sequential Gaussian simulation (SGS) and direct sequential simulation (DSS) draw values from a local probability density function that is described by the simple kriging estimate and the local simple kriging variance at unsampled locations. The local simple kriging variance, however, does not necessarily reflect the geological variability being present at subsets of the target domain. In order to address that issue, we propose a new workflow that implements two modified versions of the popular SGS and DSS algorithms. Both modifications, namely, LVM-DSS and LVM-SGS, aim at simulating values by means of introducing a local variance model (LVM). The LVM is a measurement-constrained and geology-driven global representation of the locally observable variance of a property. The proposed modified algorithms construct the local probability density function with the LVM instead of using the simple kriging variance, while still using the simple kriging estimate as the best linear unbiased estimator. In an outcrop analog study, we can demonstrate that the local simple kriging variance in sequential simulations tends to underestimate the locally observed geological variability in the target domain and certainly does not account for the spatial distribution of the geological heterogeneity. The proposed simulation algorithms reproduce the global histogram, the global heterogeneity, and the considered variogram model in the range of ergodic fluctuations. LVM-SGS outperforms the other algorithms regarding the reproduction of the variogram model. While DSS and SGS generate a randomly distributed heterogeneity, the modified algorithms reproduce a geologically reasonable spatial distribution of heterogeneity instead. The new workflow allows for the integration of continuous geological trends into sequential simulations rather than using class-based approaches such as the indicator simulation technique.

Keywords: sequential simulation; local variance model; geological heterogeneity; uncertainty estimation; subset variability

1. Introduction

Drawing conclusions from uncertain data in Earth sciences is rather usual than unusual. Each measurement in geoscientific studies is affected by measurement errors and represents only a subset of the natural variability of geological media. The natural variability is a substantial business-critical controlling factor of different types of subsurface utilization such as mining, hydrocarbon and geothermal exploitation, carbon capture and storage, or nuclear waste disposal. The physical variability of rocks is defined as the complexity or heterogeneity of a system within time

and space [1]. Even marginal discrepancies from the predicted property distributions in the subsurface can lead to inaccurate simulations of a quarry's production potential or a reservoir's recovery and life-time [2,3]. Especially the small-scale variability of rock physical properties makes field-sized predictions still challenging.

Natural heterogeneity and the corresponding property distribution in time and space can be modeled through interpolation, statistical regression, machine learning or stochastic simulation [4,5] by using a number of observations or training data. Due to technical, economic or temporal limitations, geoscientific sampling campaigns practically always end up in scarce data sets within a target domain Ω . Accordingly, estimates of properties often do not account for or misfit the observed geological structures in the field and especially conventional interpolation techniques such as kriging produce smooth transitions at sharp geological boundaries. Moreover, they may fail to reproduce the global statistics appropriately. Conventional interpolations tend to underestimate the presence of values in the upper tail of a distribution and likewise in the lower tail, too [5]. Consequently, major geological heterogeneities, such as faults, major bounding surfaces, or physicochemical anomalies, are very likely not to be reproduced appropriately by a continuous random function (RF) [6].

In contrast to conventional interpolation techniques, stochastic simulations aim to reproduce the variance and the histogram observed in the global data [7,8]. Based on either being constrained or not, stochastic simulations split up into unconditional and conditional simulations [9]. Unconditional Monte Carlo-based simulations reproduce the original histogram without spatial constraints. The realizations produced by those methods, however, are regularly far away from representing the true spatial distribution and constitute “most likely” cases at the best. Conditional simulations, in contrast, aim to reproduce the original property distribution by means of discretely sampled points together with spatial characteristics such as the observed variogram model [10].

One type of conventional simulation algorithms is represented by the sequential Gaussian simulation (SGS) in which the local variability is simulated by sampling the local probability density function (PDF) derived from the local simple kriging variance σ_{SK}^2 . This parameter results from the previously performed interpolation of the standard normally distributed data set [11]. Early field studies have proven the potential of this method to predict rock properties at unknown locations and to assess the uncertainty that can be expected in the area of interest [12–15]. More recent approaches lead to modifications of the SGS algorithm without the need to transform the original variable into standard normal space. That technique—better known as direct sequential simulations (DSS)—may, for example, sample from the global histogram rather than from the local PDF [9] or perform a quantile-quantile back-transformation into the original variable's space after the simulation. Those approaches can reproduce both the original histogram and experimental semivariogram model as well [10]. The local PDF derived from σ_{SK}^2 , however, mainly reflects the degree of uncertainty induced by the interpolation method itself and does not necessarily reflect the local variability observed on a smaller scale than Ω .

In order to enhance the accuracy of sequential simulations, we propose a new workflow, which incorporates the local variability derived from measurements on a subset of Ω into SGS and DSS under the consideration of measurement errors. The modified SGS and DSS algorithms utilize a global representation of the locally observable variance, named local variance model (LVM), in order to draw a value at an unsampled location. Accordingly, the algorithms are called LVM-SGS and LVM-DSS. Before simulation, an integer programming optimization analysis is performed in order to optimize the robustness of the underlying interpolation function. Instead of sampling from the local PDF, which is generated by means of σ_{SK}^2 , or by solving a global optimization problem, our parametric approach simulates a local PDF based on a measurement-constrained and geology-driven variance extracted from the LVM. The local PDF hereby is simulated with a Box–Muller transform [16].

The method was tested and validated in a case study, which has been conducted in a potential geothermal reservoir formation in southwestern Germany. Therefore, we measured the intrinsic permeability, representing a key parameter in many types of subsurface utilization, on a set of samples taken from an active quarry. Ω is represented by a 3-D outcrop model, which is constructed by means

of photogrammetric outcrop wall reconstruction. The model covers a volume of 9000 m³. Small-scale variability is derived from rock samples, which are taken from two representative rock cubes. Those are regarded as Ω_b and cover a volume of 0.0156 m³ and 0.008 m³, respectively. The rock cubes are taken from the same outcrop, from which the global samples are taken from. Eventually, our approach is compared to the conventional SGS and DSS algorithms and assessed by its ability to reproduce the global variogram model and the geological heterogeneity.

2. Theoretical Background

2.1. Spatial Variability

In order to reduce the probability of economic failure in mining industries, the concept of the regionalized variable had been developed by Matheron [17] in the 1960s. The regionalized variable is a function that takes a definite value at each point of space. In geological media that regionalized variable often proved to be too complex to be expressed by mathematical functions. A regionalized variable is assumed to show a more or less steady continuity in space accompanied by local fluctuations (Figure 1). In geological media, those fluctuations usually result from the physical variability observed at smaller scales.

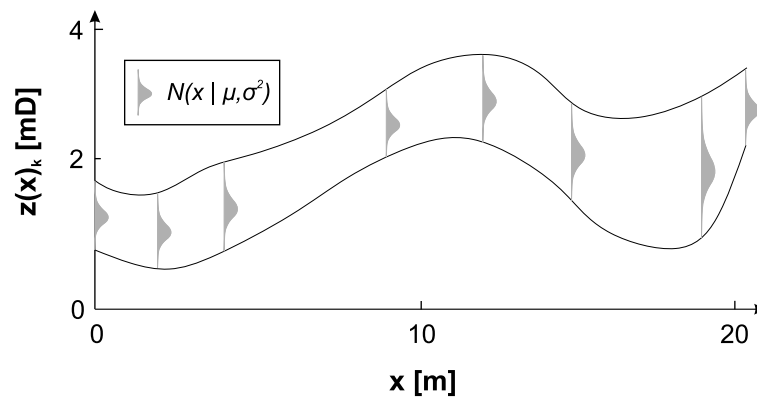


Figure 1. Conceptualization of a regionalized variable after [5] exemplarily illustrated for the intrinsic permeability.

Lithological and physical variability is subject of numerous geoscientific studies [18–21] and is commonly termed *heterogeneity*. In the Oxford Dictionary [22] the word heterogeneity is defined as a *Difference or diversity in kind from other things* or a *Composition from diverse elements or parts; multifarious composition*. In most works, this term is used to describe that an object consists of multiple subsets being different to one another in one or more attributes. Li and Reynolds [1] restrict the term to be the variability of a system property in three-dimensional space. Fitch et al. [23] provide a set of methods to quantify heterogeneity within a sample of observations including the coefficient of variation (c_v),

$$c_v = \frac{\sqrt{\sigma^2}}{\mu}, \quad (1)$$

where σ is the standard deviation and μ is the arithmetic mean and the Dykstra–Parsons coefficient (c_{dp})

$$c_{dp} = \frac{x_{50} - x_{84}}{x_{50}}, \quad (2)$$

where x_n is the n th percentile of a distribution.

The continuity of a regionalized variable is thus dependent on the continuity of the geological media and may or may not provide continuity in a mathematical sense. In this work, we will use the term property for a regionalized variable, the term field for the (quasi-)continuous spatial distribution

of a property, and the term target domain Ω for an area of interest. When we mention global and local characteristics, we refer to characteristics of Ω and its subsets Ω_b , respectively.

2.2. Geostatistical Interpolation

Geostatistical interpolation techniques aim to estimate a value at unsampled locations of a property in Ω and build the base for sequential simulations. The most popular geostatistical interpolation technique is kriging. In the following subsections, we will briefly describe the theory behind kriging and focus on its variety simple kriging (SK). Moreover, we will discuss practical computational aspects such as neighborhoods.

2.2.1. Spatial Neighborhood

As the system of linear equations for geostatistical estimations might grow very large, those algorithms require subset-sampling in order to perform reasonably. Therefore, a 3-D search ellipsoid can be used to find the neighbors of a point in a mesh [24]. This ellipsoid can be defined by six properties: azimuth α ; dip β ; plunge γ together with the radius in X r_x , Y r_y , and Z direction r_z of the ellipsoid. α , β , and γ define the ellipsoid's clockwise rotation around the Z, X, and Y axes in this exact order. Accordingly, the rotation matrix T can be defined as

$$T = \begin{pmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \beta & \sin \beta \\ 0 & -\sin \beta & \cos \beta \end{pmatrix} \begin{pmatrix} \cos \gamma & 0 & -\sin \gamma \\ 0 & 1 & 0 \\ \sin \gamma & 0 & \cos \gamma \end{pmatrix}. \quad (3)$$

After translating the mesh such that $x_x = x_y = x_z = 0$ and rotating it according to Equation (3), Equation (4) can be used to determine, whether a point x' with the transformed coordinates x'_x , x'_y and x'_z is located inside or on the boundary of the search ellipsoid (≤ 1) or not (>1).

$$\left(\frac{r_x}{x'_x}\right)^2 + \left(\frac{r_y}{x'_y}\right)^2 + \left(\frac{r_z}{x'_z}\right)^2 \leq 1 \quad (4)$$

2.2.2. Variography

The variographic analysis is a crucial prerequisite for numerous geostatistical interpolation techniques. Hereby, the experimental semivariogram represents the cumulative dissimilarity of a discrete set of point-pairs with n_c representing the count of point-pairs within the distance classes \mathbf{h} of identical distance increments (Equation (5)).

$$\gamma(\mathbf{h}) = \frac{1}{2n_c(\mathbf{h})} \sum_{k=1}^{n_c(\mathbf{h})} (z(x_k + \mathbf{h}) - z(x_k))^2 \quad (5)$$

The continuous counterpart, represented by the theoretical semivariogram γ_{theo} , is an approximation of the experimental semivariogram assuming $z(\mathbf{x})$ to be a stationary random field [25]. γ_{theo} is used to fit the experimental variogram. The spherical variogram model γ_{sph} with a nugget effect is a popular nested model used to fit the experimental semivariogram [26,27], which is calculated by

$$\gamma(\mathbf{h})_{sph} = \begin{cases} n + b \cdot \left(\frac{3|\mathbf{h}|}{2a} - \frac{|\mathbf{h}|^3}{2a^3}\right) & \text{if } 0 \leq |\mathbf{h}| < a \\ n & \text{if } |\mathbf{h}| \geq a, \end{cases} \quad (6)$$

with n being the nugget, b the sill and a the range [6]. The variogram model represents a covariance function c with the relationship $\gamma(\mathbf{h})_{theo} = c(0) - c(\mathbf{h})$, where c is a positive definite, even function and $c(0) = n + b$ in case of a spherical variogram model with nugget effect. Semivariograms can be used to quantify the spatial or time correlation of a random variable [27–29]. c and γ_{theo} are input variables for geostatistical interpolation algorithms.

