Reconstruction of Three-Dimensional Aquifer Heterogeneity From Two-Dimensional Geophysical Data

Nils Gueting · Jef Caers · Alessandro Comunian · Jan Vanderborght · Andreas Englert

Received: date / Accepted: date

Abstract Suitable training images (TI) for multiple-point statistics (MPS) are difficult to identify in real-case three-dimensional applications, posing challenges for modelers trying to develop realistic subsurface models. This study demonstrates that two-dimensional geophysical images, when employed as training and conditioning data, can provide sufficient information for three-dimensional MPS simulations. The advantage of such data-driven approach is that it does not rely on any external (possibly inappropriate) TI. The disadvantage is that three-dimensional MPS simulations must be carried out based

N. Gueting
Agrosphere (IBG-3), Institute of Bio- and Geosciences, Forschungszentrum Jülich, Germany
Tel.: +49-2461-618663
Fax: +49-2461-611768
E-mail: n.gueting@fz-juelich.de

J. Caers
Geological Sciences, Stanford University, USA

A. Comunian
Dipartimento di Scienze della Terra, Universita degli Studi di Milano, Italy

J. Vanderborght
Agrosphere (IBG-3), Institute of Bio- and Geosciences, Forschungszentrum Jülich, Germany

A. Englert
Earth Sciences, Ruhr-University Bochum, Germany

on two-dimensional information. Three different approaches (two existing, one new) are tested to overcome this problem. The two existing approaches rely on the three-dimensional reconstruction of incomplete data sets and on sequential two-dimensional simulations, respectively. The third approach is a newly proposed combination of the two former approaches. The three approaches are applied to model the three-dimensional facies structure of an alluvial aquifer based on high resolution ground penetrating radar cross-sections. The quality of a simulation outcome is evaluated based on the similarity of its multiple-point histogram (MPH) with reference MPHs, derived from the geophysical images. This evaluation reveals that the first approach (three-dimensional reconstruction) performs well close to conditioning data, but farther away from the data the simulation results deteriorate. Quite conversely, the second approach (sequential two-dimensional) performs well when only few conditioning data exist, but with increasing simulation sequence the quality decreases. The newly proposed third approach integrates the benefits from both approaches and is found to reproduce the reference MPHs significantly better than each of the two approaches stand-alone.

Keywords Aquifer heterogeneity · Multiple-point statistics · Ground-penetrating radar

1 Introduction

Flow and transport processes in the subsurface are controlled by the complex heterogeneity found in most geologic media. Geostatistical simulation techniques aim at mimicking geological heterogeneity to derive subsurface models that can be used to simulate and predict, for example, groundwater flow and contaminant transport (e.g., Deutsch and Journel 1998). Considerable attention has been devoted to the development of multiple-point statistical

(MPS) simulation approaches (Guardiano and Srivastava 1993; Strebelle 2002; Straubhaar et al. 2011; Mariethoz et al. 2010; Mariethoz and Caers 2014) because such approaches are able to model complex geometries and connectivity patterns, that cannot be modeled using traditional variogram-based simulation techniques (Gómez-Hernández and Wen 1998; Western et al. 2001; Renard and Allard 2013). The basic idea of MPS simulations is to generate models by reproducing the multiple-point spatial statistics derived from a user-defined training image (TI). Because the spatial structure of the TI is reproduced in the model, the choice of the TI is of most critical importance. As a consequence, applying MPS to real field cases can be challenging, because finding an appropriate TI is not always straight-forward, especially for three-dimensional cases (Huysmans and Dassargues 2009; Le Coz et al. 2011).

One possible approach is to use some external, conceptual TI that is anticipated to reflect the type of spatial structures that (presumably) exist at the site being modeled. Conceptual models for different geological settings may be found in data bases (Comunian and Renard 2008; Mariethoz and Caers 2014). Alternatively, the TI itself may be modeled using object-based methods (Maharaja 2008) or process-imitating models (Comunian et al. 2013). If the geological setting is unclear, Hermans et al. (2015) proposed to allow multiple alternative TIs, and narrow down the range of initially plausible TIs by validation/falsification with geophysical data.

In this study, a different approach is taken and data is utilized, instead of an external conceptual TI, to directly inform MPS simulations. The motivation for this data-driven approach is that data, collected directly at the site of interest, contains information about the actual field; and therefore it may be advantageous to rely on the data rather than on some external TI that is possibly not representative for the actual site being modeled (Ortiz and Deutsch 2004; Mariethoz and Renard 2010). Until now, data-based TIs were

obtained mainly from outcrops (Bayer et al. 2011; Huysmans and Dassargues 2012; Pickel et al. 2015). However, outcrops are often not available and, at best, located at the outer boundaries of the domain to be modeled. In contrast, state-of-the-art geophysical imaging techniques allow imaging the subsurface structure with high resolution and coverage directly at the location of interest. For example, Gueting et al. (2017) used cross-borehole ground penetrating radar (GPR) to characterize the spatial distribution of lithological facies in an alluvial aquifer. By applying a high resolution full-waveform inversion to analyze the tomograms from multiple crosshole planes, they were able to map the distribution of facies with a decimeter scale resolution, along cross-sections up to 50 m long and 10 m deep.