2.2.3. Simple Kriging

Kriging is a commonly used stochastic technique to interpolate geological rock properties in space and time [30]. The kriging estimator is the best linear unbiased estimator (BLUE) of a property as it minimizes the error variance. It incorporates the covariance structure of the globally sampled values into the weights for predicting the value $z(x_0)$ at an unsampled location x_0 [31]. Therefore, $z(x_0)$ is calculated by weighting the values of the sampled locations and building a linear combination of those what gives

$$z(x_0) = \sum_{k=1}^n w_k \cdot z(x_k), \quad (7)$$

where w_k is the weight of the sampled point x_k with the value $z(x_k)$. The kriging types primarily differ by their derivation of the weight vector. For all kriging systems, a system of linear equations must be solved as it is outlined in the following paragraphs, in which we will consider the simple kriging (SK) technique [32] and expand it by the integration of a locally varying mean [33]. Therefore, we modify Equation (7) into

$$z(x_0)_{SK} = \sum_{k=1}^n w_k \cdot z(x_k) + \left(1 - \sum_{k=1}^n w_k\right) \cdot \mu. \quad (8)$$

in which the known stationary mean μ has been added [6]. While SK assumes that μ is globally constant and known, SK with locally varying mean assumes μ to be constant only in the neighborhood of x_0 . In order to obtain the SK weights, a system of n linear equations must be solved in which n stands for the number of considered neighbors. This system of equations is defined as

$$\mathbf{A}\mathbf{w} = \mathbf{b}, \quad (9)$$

which corresponds to

$$\underbrace{\begin{pmatrix} c(x_1 - x_1) & \cdots & c(x_1 - x_n) \\ \vdots & \ddots & \vdots \\ c(x_n - x_1) & \cdots & c(x_n - x_n) \end{pmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} w_1^{SK} \\ \vdots \\ w_n^{SK} \end{pmatrix}}_{\mathbf{w}} = \underbrace{\begin{pmatrix} c(x_1 - x_0) \\ \vdots \\ c(x_n - x_0) \end{pmatrix}}_{\mathbf{b}}, \quad (10)$$

with c as covariance function and x_n as point with known value [25]. In SK each interpolated point provides a simple kriging variance σ_{SK}^2 [5], which we can calculate by means of the formula

$$\sigma_{SK}^2 = c(0) - \sum_{k=1}^n w_k c(x_k, x_0). \quad (11)$$

The quality of a kriging interpolation is dependent on the variogram model and its goodness of fit to the experimental semivariogram.

2.2.4. Consideration of Measurement Error Variance

We already saw that kriging induces a local interpolation error by itself, namely, σ_{SK}^2 . There are, however, also other components which bias the interpolation result. Besides σ_{SK}^2 , the local and unknown variability of $z(\mathbf{x})$ in Ω_b as well as the measurement error variance σ_m^2 might play an important role (Figure 2). Integrating σ_m^2 into an interpolation can be achieved by estimating the measurement error precision σ_m with a variance of σ_m^2 and incorporating it into the kriging system of linear equations giving

$$\begin{pmatrix} c(x_1 - x_1) + \sigma_1^2 & \cdots & c(x_1 - x_n) \\ \vdots & \ddots & \vdots \\ c(x_n - x_1) & \cdots & c(x_n - x_n) + \sigma_n^2 \end{pmatrix} \begin{pmatrix} w_1^{SK} \\ \vdots \\ w_n^{SK} \end{pmatrix} = \begin{pmatrix} c(x_1 - x_0) \\ \vdots \\ c(x_n - x_0) \end{pmatrix}. \quad (12)$$

In contrast to the conventional formula, σ_m^2 with regard to the considered known value at x_k is added in the diagonal of the matrix [25]. This accounts for the heteroscedastic nature of geological parameters as they commonly show a higher variability for high values and a lower variability for low values.

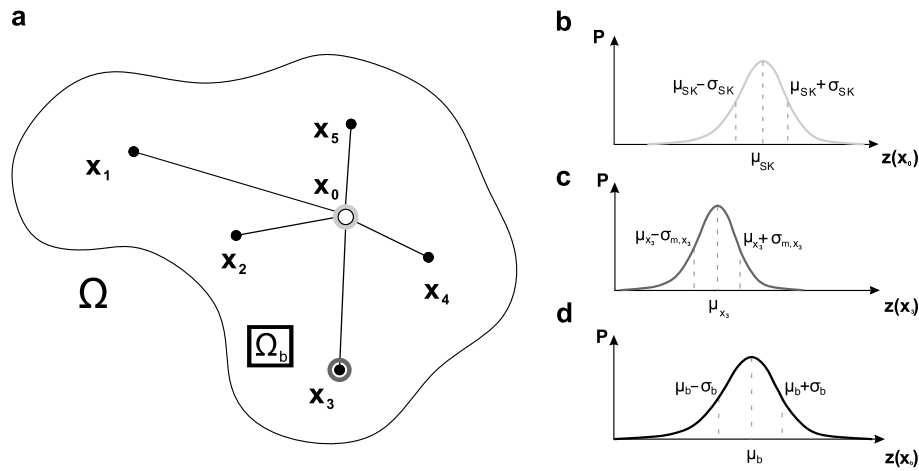


Figure 2. Schematic of the uncertainty components integrated into a predictive model of rock properties. (a) Illustration of an interpolation process using neighboring points x_k with known values to predict the unknown value at x_0 . (b–d) Schematic of the local probability density functions (PDFs) in form of a Gaussian distribution defined by σ^2 and μ for the estimated kriging error variance σ_{SK}^2 at x_0 (b), the observed measurement error σ_m^2 at the point x_3 (c) and the observed variance σ_b^2 in a subset Ω_b of Ω (d).

2.3. Sequential Simulation

In contrast to geostatistical interpolation techniques, sequential simulations aim to reproduce the global statistics in form of the considered variogram model and the global histogram. Therefore, in order to account for the spatial heterogeneity of a rock property, the sequential Gaussian simulation (SGS) and the direct sequential simulation (DSS) algorithm can be utilized for univariate simulation. SGS is based on the multi-Gaussian approach [33], which assumes that the kriging error is standard normally distributed with $\mu = 0$ and $\sigma_{SK}^2 = 1$. This requires that each one-point cumulative density function (CDF) of any linear combination of the RV is normally distributed, that all subsets of the RF are multivariate normal, that the two-point distribution is normal and that all conditional distributions of subsets of the RF are normal [33]. If the RF fulfills the requirements, then the simple kriging estimate and variance characterize the posterior cumulative CDF under consideration of the normal score variogram model. Thus, we need to transform the original distribution's CDF into standard normal space for SGS. In order to transform any point in the CDF ($F(Z(\mathbf{u}))$) of any random variable $Z(\mathbf{u})$ to a random function $Y(\mathbf{u})$ and vice versa the following equation can be applied,

$$Y(\mathbf{u}) = \phi(Z(\mathbf{u})) = G^{-1}[F(Z(\mathbf{u}))], \quad (13)$$

where G^{-1} is the inverse Gaussian CDF of $Y(\mathbf{u})$, which is also named quantile function [34], and ϕ is the inverse Gaussian CDF of $F(Z(\mathbf{u}))$. Thus, z and y correspond to the same probabilities. For each previously interpolated point x_j now a random value of the normal distribution $\mathcal{N}(\mu_{SK}, \sigma_{SK}^2)$, whose PDF defines as

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad (14)$$

is drawn as $z(x_0)$ using the Box–Muller transform [35]. We can perform this transform by applying the equation

$$z(x_0) = \sqrt{-2 \cdot \log(u_1) \cdot \cos(2\pi \cdot u_2)} \cdot \sigma + \mu, \quad (15)$$

with u_1 and u_2 as random numbers $\in [0, 1]$, σ as the standard deviation, and μ as the mean of the original distribution. The simulation is eventually back-transformed into the original space using a quantile-quantile back-transformation mapping technique. The reproduction of the covariance model, however, does not require the multi-Gaussian approach as long as the estimate and variance are derived from the SK estimation [9,10]. Thus, the conditional distribution type, which is chosen in order to simulate the variability at each point, does not necessarily need to be Gaussian. With this in mind, it is evident that a normal score transform is not needed before performing a sequential simulation. This results in the DSS approach, which commonly samples from the global PDF by determining the sampling interval from the local PDF [9].

2.4. Model Validation

2.4.1. Cross-Validation

In order to assess the quality of a realization, models, which are constructed by means of interpolation or simulation techniques, should be validated. Commonly, interpolations are validated by cross-validation. This technique is usually performed by using point removal procedures called leave-p-out cross-validation (LpO CV). For the LpO CV, p randomly selected samples are removed from the input data set of size n with $0 < p < n$ and the interpolation or simulation is performed without these samples [36]. As measures of goodness of fit, the mean-square error (MSE, Equation (16)), the root-mean-square error (RMSE, Equation (17)), and the mean-absolute error (MAE, Equation (18)) of the realization can be calculated as

$$MSE = \frac{1}{n} \sum_{k=1}^n (\hat{z}(x_k) - z(x_k))^2, \quad (16)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^n (\hat{z}(x_k) - z(x_k))^2} \quad (17)$$

and

$$MAE = \frac{1}{n} \sum_{k=1}^n |\hat{z}(x_k) - z(x_k)|, \quad (18)$$

where $\hat{z}(x_k)$ are the simulated points. While Willmott et al. [37] question the status of the triangle inequality for the RMSE, which is required for a distance function metric, Chai and Draxler [38] show that the RMSE in fact fulfills this condition. Thus, if the model errors follow a normal distribution, the RMSE is to favor over the MAE [38].