The present study builds upon the work by Gueting et al. (2017) and addresses the question if the GPR cross-sections provide enough information to reconstruct the three-dimensional aquifer structure using MPS. For this purpose, the facies distribution along the GPR cross-sections is assumed to reflect the true aquifer structure; the goal is to reconstruct the three-dimensional volume between these cross-sections. The main problem encountered with this approach is that three-dimensional MPS simulations have to be carried out on the basis of two-dimensional information. This kind of problem has already been investigated by Mariethoz and Renard (2010) and Comunian et al. (2012) who developed two approaches that enable three-dimensional MPS simulations but do not require a three-dimensional TI. Both approaches have proven their ability to generate three-dimensional models when only two-dimensional data are available (Mariethoz and Renard 2010; Comunian et al. 2012; Bayer et al. 2015; Pickel et al. 2015), but only little is known about when these approaches work or fail, and under which conditions one approach should be preferred over the other. In the present study, the performance of the two approaches is quantitatively compared based on a proxy for simulation quality that is

derived from the similarity of simulated and observed spatial patterns. Using this proxy, the benefits and limitations of the two existing approaches are illustrated and a new approach is proposed that integrates the benefits from both approaches.

2 Material and Methods

2.1 Study Site and GPR Data Set

The Krauthausen test site, set up by the research center Jülich in 1993 (Vereecken et al. 2000), is situated in the southern part of the Lower Rhine Embayment, close to the city of Düren. The 200 m × 50 m test site is equipped with a total number of 76 boreholes that reach down to the base of a shallow unconfined aquifer at approximately 10 m depth. The aquifer is composed of alluvial terrace sediments that were deposited by a local braided river system on top of older Rhine and Maas sediments (Englert 2003). The aquifer material comprises sand and gravel in varying proportions, with only small amounts of clay.

In 2013, an extensive cross-borehole GPR survey was conducted in the central part of the test site (Gueting et al. 2015, 2017). The GPR data were inverted for dielectric permittivity and electrical conductivity using a recently developed full-waveform inversion approach that is preferable, with respect to spatial resolution, over conventional ray-based inversion approaches (Klotzsche et al. 2010, 2013). To translate the electrical properties derived from GPR into hydrogeologically meaningful facies, Gueting et al. (2017) conducted a logistic regression of GPR results and co-located grain size data, and used the fitted logistic regression model to predict the distribution of lithological facies along the GPR transects. A three-dimensional plot of the obtained facies distribution along the various two-dimensional GPR transects is shown in Fig. 1.

Overall, the facies are arranged in layer-like structures that appear plausible for a braided river alluvial setting. Note that despite the generally layered architecture, significant lateral variations in number and thickness of layers can be observed, which indicates that conceptualizing the aquifer as a perfectly stratified medium (as one may be tempted to do if only one-dimensional vertical borehole logs were available) would be an oversimplification of the true aquifer structure. This highlights the value of crosshole geophysical imaging techniques, which yield two-dimensional information, along the vertical and the lateral direction.

Vertical profiles of porosity and hydraulic conductivity from direct-push tests, flowmeter measurements and grain size data, were used to hydrogeologically characterize the GPR facies. Table 1 gives the porosity and hydraulic conductivity estimates for each facies, according to the different methods. Although the absolute values from the different methods show some disagreement, there is a clear and consistent trend that the hydraulic conductivity increases from facies 1 to facies 2 to facies 3, whereas the porosity decreases. Gueting et al. (2017) compared the GPR derived facies distribution with the tracer breakthrough observed by Müller et al. (2010) and found a clear effect of the facies structure on the observed transport, which confirmed the hydrogeological relevance of the GPR derived facies. The present study takes the two-dimensional facies structure along the GPR transects obtained by Gueting et al. (2017) as given, and addresses the problem of estimating the three-dimensional facies structure. In other words, the facies distribution along the two-dimensional GPR transects (Fig. 1) is assumed to reflect the true aquifer structure; the goal is to reconstruct the three-dimensional volume between these transects.

2.2 Multiple-Point-Statistical Simulations

2.2.1 Basic Principle

This section briefly reviews the principle idea of multiple-point-statistical (MPS) simulation and its implementation in the impala algorithm (Straubhaar et al. 2011) and in the direct sampling algorithm (Mariethoz et al. 2010), which were used in the present study.

Consider an attribute S with K possible states. S may represent a categorical variable which identifies, for instance, the occurrence of different geological facies; it may also represent a continuous variable, such as porosity or permeability, whose range was divided into K classes using some kind of classification technique. The objective of MPS simulations is to assign a value $S(u)$ to each node u in a two-dimensional or three-dimensional simulation grid, by taking into account the spatial position of u relative to already informed nodes (i.e., conditioning data or already simulated nodes). The data values and geometry of N informed nodes in the neighborhood of u can be expressed as the data event d_u

$$d_u = \{S(u + h_1), ..., S(u + h_N)\}, \tag{1}$$

where $h_{1,...,N}$ are two-dimensional or three-dimensional vectors that describe the spatial position of the informed grid nodes with respect to the central node u. To simulate $S(u)$ for a given data event d_u, one is interested in the conditional probability distribution function (cpdf)

$$\text{Prob}\{S(u) = k \mid d_u\} = f(u; k \mid d_u), \qquad k = 1, 2, ..., K. \tag{2}$$

The basic idea of MPS is to infer this cpdf from a user-defined training image (TI) that is supposed to be representative for the geological setting being modeled (Strebelle 2002).

In the impala algorithm (Straubhaar et al. 2011), the cpdf is built by scanning the TI prior to simulation, and storing the number of occurrences of individual patterns in a catalog. Then the nodes in the simulation grid are sequentially visited and, for each node u (characterized by the data event d_u), the number of replicates of d_u in the TI is retrieved from the catalog, and a value for $S(u)$ is drawn randomly based on the histogram of the central node values for all replicates found in the TI.

A different strategy is used in the direct sampling algorithm (Mariethoz et al. 2010). Instead of explicitly building the cpdf by scanning the TI prior to the simulation, this step is skipped by directly sampling the TI during the simulation. As soon as a TI pattern is found which either matches exactly d_u, or whose mismatch is below a specified threshold, the scanning is stopped and the central node of the found TI pattern is directly pasted into the simulation grid. Because the TI is scanned randomly, this strategy is completely equivalent to drawing a random value from the cpdf, yet it increases simulation speed.

2.2.2 Three-dimensional Reconstruction Approach

Recognizing that an important practical limitation of MPS modeling consists in finding an appropriate TI, Mariethoz and Renard (2010) developed a data-driven reconstruction approach that uses multiple-point statistics but does not require a TI. Instead of extracting multiple-point statistics from a TI, their approach extracts multiple-point statistics directly from the available data (i.e., the informed grid nodes), which serve both as training and conditioning data.

In the present study, the approach by Mariethoz and Renard (2010) was applied, from here on called the "three-dimensional reconstruction approach", to derive three-dimensional facies models for the Krauthausen aquifer using the GPR derived facies distribution along the various two-dimensional crosshole planes as training and conditioning data (Fig. 1). For the simulation, the approximately 25 m × 50 m × 6 m aquifer volume shown in Fig. 1 is discretized on a grid of 131 × 244 × 30 nodes. This results in approximately 10^6 nodes, of which about 2% are informed by the GPR planes. The simulations were carried out with the direct sampling algorithm using a maximum neighborhood of 35 nodes, an acceptance threshold of 0 (this means that the TI is initially scanned for patterns that exactly match the data event; if no such pattern is found, the TI pattern with the lowest mismatch is accepted) and a simulation path that ensures that only nodes with at least one informed neighbor are simulated. This parametrization was chosen based on the examples given by Mariethoz and Renard (2010) who used the same parameter values for their real-case three-dimensional application (Example 3.4 in their paper).

2.2.3 Sequential Two-Dimensional Approach

In field studies, providing a three-dimensional training image will often be more challenging than providing a two-dimensional training image because two-dimensional information can often be obtained with less effort, for example, from outcrops or from geophysical surveys. This has motivated Comunian et al. (2012) to develop a method that allows conducting three-dimensional MPS simulations on the basis of two-dimensional training images. Their approach requires two perpendicular two-dimensional training images, which can be derived in practice, for example, from two outcrops with approximately perpendicular orientation (Huysmans and Dassargues 2012; Kessler et al. 2013; Pickel et al. 2015; Bayer et al. 2015). These TIs are used to perform multi-

ple sequential two-dimensional MPS simulations in a three-dimensional grid. Because the simulations are conducted along perpendicular directions, the simulated slices cross each other, and the nodes along the intersection lines with already simulated slices can be used as conditioning data for the following two-dimensional simulations. In this way, a three-dimensional volume is eventually generated "slice-by-slice" without performing an actual three-dimensional simulation at any time. As a consequence, their approach requires only two-dimensional TIs instead of three-dimensional TIs.

In the present study, two perpendicular TIs were derived from the longest longitudinal GPR transect and the longest transverse GPR transect, respectively (Fig. 2). The two-dimensional TIs were used to apply the approach by Comunian et al. (2012), from here on called the "sequential two-dimensional approach", to simulate the facies distribution in the $131 \times 244 \times 30$ nodes grid introduced already in the previous section. Following Comunian et al. (2012), the simulations were carried out with the impala algorithm using 3 multigrids and z as an auxiliary variable. Figure 3 illustrates the simulation sequence for the first 10 slices. At the beginning, the four outer surfaces are simulated (black lines). In the following, the simulation proceeds by dividing the remaining "empty spaces" into half (i.e., the slices along the central red lines are simulated next, followed by the slices along the blue lines). Note that while the TIs comprise only the two main GPR transects, all GPR transects are included as conditioning data (Fig. 1).

2.2.4 Combined Approach

As will be shown in Sects. 3.1 and 3.2, the results of this study suggest that the three-dimensional reconstruction approach is preferable over the sequential two-dimensional approach when dense conditioning data are available; but when sparse conditioning data are available, the sequential two-dimensional

approach is preferable over the three-dimensional reconstruction approach. With the goal to integrate the benefits from both approaches, a combined application of the two approaches is proposed. The basic idea is to employ each approach under the conditions where its performance is optimal: the sequential two-dimensional approach in the beginning, when only few conditioning data exist; the three-dimensional reconstruction approach later, when already simulated nodes contribute more and more conditioning data. The following 2-step workflow is proposed:

1. The first step consists of starting the sequential two-dimensional approach exactly in the same manner as already described. After a user-defined number of simulated slices, however, the sequential two-dimensional approach is stopped, and the partly informed simulation grid is saved. At this stage, the simulation grid contains (a) the initially available conditioning data and (b) the already simulated two-dimensional slices.
2. This is the starting point for the second step, which consists of running the three-dimensional reconstruction approach using the partly informed simulation grid from the first step as training and conditioning data.