2.4.2. Ergodic Fluctuations

The minimum requirement for geostatistical simulations is their ability to reproduce the original data, the global summary statistics and the global variogram model [8,39]. Ergodic fluctuations refer to the discrepancy between the model parameters and the realizations' statistics [6]. In the case of the variogram model, the discrepancy of a realization to the variogram model is related to the limitation of the integrated constraints to a limited neighborhood. This, in fact, leads to higher errors at far ranges within the simulation. In this study, we quantified the ergodic fluctuation of a realization by estimating the average MSE between the experimental semivariogram and the variogram model. If a realization's discrepancy among the experimental variogram and variogram model does not exceed the original values discrepancy, the variogram reproduction is said to be within the range of ergodic fluctuations.

3. Sequential Simulation using a Local Variance Model

In this section, we will describe how the SGS and DSS algorithms need to be modified in order to sample from a local variance model (LVM). The LVM can be described as a global representation of the locally observable variance σ_{LVM}^2 in one mesh cell. Thus, the LVM can be referred to as the local geological heterogeneity. The LVM is constructed using a mapping technique in which the value of the mapped variances is constrained by a set of measurements. Those are intended to represent the small-scale variability present at the mapped position. Subsequently, the variance is interpolated onto Ω . The basic concept of interpolating a distribution in space is illustrated in Figure 3c.

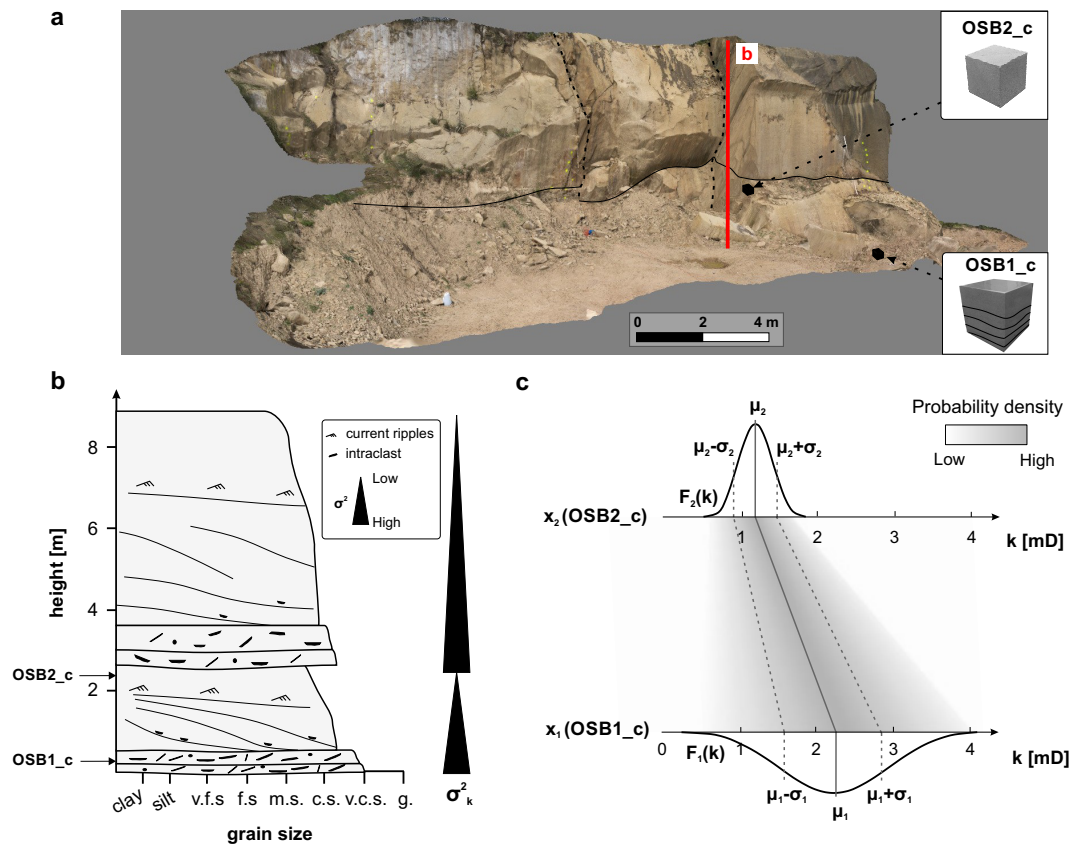


Figure 3. (a) Photogrammetric model of the investigated sandstone quarry. The outcrop is compartmentalized by two scissor faults and consists of two lacustrine-deltaic Bouma sequences [40]. (b) Sedimentological 1-D section of the sedimentary architecture observed in the outcrop. The Bouma sequence provides an erosive base. One sequence is characterized by a fining-upward trend and consists of intraclasts-rich massive sandstones at the base and trough cross-bedded and ripple cross-bedded sandstones towards top [40]. (c) Spatial interpolation of a PDF exemplary illustrated with both theoretical Gaussian distributions derived from the measurements of OSB1_c and OSB2_c.

The sequential simulations are performed on the nodes of Ω using a modification of the SGS and DSS algorithms, namely, the LVM-SGS and LVM-DSS. Our basic idea is that, if and only if the geological heterogeneity is exceeding σ_{SK}^2 at x_k , we will sample from the LVM-constructed PDF instead of from the kriging-derived PDF. Otherwise, if the interpolation error is greater than the expectable geological heterogeneity, we will sample from the kriging-derived PDF. The generalized algorithm is displayed in Algorithm 1. All analyses have been conducted with the open-source software GeoReVi [41] in which the new algorithms have been implemented as extensions in the C# programming language (Appendix A).

Algorithm 1 LVM-SGS and LVM-DSS

Given: Ω ; \mathbf{x} ; N ▷ Target domain; Sampled locations; Neighborhood information;
Initialize: \mathbf{u}_{sim} ; \mathbf{x}' ▷ Simulated locations; Spatial neighbors;
if GMV-SGS **then**
 $Y(\mathbf{x}) \leftarrow \text{Equation (13)}$ ▷ Transform to standard normal space
end if
 $\gamma(\mathbf{h}) \leftarrow \text{Equation (5)}$ ▷ Estimate the experimental variogram
 $\gamma(\mathbf{h})_{\text{sph}} \leftarrow \text{Equation (6)}$ ▷ Derive the variogram model and the covariance function
for all u_i in Ω **do**
 $\mathbf{x}' \leftarrow \text{Equation (3) \& Equation (4)}$ ▷ Determine the neighborhood with N applied to \mathbf{x} & \mathbf{u}_{sim}
 $\mu_{SK} \leftarrow \text{Equation (8) using } \gamma(\mathbf{h})_{\text{sph}}$ ▷ From \mathbf{x}'
 $\sigma_{SK}^2 \leftarrow \text{Equation (11) using } \gamma(\mathbf{h})_{\text{sph}}$ ▷ From \mathbf{x}'
 Allocate $\sigma_{LVM(x'_i)}^2$
 if $\sigma_{SK}^2 \geq \sigma_{LVM(x'_i)}^2$ **then**
 $z(u_i) \leftarrow \text{Equation (15) from } \mathcal{N}(\mu_{SK}, \sigma_{SK}^2)$ ▷ Draw a value with σ_{SK}^2
 else
 $z(u_i) \leftarrow \text{Equation (15) from } \mathcal{N}(\mu_{SK}, \sigma_{LVM}^2)$ ▷ Draw a value with σ_{LVM}^2
 end if
 Add $z(u_i)$ to \mathbf{u}_{sim}
end for
 $F(Z(\mathbf{u})) \leftarrow \text{Equation (13),}$ ▷ Back-transform the simulated values into the original space

3.1. Case Study

In order to test and evaluate the new workflow with the modified algorithms, we conducted an outcrop analogue study in a quarry in Germany. In the following subsections, we will outline the object of investigation, the sampling strategy and the modeling techniques used to implement the LVM-SGS and LVM-DSS algorithms. We decided to use the intrinsic permeability k for the implementation as that property plays a critical role in numerous types of subsurface utilization—especially with regard to subsurface reservoirs.

3.1.1. Object of Investigation

An actively quarried sandstone outcrop (long. 7.647546, lat. 49.523821) in Obersulzbach, which is located in the Saar-Nahe basin in southwestern Germany, has been selected as object of investigation (Figure 3a). The outcrop exposes the Disibodenberg Formation of the innervariscan Rotliegend Group, which constitutes a deeply buried [42] potential hydrothermal reservoir unit [43] in the northern Upper Rhine Graben. The Disibodenberg Formation in the quarry is composed of two Bouma sequences (Figure 3b) from a lacustrine delta, which deposited during Permian times. There were two selection criteria being decisive for selecting the quarry. On the one hand side, the sedimentary beds are ≥ 2 m thick and laterally continuous. Moreover, the outcrop is actively mined, which reduces the impact of recent weathering onto the permeability. Moreover, it was possible to extract both rock samples from the outcrop wall as well as oriented rock cubes from different representative lithofacies types in order to conduct multi-scale three-dimensional investigations. The outcrop measures $50 \times 15 \times 10$ m and thus owns the size of a typical cell in common static and dynamic reservoir models (see, e.g., in [44]).

3.1.2. Sampling Strategy

Numerous studies showed that the physical variability in geological media must be integrated as a function of measurement volume, also known as the representative elementary volume (REV) [45].

The REV denotes a volume, at which a representative amount of heterogeneity is captured by one measurement [46] minimizing the smaller-scale fluctuations. Therefore, a multi-scale approach based on the concept of the REV has been implemented. Accordingly, 39 cylindrical rock samples with diameters and lengths of four centimeters were extracted from the outcrop wall. The samples were taken from six 1-D profiles covering the entire quarry area (Figure 4a). More information regarding the sample positions and orientations can be found in Linsel [41]. Those samples were used for the global field simulations.

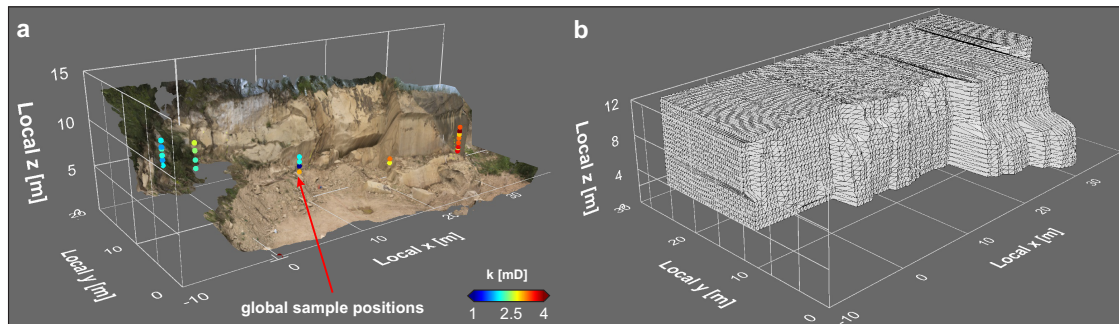


Figure 4. (a) Photogrammetric model of the investigated sandstone quarry in Obersulzbach, Germany. Sample locations are displayed as spheres, whose color indicates the observed permeability value at the sample locations. (b) Hexahedral non-orthogonal mesh of the investigated outcrop generated by an IDW interpolation using the nodes of the photogrammetric model as constraints.