Note that this workflow is not at all restricted to the data set used in the present study, but can be applied, without any additional implementations, whenever the sequential two-dimensional approach can be applied.

2.3 Evaluating the Quality of a Simulation Outcome

In the present study, the quality of a simulation outcome is evaluated based on its multiple-point histogram (MPH). The MPH indicates the frequency of spatial patterns in a grid, for a user-defined template (Boisvert et al. 2007). For example, for a variable with 3 possible states and a 3×3 nodes template, there would be a total number of 19,683 (3^9) possible patterns. The MPH

is obtained by scanning the entire grid for each pattern, and recording their number of occurrences (Boisvert et al. 2007).

MPS simulations aim at reproducing the spatial patterns found in a TI. Because grids with similar spatial patterns are characterized by similar MPHs, successful pattern reproduction will result in an MPH that is very similar to the MPH of the TI. However, as the simulation proceeds and the simulation grid gets increasingly filled by informed nodes, the simulation may produce data events that are incompatible with the TI. Such data events deteriorate the simulation and lead to a deviation of the simulated MPH from the MPH of the TI. Taking the MPH of the TI as a reference and quantifying the deviation of a simulated MPH from this reference MPH, thus provides information about the quality of a simulation outcome. A quantitative metric for the dissimilarity of two MPHs, p and q, is the Jensen-Shannon (JS) divergence (Tan et al. 2014)

$$d(p,q) = \frac{1}{2}\sum_i p_i \log\left(\frac{p_i}{q_i}\right) + \frac{1}{2}\sum_i q_i \log\left(\frac{q_i}{p_i}\right). \tag{3}$$

If more than two MPHs shall be compared, a distance matrix D can be obtained by computing Eq.(3) for all possible pairs of MPHs. In the present study, the MPH-distance to the GPR derived facies distribution is used to evaluate the quality of an MPS simulation outcome. More precisely, the 15 available GPR planes are used to obtain an ensemble of 15 reference MPHs, and it is assumed that the quality of a simulation outcome is reflected in the distance of its MPH to these reference MPHs. A small distance indicates consistent spatial structures, and thus, high-quality simulation results. A large distance indicates inconsistent structures, and thus, poor simulation results.

Because the simulation outcome is three-dimensional, but the GPR planes are two-dimensional, the simulated three-dimensional block is decomposed into two-dimensional vertical slices. As the simulation grid has dimensions of 131

$\times$ 244 $\times$ 30 nodes, this results in 131 slices along the longitudinal direction (yz) and in 244 slices along the transverse direction (xz). Because the goal is to investigate in detail if the simulation quality differs with the location in the simulation domain, each longitudinal slice is additionally split into 7 segments, and each transverse slice is split into 4 segments. This results in a total number of 1,893 segments with a length of approximately 6 m. By calculating the MPH (for a 3 $\times$ 3 template) for each simulated segment and for each GPR plane, and by computing Eq.(3) for all possible pairs, a distance matrix D is obtained which describes the dissimilarity among all simulated segments and GPR planes. The information in D is visualized using multi-dimensional scaling (MDS) (Caers 2011). MDS approximates the MPH-distance by a lower-dimensional Euclidean distance in Cartesian space, which facilitates the visualization of D.

In addition, the minimal MPH-distance to a GPR plane is proposed as a quantitative proxy for the quality of a simulated segment. A small MPH-distance between a simulated segment and a GPR plane indicates consistent spatial patterns, and thus, high-quality simulation results. A large distance indicates inconsistent patterns, and thus, poor simulation results. As there are 15 GPR planes, there are 15 individual distances $d_{1,\ldots,15}$ available for each simulated segment (one distance to each GPR plane). The smallest distance $d_{\min} = \min(d_{1,\ldots,15})$ is proposed as a quality measure, because in this way a simulated segment is assigned a high quality if it is similar to any one of the GPR planes. Alternatively, one could use the sum of distances to all GPR planes. However, this would mean that if several GPR planes are similar to each other, a simulated segment which is similar to them is better than a simulated segment which is similar to a single GPR plane that itself is different from the other GPR planes.

3 Results

3.1 Three-Dimensional Reconstruction Approach (Stand-Alone)

The facies model obtained from the three-dimensional reconstruction approach is shown in Fig. 4. The simulated facies distribution exhibits three-dimensional structures that are locally consistent with the conditioning data, that is, the spatial structures in the GPR planes continue in the simulation grid (Fig. 4(b)). Overall, however, the facies distribution appears patchy. Disconnected structures and isolated pixels can be observed. Particularly further away from the conditioning GPR planes, the simulated facies structures seem to deteriorate. This is most obvious for the distribution of facies 3, which seems to follow the GPR transects in the simulation grid (Fig. 4(a)). In the rest of the simulation domain, facies 3 is less abundant. This is also reflected in the global facies proportions (Table 2). While facies 3 accounts for 14% of the conditioning data, it accounts for only 7% of the simulated three-dimensional volume. For all three facies, the differences between reference and simulated facies proportions sum up to 24%.

As described in Sect. 2.3, the consistency between the spatial structures in the simulation grid and the spatial structures in the GPR images is assessed by decomposing the simulated three-dimensional block into multiple two-dimensional segments and calculating the JS-divergence (Eq.(3)) between the MPHs of these segments and the MPHs of the GPR images. The resulting distance matrix is visualized using MDS. Figure 5 shows the MDS plot for the three-dimensional reconstruction result. The GPR images are plotted in red. The two-dimensional segments extracted from the simulated three-dimensional block are plotted in black. On the one hand, it can be observed that the distribution of black data points comprises the area where the GPR

images plot. This shows that some part of the simulated segments are consistent with the GPR data (meaning that they are characterized by an MPH similar to the MPH of the GPR images). On the other hand, the black data points are scattered over a much larger area, including regions far away from where the GPR images plot. These simulated segments are not consistent with the GPR data (i.e., they contain spatial structures which are not contained in the GPR images).

Figure 5 includes images of three individual two-dimensional segments that plot in different regions of the MDS map. As expected, the spatial structures in the segments can be seen to deteriorate with increasing distance to the GPR planes in the MDS plot. This supports the idea that the quality of a simulated segment can be evaluated based on the segment's distance (in terms of the MPH) to GPR planes. As described in Sect. 2.3, the minimal MPH-distance to a GPR plane, $d_{\min}$, is used as a proxy for the quality of a simulated segment. Note that also the arithmetic and harmonic mean of the MPH-distances to GPR planes were tested as a quality proxy, but the results were very similar to the results obtained using the minimal distance $d_{\min}$. Here, only the results that were obtained with $d_{\min}$ are shown. Figure 6 shows in green and red color, the 33% of the longitudinal and transverse segments that are characterized by the smallest and largest $d_{\min}$-values, respectively. While the green segments are distributed closely around the GPR planes, the red slices tend to plot farther away from the GPR planes (e.g., in the corners of the domain). This suggests that the three-dimensional reconstruction approach yields better results close to the conditioning data than farther away from the conditioning data. This observation is in agreement with the findings of Mariethoz and Renard (2010), who discussed the benefits and limitations of the three-dimensional reconstruction approach and noted that only at locations where sufficient information are available, complex spatial structures can be inferred from the data (Mariethoz

and Renard 2010). The present study provides quantitative evidence for this behavior using the JS-divergence between simulated and observed MPH as a proxy for simulation quality.

3.2 Sequential Two-Dimensional Approach (Stand-Alone)

The three-dimensional facies model obtained from the sequential two-dimensional approach is shown in Fig. 7. Overall, the facies model is characterized by layer-like structures that appear to be connected across the entire simulation grid. The conditioning GPR planes are well integrated in the simulation grid (Fig. 7(b)). However, particularly in the inner volume of the three-dimensional block (Fig. 7(b)), noisy simulation results with unclean facies boundaries can be observed. Compared with the three-dimensional reconstruction result (Fig. 4), the facies distribution obtained from the sequential two-dimensional approach is more continuous. Most notably, a more or less continuous top layer of facies 3 is obtained, which was not obtained from the three-dimensional reconstruction approach. As a consequence, the global facies proportions obtained from the sequential two-dimensional approach better match the facies proportions of the conditioning data. The deviations for all three facies sum up to only 6%, which is a distinct improvement relative to the 24% deviation obtained with the three-dimensional reconstruction approach (Table 2).

Figure 8 shows the MDS plot for the sequential two-dimensional simulation results. The distribution of the data points shows that the ensemble of two-dimensional segments extracted from the three-dimensional simulated block comprises both segments that are consistent (in terms of the MPH) with the GPR planes, but also segments that are not consistent with the GPR planes. This is similar to what was observed for the three-dimensional reconstruction approach. What is different for the sequential two-dimensional simulation re-

sult, is how the consistent/inconsistent segments are distributed in the model domain. This is shown in Fig. 9(a), where again the 33% of the longitudinal and transverse segments with the smallest and largest $d_{\min}$-values are plotted in green and red color, respectively. In contrast to the three-dimensional reconstruction result, where the green segments were clearly aligned along the GPR planes (Fig. 6), here, the green and red segments are quite evenly distributed over the entire model domain.

The quality of the sequential two-dimensional simulation results, thus, does not seem to be controlled by the proximity to conditioning data. Instead, the quality seems to be a function of the simulation sequence (Fig. 9(b)). This can be explained as follows. The sequential two-dimensional approach generates three-dimensional volumes by performing multiple two-dimensional simulations, considering already simulated slices as conditioning data for the following simulations. With increasing simulation sequence, the two-dimensional simulations become more and more constrained by the already simulated pixels along the intersection lines with previously simulated slices. As discussed by Comunian et al. (2012) this increasing amount of conditioning data eventually produces data events that are incompatible with the TIs. Such incompatible data events deteriorate the simulation results and are the reason why a correlation between simulation sequence and quality can be seen (Fig. 9(b)).