The quarry contains sequences from a prodelta mouthbar deposited as turbiditic densites. The sequences graduate from a high-energetic depositional environment at the base to a low-energetic environment at the top as the flow velocity is steadily declining. The sequences consist of heterogeneous, intraclast-rich sandstones at the base and of trough cross-bedded, ripple cross-bedded and homogeneous sandstones at the top. Consequently, the sequences can be declared as Bouma sequences containing the Bouma A to Bouma E intervals in a fluvial-dominated lacustrine-deltaic depositional environment [47]. Based on that, we assumed that the variability within one Bouma sequence is highest at the base and lowest at the top (Figure 3b).

Accordingly, two rock cubes of $0.2 \times 0.2 \times 0.2$ m (OSB2_c) and $0.25 \times 0.25 \times 0.25$ m (OSB1_c) were taken—one from the top (Bouma E) and one from the base (Bouma A) of one sequence—in order to capture both the highest and the lowest variability. The locations of the cubes within the quarry and the strata are shown Figure 3a,b.

We selected two types of lithofacies: OSB1_c, a discontinuously cross-bedded, intraclast-rich lithofacies type and OSB2_c, a homogeneous lithofacies type without macroscopically observable internal bounding surfaces. In total, 79 rock cylinders were extracted from rock cube OSB1_c and 29 from OSB2_c. More information regarding the sampling process can be found in Linsel et al. [40]. Those samples were used for constraining the LVM.

3.1.3. Laboratory Measurements

The cylinder samples were cut, oven-dried at $105\text{ }^{\circ}\text{C}$ and measured in the laboratory for determining the intrinsic gas permeability k at unsaturated conditions. k can be considered one of the key parameters of geothermal reservoir rocks with regard to hydrothermal systems in porous aquifers [48]. k was measured with the Hassler cell Darmstadt permeameter. The device's functionality is described in detail in Filomena et al. [49]. The permeability is provided in the industry-standard unit millidarcy (mD), where 1 mD corresponds to $9.869 \times 10^{-16} \text{ m}^2$. The permeability measurement provides an error variance between 0 and 0.15 mD^2 in the range of the observed values [50].

3.1.4. Mesh Generation

In order to construct Ω , the outcrop wall is modeled using a photogrammetric representation that was downsampled into 40×20 faces and subsequently interpolated using Shepard's p -value IDW interpolation, which we can write as

$$z(x_0) = \frac{\sum_{k=1}^n \left(1/d_k^p\right) \cdot z(x_k)}{\sum_{k=1}^n 1/d_k^p}, \quad (19)$$

where d is the Euclidean distance between the the point with the known value x_k and the point with the unknown value x_0 and p is an exponent factor to influence the weights non-linearly. IDW has been applied with a short search radius of five meters and a power parameter of four. The interpolation result has an RMSE of 0.024 m, which can be considered low for the surface interpolation. The resulting outcrop surface is used as a bounding surface for a hexahedral mesh, which represents Ω , that is composed of 75,240 cells (Figure 4b, Table 1). The rock cubes, which represent Ω_b , are constructed by an orthogonal, hexahedral mesh containing 25,230 (OSB2_c) and 64,000 cells (OSB1_c), respectively. The volume of an average cell of the outcrop mesh is roughly eight times the volume of OSB1_c and 15 times the volume of OSB2_c (Table 1).

Table 1. Statistical characteristics of the outcrop mesh and both cube meshes (n_n = number of nodes, n_c = number of cells, V = volume of the mesh, \bar{V}_c = average volume of a mesh cell).

Object	n_n [-]	n_c [-]	V [m ³]	\bar{V}_c [m ³]
Outcrop (Ω)	82,000	75,240	9000	0.12
OSB1_c (Ω_b)	68,921	64,000	0.0156	6.19×10^{-7}
OSB2_c (Ω_b)	31,500	25,230	0.008	1.25×10^{-7}

The variance σ_c^2 derived from the measurements conducted on the samples from the rock cubes is assumed to represent the variance $\sigma_{\Omega_b}^2$ that can be expected in one cell of the outcrop mesh so that

$$\sigma_{LVM}^2 \approx \sigma_{\Omega_b}^2, \quad (20)$$

with σ_{LVM}^2 being the local sample variance, which we can calculate by means of the formula

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2, \quad (21)$$

where n is the total number of samples, μ is the mean and x_i is the sample at the i th location.

4. Results

4.1. Spatial Variability

The variogram analysis reveals a range of 0.3 m and 0.2 m for the rock cube samples OSB1_c and OSB2_c, respectively, and a range of 18 m for the outcrop samples (Figure 5a,d,g). The sill is slightly higher in the outcrop region as it is in the rock cubes. Moreover, the outcrop samples show a weak nugget effect. Generally, a scale effect can be observed in which the variance increases with the considered volume. This effect is also present in the descriptive statistics (Figure 5c,f,i).

The measurements from the outcrop region show a c_v of 0.28 and a c_{dp} of 0.31. The histogram indicates a normal distribution of k ranging from 0.7 mD to 4.6 mD (Figure 5b). A two-sided Kolmogorov–Smirnov test [51], which is based on an implementation of Simard and Ecuyer [52], confirmed the hypothesis that all samples come from a normal distribution. The application of Tukey's

outlier detection method [53] reveals no statistical outliers in the sample. By applying the classification scheme of Corbett and Jensen [54], the sample can be classified as being very homogeneous.

The local histogram of k from OSB1_c shows a bimodality in the distribution ranging from 0.7 to 3.9 mD (Figure 5e). OSB2_c's histogram shows a unimodal range from 0.8 to 1.5 mD (Figure 5h). Again, no statistical outliers can be detected. The local variability of OSB1_c is significantly higher than that of OSB2_c. k of OSB1_c provides a c_v of 0.3 and a c_{dp} of 0.4 while measurements from OSB2_c show values of 0.2 for c_v and 0.18 for c_{dp} . c_v and c_{dp} of OSB1_c tend to cover the variability of the global data. This result is in good agreement with the REV theory from Nordahl and Ringrose [45]. Both rock cubes can be classified being very homogeneous as well.

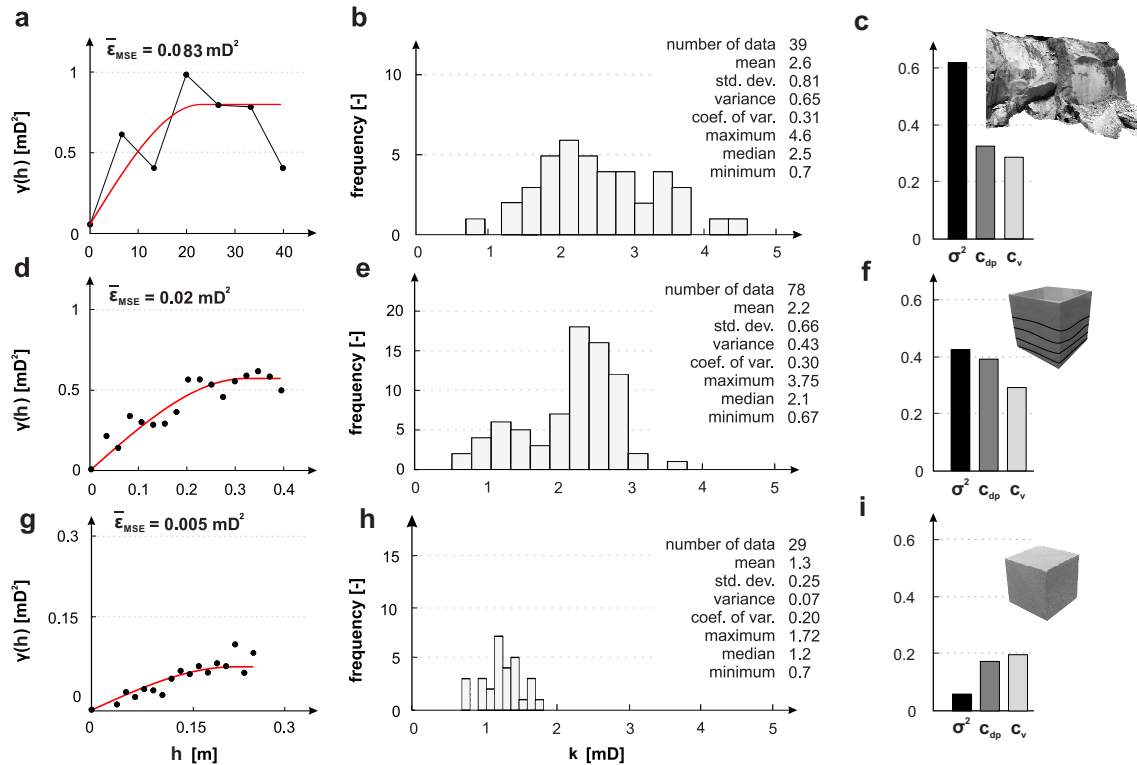


Figure 5. Empirical variogram and variogram model, empirical histogram, and heterogeneity-indexes derived from the k measurements for the outcrop (a–c), and the rock cubes OSB1_c (d–f) and OSB2_c (g–i). A scale-effect is observable in the heterogeneity-indicating coefficient of variation, the Dykstra–Parson coefficient and the sample variance. All variogram models are described by a spherical model with nugget effect. The variogram model for (a) is described by $n = 0.05 \text{ mD}^2$, $a = 23 \text{ m}$ and $b = 0.75 \text{ mD}^2$ with n as nugget, a as range, and b as sill. The model for (d) is described by $n = 0 \text{ mD}^2$, $a = 0.3 \text{ m}$ and $b = 0.58 \text{ mD}^2$ while the model of (g) is described by $n = 0.005 \text{ mD}^2$, $a = 0.18 \text{ m}$ and $b = 0.08 \text{ mD}^2$.

Thus, we can observe a significant small-scale variability. The bedding structures in OSB1_c are well preserved in the permeability field of the k interpolation, which gradually increases from low values between 0.7 and 2 mD in the lower beds to higher values between 2 and 4 mD in the upper beds (Figure 6a). In OSB2_c the trend is running diagonally through the rock cube (Figure 6b); however, no macroscopic bounding surfaces are visible, which could have had a control on the field of k here. It should be noted, however, that the range of k is significantly smaller here compared to OSB1_c.

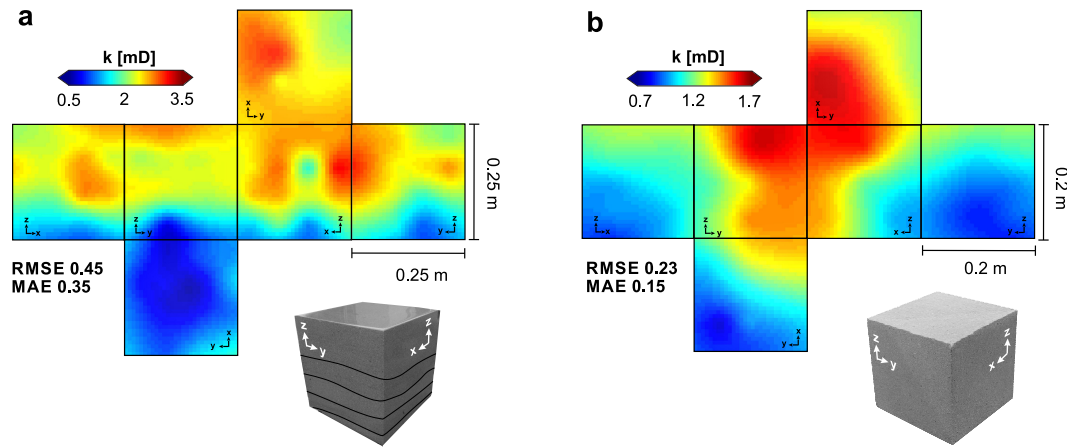


Figure 6. Spatial distribution of the intrinsic permeability in the rock cubes OSB1_c (a) [40] and OSB2_c (b) interpolated using the SK method.

4.2. Constructing the LVM

The LVM is constructed by means of a 3-D architectural element mapping of both Bouma sequences in the quarry. The base and the top of the sequences are mapped which are being used to constrain the LVM by the locally observable variance σ_{LVM}^2 . The exploratory data analysis reveals that the variance of k in OSB1_c is five times larger than that of OSB2_c. This is in accordance with the sedimentological mapping, which indicates a higher heterogeneity at the base of the Bouma sequence.

It is assumed that OSB1_c represents the most heterogeneous and OSB2_c the most homogeneous lithofacies type in the Bouma sequences as it is illustrated in Figure 3b. Accordingly, the positions of those lithofacies types are mapped throughout the quarry area and parameterized with σ_{LVM}^2 , which has been determined by the k measurements of OSB1_c and OSB2_c. Thus, we use $\sigma_{LVM}^2 = 0.43$ for mapping the base boundaries of the sequences throughout the outcrop area. Likewise, $\sigma_{LVM}^2 = 0.07$ is used as a local variance for the topmost boundary of the single sequences. The mapping locations of σ_{LVM}^2 are shown in Figure 7a. The points mapped onto the outcrop model are subsequently interpolated onto Ω by using a SK-based interpolation procedure for parametric PDFs (Figure 3c). The interpolation is conducted using 5 neighbors, a range of five meters, a sill of 0.005, a nugget of 0 and a plunge of 10° as the strata gently dip towards south. Figure 7b shows the constructed LVM which is being used by the sequential simulation algorithms. It should be noted that we have a decent offset in the LVM in the area of the central fault zones.

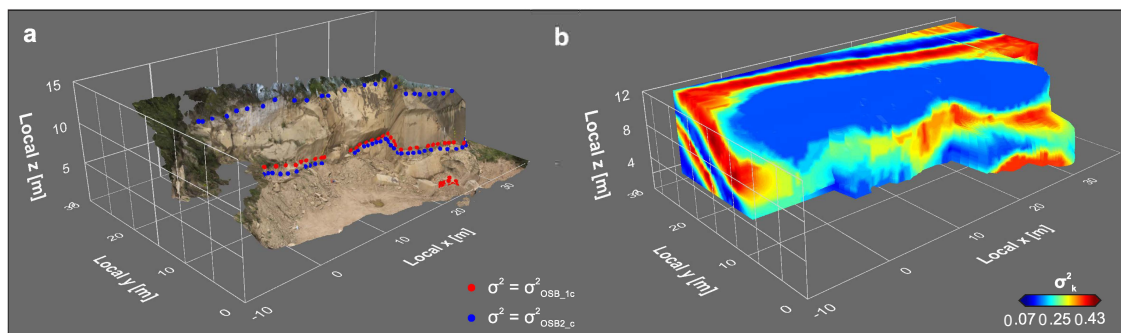


Figure 7. (a) Mapping of the local variance with regard to the observed geological structure. The highest variance is indicated by red spheres whereas the lowest variance is indicated by blue ones. The variance is derived from the rock cube measurements of OSB1_c—representing the most heterogeneous lithology at the bottom of the Bouma sequences (red)—and OSB2_c—likewise representing the most homogeneous lithology at the top of the Bouma sequences (blue). (b) The 3-D local variance model (LVM) representing the locally observable variance, which is constrained by the mappings shown in (a).

4.3. Optimizing the BLUE for Sequential Simulation

Prior to sequential simulation, the optimal SK conditions with regard to the integrated measurement error variance σ_m^2 and the selected neighborhood are determined. Therefore, a simple integer programming optimization is performed using varying measurement error variances ($0.0 \text{ mD}^2 \leq \sigma_m^2 \leq 0.15 \text{ mD}^2$) and a varying number of neighbors ($10 \leq n_n \leq 20$) as inequality constraints. We can express the optimization problem as

$$\begin{aligned} \min_{\sigma_m^2 \in \mathbb{R}, n_n \in \mathbb{N}} \quad & \epsilon_{SK}(\sigma_m^2, n_n) \\ \text{subject to} \quad & 0 \leq \sigma_m^2 \leq 0.15 \\ & 10 \leq n_n \leq 20, \end{aligned} \quad (22)$$

in which the SK error ϵ_{SK} in form of the RMSE and MAE must be minimized. The response surface of the numerical optimization process indicates that the SK error is generally declining when σ_m^2 is increasing. The lowest errors are produced with an n_n of 10, 11, and 20. This sensitivity of the SK error on the number of neighboring points is not unusual. The numerical optimization reveals that the optimal conditions for SK are met at $n_n = 20$ and $\sigma_m^2 = 0.15$ which yields a RMSE of 0.708 mD (Figure 8). The interpolation error can be reduced by 16.5% for the RMSE and by 18.5% for the MAE. The final SK realization and the spatial distribution of σ_{SK}^2 for that exact model is illustrated in Figure 9. It should be noted that the spatial distribution of σ_{SK}^2 in a sequential simulation is different as previously simulated locations are considered as well.

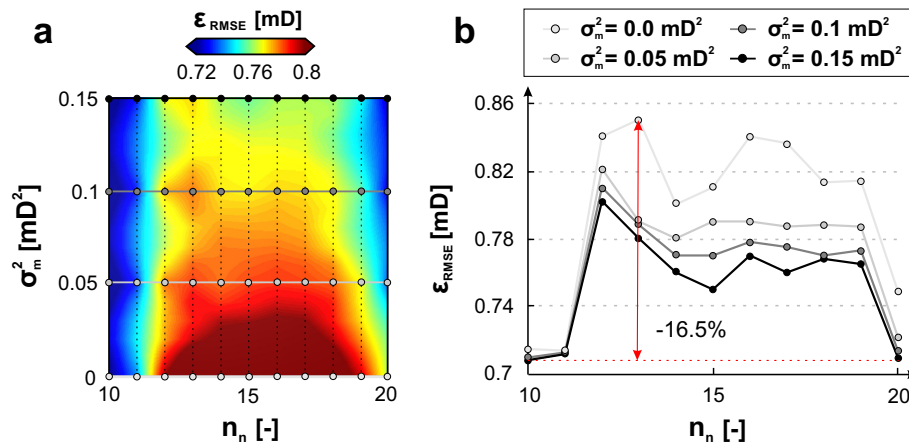


Figure 8. Results of the linear integer programming optimization using the marked sampling points. The interpolation error ϵ_{RMSE} is minimized using the inequality constraints given in Equation (22). (a) RMSE response surface with regard to the incorporated measurement error variance σ_m^2 and the maximum number of neighbors n_n using a leave-one-out cross-validation. (b) Cross sections through the response surface of (a).

The final modeling variables for the sequential simulations are given in Table 2. For SGS and LVM-SGS, the original data are transformed into standard normal space with $\mu = 0$ and $\sigma = 1$. The transformation leads to an adaption of the considered variogram model as the sill is now 1 and not 0.75 with a nugget of 0 instead of 0.05.

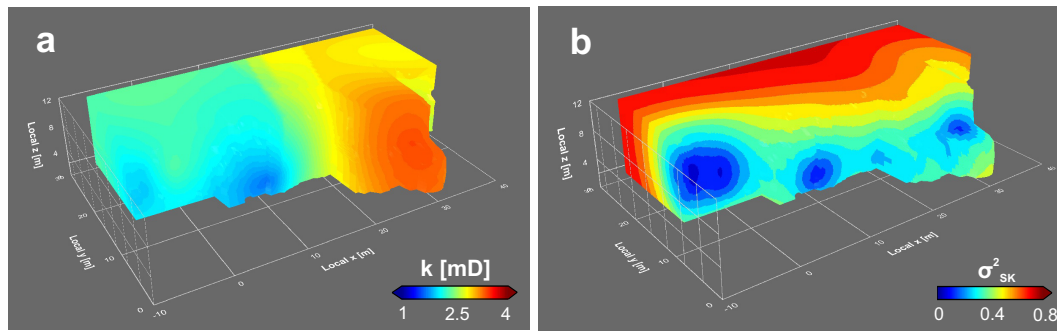


Figure 9. (a) Simple kriging estimate (b) and the local simple kriging variance for one SK realization.

Table 2. Modeling variables for the sequential simulations.

Variable	SGS & LVM-SGS	DSS & LVM-DSS
BLUE	SK	SK
Normal score transform	yes	no
Quantile-quantile back transform	yes	yes
Range x	50 m	50 m
Range y	50 m	50 m
Range z	15 m	15 m
Nugget	0.05	0
Sill	0.75 mD ²	1 mD ²
Range	23 m	23 m
Max. number of neighbors	20	20
Azimuth	0°	0°
Dip	0°	0°
Plunge	10°	10°
Measurement error variance	0.15 mD ²	0.15 mD ²

4.4. σ_{SK}^2 versus σ_{LVM}^2

The statistical and spatial characteristics of σ_{SK}^2 and σ_{LVM}^2 differ tremendously. σ_{SK}^2 is unimodally distributed, whereas σ_{LVM}^2 provides a bimodal distribution (Figure 10a). It is evident that σ_{SK}^2 covers the total range of the considered covariance model while σ_{LVM}^2 's range is more limited. The probability of simulating variances between 0.2 and 0.43 mD² is higher when sampling from the LVM instead of the local SK variance (Figure 10b). The median between σ_{LVM}^2 and σ_{SK}^2 differs by ≈ 0.08 mD², which indicates that the variability simulated in a realization of conventional sequential simulation algorithms is systematically underestimated.

With regard to the variogram model, σ_{LVM}^2 has a range of 5 m and a sill of 0.36 mD², and σ_{SK}^2 has a range of 0.3 m and a sill of 0.1 mD². Thus, σ_{SK}^2 seems to be spatially uncorrelated and random. However, the grade of variability in the eastern part of the outcrop is slightly higher than in the western part. Therefore, in contrast to σ_{LVM}^2 , σ_{SK}^2 obviously does not provide the simulation algorithm with a spatial trend when simulating the local variability.