In conclusion, the MPH based analysis of the facies models obtained from the three-dimensional reconstruction approach and from the sequential two-dimensional approach, highlighted some important differences between the two approaches. The three-dimensional reconstruction approach was found to yield good results close to the conditioning data, but further away from the conditioning data, the quality of the simulation results decreased. In contrast, the sequential two-dimensional approach was able to produce good results far away from the conditioning data, but with increasing simulation sequence (i.e.,

increasing amount of conditioning data), the quality of the simulation results decreased. So basically, while the presence of conditioning data improves the three-dimensional reconstruction results, it deteriorates the sequential two-dimensional simulation results. Given these complimentary requirements, it appears promising to combine the two approaches in order to benefit from the advantages of both approaches.

3.3 Combined Approach

As described in Sect. 2.2.4, a combined application of the two approaches was proposed by starting with the sequential two-dimensional approach, and after a couple of two-dimensional simulations, switching to the three-dimensional reconstruction approach. Obviously, the question arises regarding the optimal point to switch from the one approach to the other. This question will be addressed later in Sect. 4. First, the results obtained by choosing a switching point after 10 slices simulated by the sequential two-dimensional approach will be analyzed. Figure 10 shows the resulting three-dimensional facies model. Visual comparison with the models obtained from the two approaches stand-alone (Fig. 4 and Fig. 7) indicates that the combined approach might indeed be an improvement because the obtained model shows neither the disconnected patches nor the noisy facies boundaries, which were the main issues in the models obtained from the two approaches stand-alone. The combined approach also better reproduces the reference facies proportions (Table 2).

To compare the facies models from the three different approaches in more detail, MDS plots were constructed by calculating the MPH-distance between the two-dimensional segments extracted from all three models and plotting them together in the same MDS plot. The resulting MDS maps are shown in Fig. 11(a). Individual segments from the three models are compared in

Fig. 11(b). Overall, it can be observed that the segments obtained from the combined approach tend to plot closer to the GPR planes in the MDS maps than the segments obtained from the other two approaches. This suggests that, of the three approaches, the combined approach produces spatial structures that best resemble the GPR images. However, due to the strongly overlapping data points, it is difficult to decide on the basis of the MDS maps if one approach outperforms another. Moreover, the distances in the MDS maps are not the actual distances but their Euclidean approximations. It appears more appropriate to use the actual distances for quality assessment. Recall that in Sects. 3.1 and 3.2 the actual distance $d_{\min}$ was used to identify differences in the quality of segments belonging to the same three-dimensional model (Figs. 6 and 9). Now, $d_{\min}$ will be used to assess differences in the quality of segments belonging to different three-dimensional models. Figure 12 shows $d_{\min}$ for all segments obtained from (a) the three-dimensional reconstruction approach, (b) the sequential two-dimensional approach and (c) the combined approach. By sorting the individual segments (x-axis) with respect to $d_{\min}$, distribution curves for $d_{\min}$ are obtained that allow for a quantitative comparison of the quality of the three different facies models. Because low values for $d_{\min}$ express a strong statistical similarity between the simulated spatial structures and the structures in a GPR image, they indicate high quality simulation results. In contrast, high values for $d_{\min}$ indicate low quality simulation results. It is obvious in Fig. 12, that the combined approach leads to the distribution curve that is characterized by the lowest $d_{\min}$ values. This shows that the combined approach produces spatial structures that are in better agreement with the GPR images than both the three-dimensional reconstruction and the sequential two-dimensional approach stand-alone. The distinct improvement is particularly intriguing because the $d_{\min}$-curves from the two approaches stand-alone look quite similar (with slight advantages for the sequential two-

dimensional approach, though). Due to the similarity of the two curves, it appears useless to combine the two approaches to get a better result. But the point is that due to the opposite effect that conditional data has on the two approaches (Figs. 6 and 9), the two approaches produce "good" and "bad" segments in different spatial parts of the simulation domain. Therefore, the effect of combining the two approaches is that the "bad" segments of the one approach are replaced by the "good" segments of the other approach. This eventually leads to the distinct improvement that can be observed in Fig. 12 for the combined approach.

4 Discussion

4.1 General Validity of Findings

In Sect. 3, one simulation outcome obtained from each approach was analyzed. But MPS offers a stochastic modeling framework, that is, one simulation outcome is only one out of multiple realizations that can be obtained from each approach. Therefore, it is important to test if the results are valid for other realizations. Figure 13 shows the distribution of $d_{\min}$ for three realizations obtained from each approach. The distribution curves from the individual realizations are strikingly similar. Note that this similarity is not a consequence of nearly identical three-dimensional facies models. In fact, the spatial arrangement of facies in the simulation grid looks quite different from one realization to the other (not shown here). Nevertheless, in terms of $d_{\min}$, which expresses the MPH-based similarity to the GPR images regardless of the exact location of specific structures, the realizations yield nearly identical results. This finding strongly suggests that the results presented in Sect. 3 are generally valid, even though a number of only three realizations is arguably not sufficient to ascertain general validity.

4.2 Optimal Switching Point

As foreshadowed in Sect. 3.3, it will be investigated now if there is an optimal point to switch from the sequential two-dimensional approach to the three-dimensional reconstruction approach. Principally, an optimal switching point should be not too early, because otherwise there is not sufficient conditioning data available for the three-dimensional reconstruction approach, but also not too late, because otherwise flawed simulation results, obtained with increasing simulation sequence from the sequential two-dimensional approach, are irrevocably included into the simulation grid. In fact, avoiding the latter is critical because, once simulated, defective structures get reproduced by the three-dimensional reconstruction approach, which uses the already simulated slices as training data. To identify an optimal switching point, the performance of the combined approach was tested for different switching points. This was achieved by running the sequential two-dimensional approach and saving the partly informed simulation grid after 4, 10, 18, 34, 66 and 130 simulated slices (Fig. 14(a-f)). The partly informed grids were then used as starting points for the three-dimensional reconstruction approach.