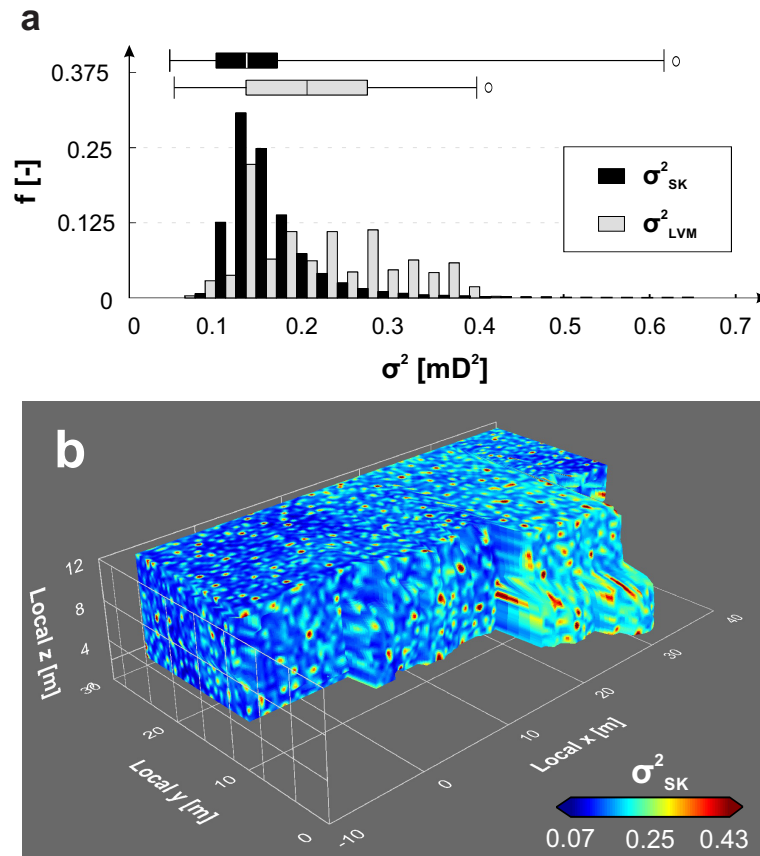


Figure 10. (a) Comparison of the empirical histograms of the σ^2_{SK} model produced in a DSS realization with the LVM and (b) the empirical distribution of σ^2_{SK} produced in the realization of (a).

4.5. Model Validation

All algorithms reproduce the considered variogram model within the range of ergodic fluctuations after back-transformation (Figure 11a–d). The quality of variogram reproduction has been evaluated by calculating the average mean square error $\bar{\epsilon}_{MSE}$ of all realizations between the experimental variogram and the variogram model. The best reproduction is produced by the LVM-DSS and LVM-SGS algorithms, while the latter one provides the lowest degree of ergodic fluctuations with $\bar{\epsilon}_{MSE} = 0.066$ mD². All realizations reproduce short-range dissimilarities well but slightly underestimate the dissimilarity at medium ranges. DSS and SGS tend to gentle underestimation at far ranges which is a drawback of limited neighborhoods. This effect, however, is less expressed in the LVM-based algorithms. For both types of sequential simulation, the LVM-based algorithm outperforms the conventional conditional simulation approaches.

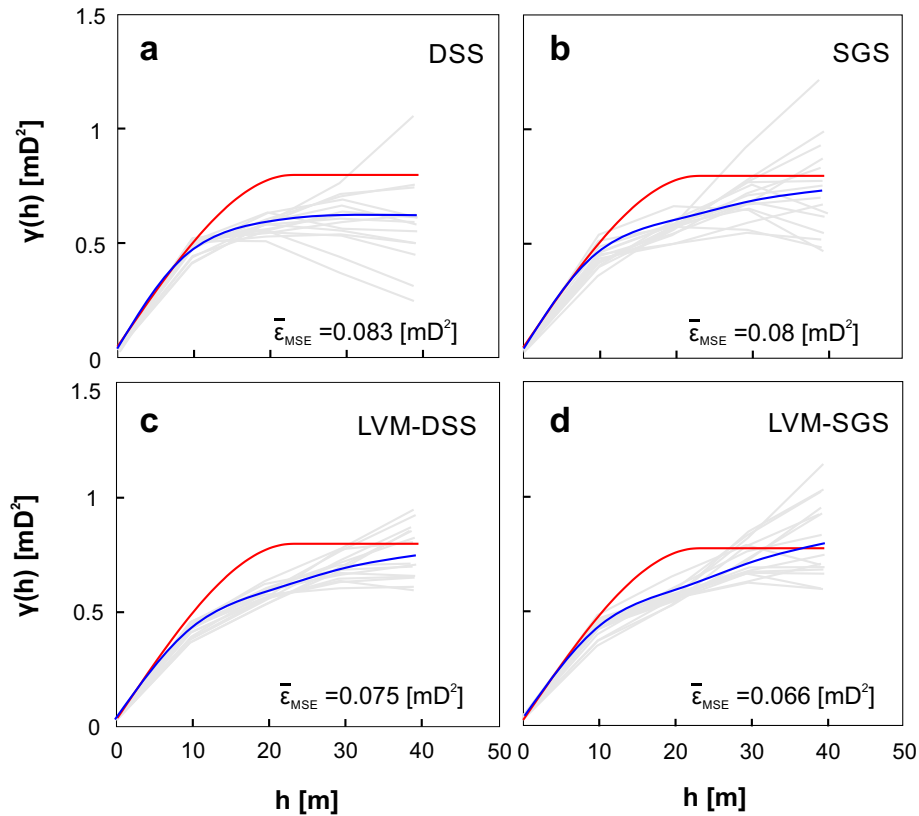


Figure 11. Experimental variograms (gray) for 15 realizations of DSS (a), SGS (b), LVM-DSS (c) and LVM-SGS (d) plotted together with the average over all realizations (blue) and the considered variogram model (red), which is described by a nugget of 0.05 mD², a range of 23 m and a sill of 0.75 mD².

Visual Outputs

It is evident that all simulation algorithms provide visually comparable results (Figure 12). It should be noted that the quadrilaterals of the 3-D models are subdivided using the Catmull–Clark scheme [55] for visualization. Within this scheme, a new point in a quadrilateral is calculated by

$$x_j^{k+1} = \frac{1}{n} \sum_{i=0}^{n-1} x_i^k, \quad (23)$$

with x_j^{k+1} as the new point at subdivision step $k + 1$ in the center of the element j with n vertices at the subdivision step k . This technique smooths the observable patterns in the models. There is an obvious trend in all realizations, which indicates that the highest values are located in the eastern part of the quarry and the lowest values in the western part. Having in mind that the applied algorithms are conditional, this trend is in well accordance with the constraints as given by the global measurements, which also provide the highest values in the eastern part of the quarry and the lowest values in the western part (Figure 4a). The trend is most clearly depicted in the DSS and LVM-DSS realizations (Figure 12). SGS and DSS tend to construct homogeneous regions more likely than their LVM equivalents. Thus, those algorithms might indicate a homogeneity, which is likely not present in the strata. Moreover, the heterogeneity of the LVM equivalents is more realistically oriented along the bounding surfaces in the quarry than the models produced by the conventional algorithms.

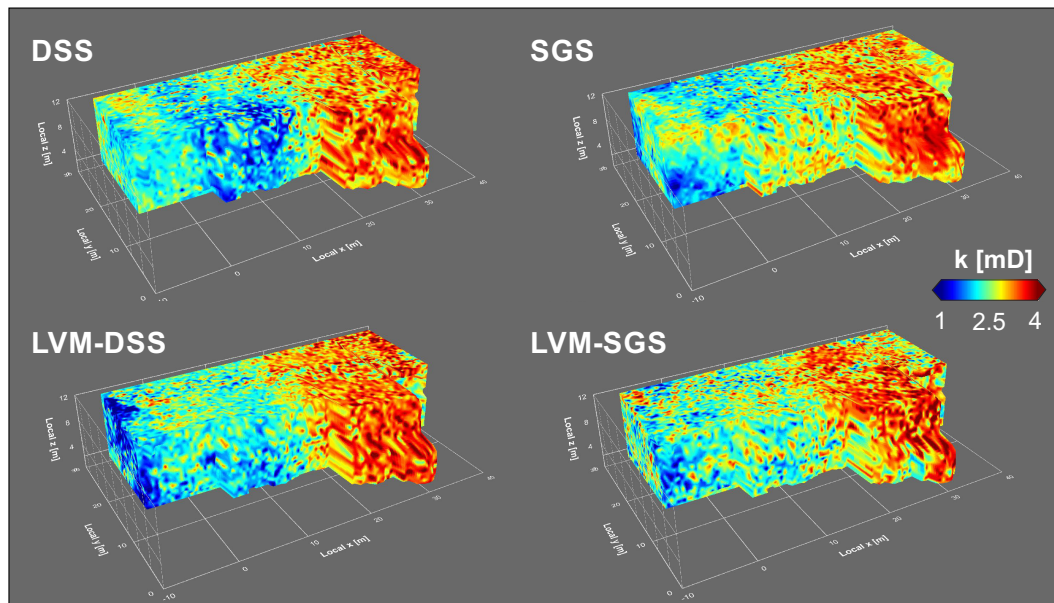


Figure 12. Exemplary model visualizations for the DSS, SGS, LVM-DSS and LVM-SGS realizations.

5. Discussion

In this study, we present a workflow that accounts for the locally observable geological variability in modified versions of conventional sequential simulation algorithms. Our approaches produce similar outputs as the conventional algorithms and reproduce the global variance model together with the global summary statistics, which are important criteria for the validity of a statistical simulation [8,10,39]. Our results are confirming the concept of the REV [45], in which the complexity of a continuous random variable is increasing with reducing the scale of observation. Moreover, we can confirm that σ_{SK}^2 constitutes no measure for the local estimation accuracy [56] as it is only reflecting the spatial configuration of the constraining data points being simultaneously independent on the constraints' values [6]. There are, however, two points which must be raised in order to discuss the benefits as well as the drawbacks of our approach.

5.1. Construction of the LVM

The main source of errors in the proposed workflow is based on the construction of the LVM. The LVM has been derived by an integrated approach of measuring the local variability in the most homogeneous and most heterogeneous lithofacies types in the sedimentary succession. The statistical analysis revealed that this assumption proved to be true as the heterogeneity measures in OSB1_c indicate a way higher variability as is present in OSB2_c. This, in fact, is building the basis for this study. The variance has been assumed to be constant at the base and at the top of a Bouma sequence. This assumption is limited by the number of samples taken within this case study. By constructing the LVM with an SK interpolation, we assume that the variance in one sedimentary Bouma sequence is continuous in a mathematical sense. This assumption might be proved to be too simple in future studies. In order to validate those results, more local samples would be necessary to constrain the LVM. This is a drawback in comparison to conventional SGS and DSS algorithms as those are not dependent on estimating a global variability model.