Figure 15 shows the distribution of $d_{\min}$ obtained for the different switching points S_P. While relatively high values for $d_{\min}$ are obtained for very early and late switching points, lower values for $d_{\min}$ are obtained for the intermediate switching points. This shows that the intermediate switching points yield results that are in better agreement (in terms of the MPH) with the GPR images than the results obtained with the early and late switching points.

What is additionally included in red color in Fig. 15 is the distribution of $d_{\min}$ derived for the GPR images themselves by calculating for each GPR image the minimal JS-divergence to another GPR image. To enable the comparison of the distribution of $d_{\min}$ for simulation results and GPR images, the x-axis

in Fig. 15 is expressed as the rank number n_i of the sorted segments/images divided by the total number n_{tot} of segments/images (i.e., n_{tot}=1,893 for the simulated segments and n_{tot}=15 for the GPR images). This comparison is based on the idea that lower values for d_{min} might not always indicate better simulation results. Consider the extreme case that d_{min} is zero for each two-dimensional segment extracted from a simulated three-dimensional model. In this case, each two-dimensional segment would be a mere copy of one of the GPR images. Clearly, this should not be considered a good simulation result. If it is assumed that the 15 GPR images are representative for the entire aquifer, then the distribution of d_{min} for the GPR images can be considered a reference distribution, and the distribution of d_{min} for the simulated segments should ideally follow this reference distribution. In Fig. 15, the majority of d_{min} values obtained from the GPR images plot below the distribution curves from the simulations, which shows that the similarity among the GPR images is more pronounced than the similarity between simulation results and GPR images. Nevertheless, the GPR image based d_{min} distribution is not distinctly different from the simulation results, and particularly the simulation results obtained with the switching point S_P=10 show a fairly good match with the GPR image based reference distribution.

In conclusion, the results suggest that for the specific case considered here, the optimal point to switch from the sequential two-dimensional approach to the three-dimensional reconstruction approach is when approximately 10 to 34 two-dimensional slices have been simulated, which corresponds to approximately 5% to 20% of the nodes in the three-dimensional simulation grid. However, this percentage may depend on the shape and extension of the site-specific spatial structures, on the dimensions of the simulation grid, as well as on the amount and location of initially available conditioning data. Generally, the optimal switching point can always be found by trial-and-error as presented

in the present study. This procedure, however, is not very efficient because it involves performing and analyzing simulations for several switching points. A more elegant solution would be to define a stop criterion for the sequential two-dimensional approach using for example a threshold value that expresses a maximally acceptable MPH-distance between a simulated slice and the corresponding two-dimensional TI. As soon as the sequential two-dimensional approach produces a slice (or a specified number of slices) exceeding that threshold, the sequential two-dimensional approach could be automatically stopped and the simulation continued using the three-dimensional reconstruction approach. However, more research is required to properly define and implement such threshold.

4.3 Possible Future Research Directions

More research should be carried out to further test and optimize the combined approach presented here. One potential improvement would be to develop a strategy against incompatible "simulation fronts" originating from different conditioning slices. This problem was already recognized by Mariethoz and Renard (2010) who applied the three-dimensional reconstruction approach to reconstruct a synthetic three-dimensional test case and observed unrealistic sharp transitions in the simulated facies distribution at the equidistance of two parallel conditioning slices. They attributed these artifacts to the custom path (i.e., the sequence in which nodes are simulated) used for the reconstruction. This path randomly visits nodes that have at least one informed neighbor. This means that first the direct neighborhood of the conditioning data is simulated, then the direct neighborhood of the simulated nodes is simulated, etc. As a result, the simulation gradually moves away from the conditioning data. When simulated nodes originating from distant condition-

ing slices meet, incompatibilities can occur and lead to sharp transitions in the simulation grid. In the present study, the same custom path was used for the three-dimensional reconstruction approach, and occasionally, sharp transitions in the simulated structures between two conditioning slices could be observed. One example for such a sharp transition can be seen in Fig. 11(b), where the $1,000^{\text{th}}$ closest segment for the combined approach shows an upper blue facies layer that abruptly ends as if vertically cut. Probably, this is due to incompatible simulation fronts originating from conditioning slices located left and right of this segment. As noted by Mariethoz and Renard (2010), a possible strategy to avoid such incompatible simulation fronts would be to use larger neighborhoods for the simulations because this would increase the probability that simulated nodes coming from another slice belong to the neighborhood of the node to be simulated. Alternatively, one could use syn-processing (Mariethoz et al. 2010) to remove inconsistent patterns through recursive un- and re-simulation of nodes. Unfortunately, both strategies would greatly increase the computational demands. Another possibility would be to relax the condition for the path so that not only nodes in the direct neighborhood of informed nodes are simulated but also nodes which are some (small) distance away from informed nodes.