5.2. Comparison of the Spatial Distribution of the Local Variance

Figure 13a,b illustrates the relationship between the LVM and a DSS realization (a) and an LVM-DSS realization, respectively (b). Although the overall trend remains identical among both types of simulation, the spatial distribution of local variability is uncorrelated and inherently different. In the DSS realization, the heterogeneity within the region is randomly distributed. The most heterogeneous

areas in the LVM-DSS realization reside in the light areas—in which σ_{LVM}^2 is high—whereas the most homogeneous regions reside in the dark ones—where σ_{LVM}^2 is low. As the spatial distribution of σ_{SK}^2 is primarily dependent on the distance to the constraining neighbors, the SGS and DSS algorithms, in contrast to their LVM-based modifications, cannot account for a realistic spatial distribution of the local geological variability. This observation is conceptually illustrated in Figure 13c, which shows the spatial relationship between σ_{SK}^2 and σ_{LVM}^2 , as implied by the results of our study. It is evident that the conventional algorithms underestimate the local geological variability in close ranges to conditional data. It is also evident that σ_{SK}^2 systematically underestimates the natural variability present in the geological medium, which is investigated in this study (Figure 10a). Therefore, SGS and DSS might not be able reproduce the total geological variability as shown in this study, which is an advantage of the proposed algorithms instead.

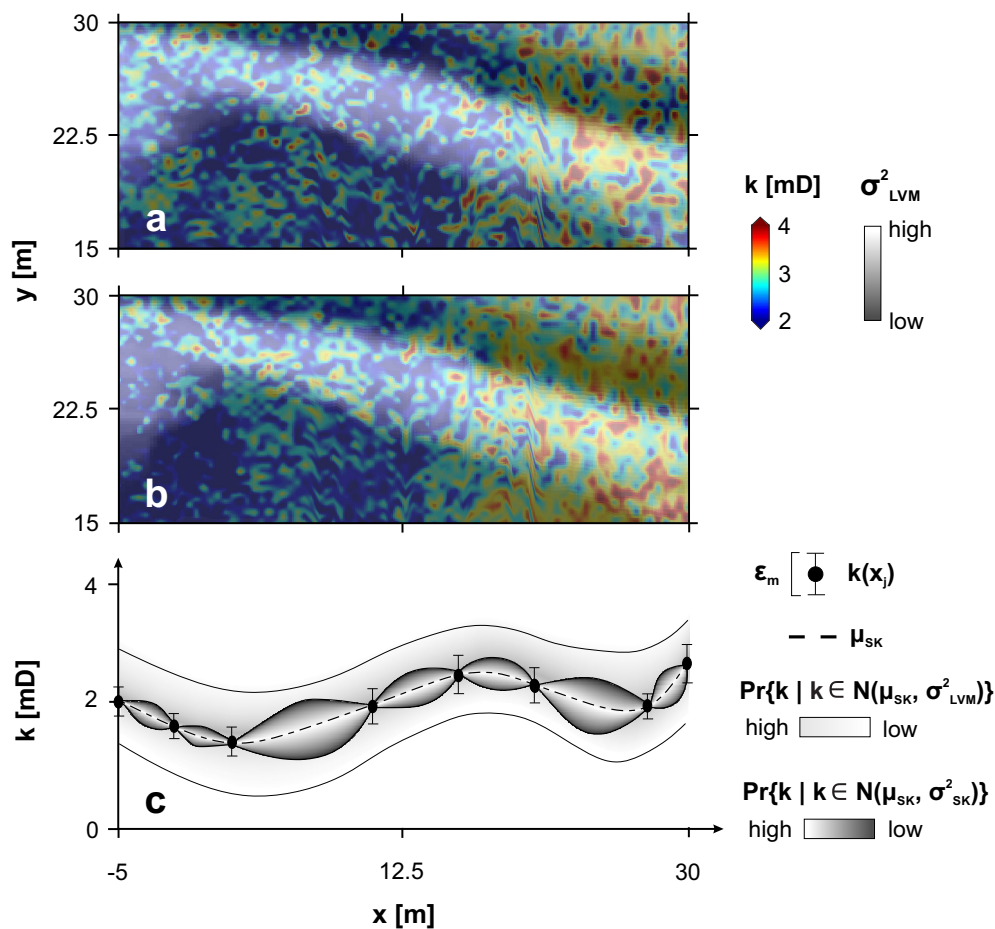


Figure 13. Top-view onto a representative simulation result of DSS (a) and LVM-DSS (b) superimposed by a gray-scale representation of the LVM with an opacity of 0.6. It is evident that the LVM-based algorithms' heterogeneity is highest in that area of the LVM in which it provides the highest local variance as well. The conventional approach, however, does not reflect the expected variance in space. (c) Conceptual illustration showing the spatial distribution of the constraining measurements $k(x_j)$ and the spatial relationship between the simple kriging estimate μ_{SK} with the measurement error ϵ_m and the two parameters used to simulate k in this study namely σ_{SK}^2 and σ_{LVM}^2 . Pr stands for the probability of k under the condition that k belongs to the Gaussian distribution described by μ_{SK} together with either σ_{SK}^2 or σ_{LVM}^2 .

6. Conclusions

In this study, we propose a new workflow, which incorporates the locally observed variability σ_{LVM}^2 into sequential simulations. We could demonstrate that the local simple kriging variance σ_{SK}^2 differs from σ_{LVM}^2 in local volumes of the target region. Therefore, the DSS and SGS algorithms have been modified by the replacement of σ_{SK}^2 through the measurement-derived σ_{LVM}^2 within one mesh cell. This replacement has been done if and only if $\sigma_{LVM}^2 \geq \sigma_{SK}^2$. The LVM has been constructed by means of geological mapping and the assumption that the variability is highest in the most heterogeneous lithology and lowest in the least heterogeneous lithology in a Bouma sequence. The proposed approach can be used in any type of spatial property simulation but is especially tailored for geological media.

The LVM-DSS and LVM-SGS approaches reproduce the observed variability in the sedimentary succession adequately yet reproducing the minimum required statistical measures of a valid simulation including the global histogram, the global heterogeneity, and the variogram model. Moreover, in contrast to their conventional representatives, the LVM-based algorithms account for the spatial distribution of the expected local variance adequately. Once the LVM is derived, it may be integrated into other geostatistical simulation algorithms such as the turning bands method [57–60].

From our results we conclude the following.

1. The distance metrics RMSE and MAE in spatial interpolations can be optimized with regard to the measurement error variance and the optimal neighborhood.
2. Geological samples always represent a small subset of the local variability, which should be accounted for by high-resolution sampling at a random basis at the least.
3. The simple kriging variance does not necessarily account for the magnitude of local variability in geological media and definitely does not account for its spatial distribution.
4. The fact that the local simple kriging variance does not reflect a geological trend might lead to unforeseen problems when using sequential simulation-derived models as a basis for subsurface utilization processes because the full geological heterogeneity might not have been taken into account properly.
5. By introducing a measurement-constrained, geology-driven local variance model, the spatial distribution of the variance that is expected in the investigated quarry can be integrated into sequential simulations. This allows to simulate the geological variability, which might be greater than the simulated variability in conventional sequential simulation algorithms.

Future research should focus on comparing σ_{SK}^2 and σ_{LVM}^2 under the consideration of other physicochemical properties, other geological settings, and other scales. This might require adapting the assumptions on the spatial continuity of the variability which should, however, always be based on reliable geological analyses.

Author Contributions: Conceptualization, Adrian Linsel; methodology, Adrian Linsel; software, Adrian Linsel; validation, Adrian Linsel; investigation, Adrian Linsel, Joshua Haas and Sebastian Wiesler; data curation, Joshua Haas and Sebastian Wiesler; writing—original draft preparation, Adrian Linsel; visualization, Adrian Linsel; supervision, Kristian Bär and Matthias Hinderer. All authors have read and agreed to the published version of the manuscript.

Funding: A.L. has received financial support by a PhD scholarship from the Friedrich-Ebert-Stiftung, Germany, which is gratefully acknowledged.

Acknowledgments: We are grateful for the rock cube preparation by the IWAR (Technische Universität Darmstadt, Germany). We would like to thank three anonymous reviewers for their valuable comments on an earlier version of this paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CDF	Cumulative distribution function
DSS	Direct sequential simulation
LVM	Local variance model
LpO CV	Leave-p-out cross-validation
MAE	Mean-absolute-error
MSE	Mean-square error
PDF	probability density function
REV	Representative elementary volume
RMSE	Root-mean-square error
RF	Random function
RV	Random variable
SGS	Sequential Gaussian simulation
SK	Simple kriging

Appendix A. Code and Data Availability

GeoReVi is an open-source software for Windows systems available under <https://github.com/ApirsAL/GeoReVi>. Data is available under <https://www.doi.org/10.6084/m9.figshare.11791407.v2>.

References

- Li, H.; Reynolds, J. On Definition and Quantification of Heterogeneity. *Oikos* **1995**, *73*, 280–284. [[CrossRef](#)]
- Crooijmans, R.A.; Willems, C.J.L.; Nick, H.M.; Bruhn, D.F. The influence of facies heterogeneity on the doublet performance in low-enthalpy geothermal sedimentary reservoirs. *Geothermics* **2016**, *64*, 209–219. [[CrossRef](#)]
- Rodrigo-Illarri, J.; Reisinger, M.; Gómez-Hernández, J.J. Influence of Heterogeneity on Heat Transport Simulations in Shallow Geothermal Systems. In *Geostatistics Valencia 2016*; Springer: Berlin, Germany, 2017; pp. 849–862. [[CrossRef](#)]
- Shepard, D. A Two-Dimensional Interpolation Function for Irregularly-Spaced Data. In Proceedings of the 1968 ACM National Conference, New York, NY, USA, 27–29 August 1968; pp. 517–524. [[CrossRef](#)]
- Webster, R.; Margaret, A.O. *Geostatistics for Environmental Scientists*, 2nd ed.; Wiley & Sons, Inc.: Hoboken, NJ, USA, 2007; p. 330.
- Deutsch, C.V.; Journel, A. *GSLIB: Geostatistical Software Library and User's Guide*; Oxford University Press: Oxford, UK, 1998.
- Tran, T.T. Improving variogram reproduction on dense simulation grids. *Comput. Geosci.* **1994**, *20*, 1161–1168. [[CrossRef](#)]
- Leuangthong, O.; McLennan, J.A.; Deutsch, C.V. Minimum Acceptance Criteria for Geostatistical Realizations. *Nat. Resour. Res.* **2004**, *13*, 131–141. [[CrossRef](#)]
- Soares, A. Direct Sequential Simulation and Cosimulation. *Math. Geol.* **2001**, *33*, 911–926. [[CrossRef](#)]
- Robertson, R.K.; Mueller, U.A.; Bloom, L.M. Direct sequential simulation with histogram reproduction: A comparison of algorithms. *Comput. Geosci.* **2006**, *32*, 382–395. [[CrossRef](#)]
- Journel, A.; Alabert, F. Non-Gaussian data expansion in the Earth Sciences. *Terra Nova* **1989**, *1*, 123–134. [[CrossRef](#)]
- Verly, G. Sequential Gaussian Simulation: A Monte Carlo Method for Generating Models of Porosity and Permeability. In *Generation, Accumulation and Production of Europe's Hydrocarbons III*; Spencer, A.M., Ed.; Springer: Berlin/Heidelberg, Germany, 1993; pp. 345–356.
- Ersoy, A.; Yünsel, T.Y. Geostatistical Conditional Simulation for the Assessment of the Quality Characteristics of Cayırhan Lignite Deposits. *Energy Explor. Exploit.* **2006**, *24*, 391–416. [[CrossRef](#)]
- Delbari, M.; Afrasiab, P.; Loiskandl, W. Using sequential Gaussian simulation to assess the field-scale spatial uncertainty of soil water content. *CATENA* **2009**, *79*, 163–169. [[CrossRef](#)]

15. Pinheiro, M.; Emery, X.; Miranda, T.; Lamas, L.; Espada, M. Modelling Geotechnical Heterogeneities Using Geostatistical Simulation and Finite Differences Analysis. *Minerals* **2018**, *8*, 52. [\[CrossRef\]](#)
16. Box, G.E.P.; Muller, M.E. A Note on the Generation of Random Normal Deviates. *Ann. Math. Statist.* **1958**, *29*, 610–611. [\[CrossRef\]](#)
17. Matheron, G. Principles of Geostatistics. *Econ. Geol.* **1963**, *58*, 1246–1266. [\[CrossRef\]](#)
18. Anyiam, O.A.; Andrew, P.J.; Okwara, I.C. Assessment of the heterogeneity and petrophysical evaluation of reservoirs in the “Akbar Field”, Niger Delta, Nigeria. *J. Pet. Explor. Prod. Technol.* **2017**, *7*, 1035–1050. [\[CrossRef\]](#)
19. Michie, E.A.H.; Haines, T.J. Variability and heterogeneity of the petrophysical properties of extensional carbonate fault rocks, Malta. *Pet. Geosci.* **2016**, *22*, 136–152. [\[CrossRef\]](#)
20. Mukerji, T.; Mavko, G.; Rio, P. Scales of Reservoir Heterogeneities and Impact of Seismic Resolution on Geostatistical Integration. *Math. Geol.* **1997**, *29*, 933–950. [\[CrossRef\]](#)
21. De Ros, L.F.; Scherer, C.M.S. Stratigraphic Controls on the Distribution of Diagenetic Processes, Quality and Heterogeneity of Fluvial-Aeolian Reservoirs from the Recôncavo Basin, Brazil. In *Linking Diagenesis to Sequence Stratigraphy*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2013; pp. 105–132. [\[CrossRef\]](#)
22. Oxford English Dictionary. *Heterogeneity*; Oxford University Press: Oxford, UK, 2014.
23. Fitch, P.J.R.; Lovell, M.A.; Davies, S.J.; Pritchard, T.; Harvey, P.K. An integrated and quantitative approach to petrophysical heterogeneity. *Mar. Pet. Geol.* **2015**, *63*, 82–96. [\[CrossRef\]](#)
24. Remy, N.; Boucher, A.; Wu, J. *Applied Geostatistics with SGeMS: A User's Guide*; Cambridge University Press: Cambridge, UK, 2009. [\[CrossRef\]](#)
25. Wackernagel, H. *Multivariate Geostatistics*, 3rd ed.; Springer: Berlin/Heidelberg, Germany, 2003; p. 388. [\[CrossRef\]](#)
26. Armstrong, M. Experimental Variograms. In *Basic Linear Geostatistics*; Springer: Berlin/Heidelberg, Germany, 1998; pp. 47–58. [\[CrossRef\]](#)
27. Ringrose, P.; Bentley, M. *Reservoir Model Design*, 1st ed.; Springer: Dordrecht, The Netherlands, 2015; p. 249. [\[CrossRef\]](#)
28. Gu, Y.; Rühaak, W.; Bär, K.; Sass, I. Using seismic data to estimate the spatial distribution of rock thermal conductivity at reservoir scale. *Geothermics* **2017**, *66*, 61–72. [\[CrossRef\]](#)
29. Rühaak, W.; Guadagnini, A.; Geiger, S.; Bär, K.; Gu, Y.; Aretz, A.; Homuth, S.; Sass, I. Upscaling thermal conductivities of sedimentary formations for geothermal exploration. *Geothermics* **2015**, *58*, 49–61. [\[CrossRef\]](#)
30. Rühaak, W. 3-D interpolation of subsurface temperature data with measurement error using kriging. *Environ. Earth Sci.* **2015**, *73*, 1893–1900. [\[CrossRef\]](#)
31. Bailey, T.; Gatrell, A. *Interactive Spatial Data Analysis*; Longman Group Limited: Harlow, UK, 1995; p. 432.
32. Journel, A.G. Nonparametric estimation of spatial distributions. *J. Int. Assoc. Math. Geol.* **1983**, *15*, 445–468. [\[CrossRef\]](#)
33. Goovaerts, P. *Geostatistics for Natural Resources Evaluation*; Oxford University Press: Oxford, UK, 1997.
34. Remy, N. *Algorithmic and Software Methods for a Better Integration of the Geological Information into Numerical Models*; Stanford University: Stanford, CA, USA, 2004.
35. Ökten, G.; Göncü, A. Generating low-discrepancy sequences from the normal distribution: Box—Muller or inverse transform? *Math. Comput. Model.* **2011**, *53*, 1268–1281. [\[CrossRef\]](#)
36. Celisse, A. Optimal cross-validation in density estimation with the L2-loss. *Ann. Stat.* **2014**, *42*, 1879–1910. [\[CrossRef\]](#)
37. Willmott, C.J.; Matsuura, K.; Robeson, S.M. Ambiguities inherent in sums-of-squares-based error statistics. *Atmos. Environ.* **2009**, *43*, 749–752. [\[CrossRef\]](#)
38. Chai, T.; Draxler, R.R. Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* **2014**, *7*, 1247–1250. [\[CrossRef\]](#)
39. Emery, X. Testing the correctness of the sequential algorithm for simulating Gaussian random fields. *Stoch. Environ. Res. Risk Assess.* **2004**, *18*, 401–413. [\[CrossRef\]](#)
40. Linsel, A.; Wiesler, S.; Hornung, J.; Hinderer, M. High-Resolution Analysis of the Physicochemical Characteristics of Sandstone Media at the Lithofacies Scale. *Solid Earth Discuss.* **2020**, *2020*, 1–28. [\[CrossRef\]](#)
41. Linsel, A. ApirsAL/GeoReVi: GeoReVi v1.0.0 Pre-Release. Available online: <https://zenodo.org/record/3541136#.XvhBb3ERWUk> (accessed on 20 December 2019). [\[CrossRef\]](#)

42. Becker, A.; Schwarz, M.; Schäfer, A. Lithostratigraphische Korrelation des Rotliegend im östlichen Saar-Nahe-Becken. *Jahresberichte Und Mitteilungen Des Oberrheinischen Geologischen Vereins* **2012**, *94*, 105–133. [\[CrossRef\]](#)
43. Aretz, A.; Bär, K.; Götz, A.E.; Sass, I. Outcrop analogue study of Permocarboniferous geothermal sandstone reservoir formations (northern Upper Rhine Graben, Germany): Impact of mineral content, depositional environment and diagenesis on petrophysical properties. *Int. J. Earth Sci.* **2015**, *105*, 1431–1452. [\[CrossRef\]](#)
44. Farkhutdinov, A.; Goblet, P.; de Fouquet, C.; Cherkasov, S. A case study of the modeling of a hydrothermal reservoir: Khankala deposit of geothermal waters. *Geothermics* **2016**, *59*, 56–66. [\[CrossRef\]](#)
45. Nordahl, K.; Ringrose, P.S. Identifying the Representative Elementary Volume for Permeability in Heterolithic Deposits Using Numerical Rock Models. *Math. Geosci.* **2008**, *40*, 753. [\[CrossRef\]](#)
46. Nordahl, K.; Messina, C.; Berland, H.; Rustad, A.B.; Rimstad, E.; Martinius, A.W.; Howell, J.A.; Good, T.R. Impact of multiscale modelling on predicted porosity and permeability distributions in the fluvial deposits of the Upper Lunde Member (Snorre Field, Norwegian Continental Shelf). In *Sediment-Body Geometry and Heterogeneity: Analogue Studies for Modelling the Subsurface*; Geological Society of London: London, UK, 2014; Volume 387, p. 25. [\[CrossRef\]](#)
47. Middleton, G.V. Sediment Deposition from Turbidity Currents. *Annu. Rev. Earth Planet. Sci.* **1993**, *21*, 89–114. [\[CrossRef\]](#)
48. Agemar, T.; Weber, J.; Schulz, R. Deep Geothermal Energy Production in Germany. *Energies* **2014**, *7*, 4397–4416. [\[CrossRef\]](#)
49. Filomena, C.M.; Hornung, J.; Stollhofen, H. Assessing accuracy of gas-driven permeability measurements: A comparative study of diverse Hassler-cell and probe permeameter devices. *Solid Earth* **2014**, *5*, 1–11. [\[CrossRef\]](#)
50. Bär, K. *Untersuchung der tiefeGeothermischen Potenziale von Hessen*; Technische Universität Darmstadt: Darmstadt, Germany, 2012; p. 297.
51. Massey, F.J. The Kolmogorov-Smirnov Test for Goodness of Fit. *J. Am. Stat. Assoc.* **1951**, *46*, 68–78. [\[CrossRef\]](#)
52. Simard, R.; Ecuyer, P. Computing the Two-Sided Kolmogorov-Smirnov Distribution. *J. Stat. Softw.* **2011**, *1*. [\[CrossRef\]](#)
53. Tukey, J. *Exploratory Data Analysis*; Addison-Wesley: Reading, MA, USA, 1977; p. 712.
54. Corbett, P.; Jensen, J.L. Estimating the mean permeability: How many measurements do you need? *First Break* **1992**, *10*, 5. [\[CrossRef\]](#)
55. Catmull, E. *A Subdivision Algorithm for Computer Display of Curved Surfaces*; University of Utah: Salt Lake City, Utah, 1974.
56. Journel, A.G. Geostatistics: Models and tools for the earth sciences. *Math. Geol.* **1986**, *18*, 119–140. [\[CrossRef\]](#)
57. Matheron, G. The intrinsic random functions and their applications. *Adv. Appl. Probab.* **1973**, *5*, 439–468. [\[CrossRef\]](#)
58. Mantoglou, A.; Wilson, J.L. The Turning Bands Method for simulation of random fields using line generation by a spectral method. *Water Resour. Res.* **1982**, *18*, 1379–1394. [\[CrossRef\]](#)
59. Emery, X.; Lantuéjoul, C. TBSIM: A computer program for conditional simulation of three-dimensional Gaussian random fields via the turning bands method. *Comput. Geosci.* **2006**, *32*, 1615–1628. [\[CrossRef\]](#)
60. Paravarzar, S.; Emery, X.; Madani, N. Comparing sequential Gaussian and turning bands algorithms for cosimulating grades in multi-element deposits. *Comptes Rendus Geosci.* **2015**, *347*, 84–93. [\[CrossRef\]](#)

Sample Availability: The investigated rock samples are available at the Institute of Applied Geosciences Darmstadt and can be requested under linsel@geo.tu-darmstadt.de. Moreover, the samples are registered in the System for Earth Sample Registration (SESAR, www.geosamples.org).



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).