Another potential improvement lies in modifying the sequence of slices simulated with the sequential two-dimensional approach. In the present study, alternating simulations along xz and yz were performed to gradually fill up the simulation domain (compare Fig. 14). As noted by Comunian et al. (2012), it can be beneficial when conditioning data are available, to define a customized sequence by preferentially simulating slices that include as many conditioning data as possible. In addition, when conditioning data are available along two-dimensional planes (such as from GPR) it appears particularly useful to avoid the simulation of slices located in close distance to conditioning planes with

approximately parallel orientation, because in this case, it is very likely that the simulated slice and the conditioning plane contain inconsistent structures, which are difficult to "link" in the later simulations.

Finally, future studies should investigate the importance of the work presented here by conducting flow and transport simulations. In the present study, the quality of simulated facies models was evaluated based on spatial structures (i.e., based on the consistency, in terms of MPH, with available data). Often, the ultimate goal of a subsurface facies model is to enable flow and transport modeling (e.g., in the context of groundwater extraction, oil production, contaminant transport and remediation, etc.). Therefore, it would be interesting to evaluate the quality of a facies model on the basis of flow and transport simulations instead of its spatial structure only. In fact, the facies models derived in the present study provide an excellent opportunity to carry out such flow/transport based evaluations, because porosity and hydraulic conductivity values are available for each facies (Table 1) and it would be straight-forward to use them to parametrize the obtained facies models for flow and transport simulations. Simulation results could be compared with pumping and tracer test data available from previous studies at the Krauthausen site (Li et al. 2007, 2008; Müller et al. 2010). This comparison would provide the opportunity to verify/falsify individual facies models and, on this basis, decide which approach yields the most realistic facies distribution. In a similar manner, facies models from simpler modeling approaches (e.g., homogeneous, perfectly layered, multi-Gaussian) could be tested and used to evaluate the benefit of the MPS based modeling approach presented here.

5 Conclusions

The present study demonstrates that high resolution two-dimensional geophysical images can provide enough information to derive reasonable three-dimensional models of the subsurface using MPS simulations. The advantage of such data-driven approach is that no external TI (which possibly does not properly reflect the spatial structure of the actual field being modeled) is required. The disadvantage is that three-dimensional MPS simulations have to be carried out on the basis of two-dimensional information. Three different approaches (two existing, one new) were tested to overcome this problem. To quantitatively compare the performance of the three approaches, the quality of a simulation outcome is approximated by the similarity of its multiple-point histogram (MPH) with reference MPHs, derived from the geophysical images. In this way, a simulation outcome is considered as "good" if the simulated spatial patterns are consistent with the type of spatial patterns observed along the geophysical transects. This evaluation reveals that the three-dimensional reconstruction approach proposed by Mariethoz and Renard (2010) produces good results close to the conditioning data, but farther away from the conditioning data the simulation results deteriorate. Quite conversely, the sequential two-dimensional approach proposed by Comunian et al. (2012) produces good results when only few conditioning data exist, but with increasing simulation sequence the quality decreases because the simulations become too strongly constrained by conditioning data.

To integrate the benefits from both approaches, a new combination of the two existing approaches is proposed. The combination is achieved by starting with the sequential two-dimensional approach, and after a couple of two-dimensional simulations, switching to the three-dimensional reconstruction approach. In this way, each approach is employed under the conditions where

its performance is optimal: the sequential two-dimensional approach is used when only few conditioning data exist; the three-dimensional reconstruction approach is used when more and more conditioning data become available. The results clearly show that the combined approach is able to reproduce the reference MPHs better than each of the two approaches stand-alone. As the combined approach does not require any additional implementations, it is easy to use and straight-forward to apply. The only requirement for its applicability is that suitable two-dimensional training images are available (i.e., similar requirements as for the sequential two-dimensional approach). More research should be carried out to further test and optimize the combined approach presented here. In particular, its benefits should be evaluated on the basis of flow and transport simulations.

Acknowledgements This work was supported in part by the TERrestrial ENvironmental Observatories (TERENO) and in part by the Transregional Collaborative Research Centre 32 (TR32) Patterns in Soil-Vegetation-Atmosphere Systems: Monitoring, Modelling, and Data Assimilation. We want to thank Philippe Renard, Julien Straubhaar and Gregoire Mariethoz for fruitful discussions and for providing the MPS algorithms impala and direct sampling. We also want to thank two anonymous reviewers for their valuable suggestions.

References

Bayer P, Huggenberger P, Renard P, Comunian A (2011) Three-dimensional high resolution fluvio-glacial aquifer analog: Part 1: Field study. J Hydrol 405(1-2):1–9

Bayer P, Comunian A, Höyng D, Mariethoz G (2015) High resolution multi-facies realizations of sedimentary reservoir and aquifer analogs. Sci Data 2:150033

Boisvert JB, Pyrcz MJ, Deutsch CV (2007) Multiple-point statistics for training image selection. Nat Resour Res 16(4):313–321

Caers J (2011) Modeling Uncertainty in the Earth Sciences. Wiley

Comunian A, Renard P (2008) Introducing wwhypda: A world-wide collaborative hydrogeological parameters database. Hydrogeol J 17(2):481–489

Comunian A, Renard P, Straubhaar J (2012) 3D multiple-point statistics simulation using 2D training images. Comput Geosci 40:49–65

Comunian A, Jha SK, Giambastiani BMS, Mariethoz G, Kelly BFJ (2013) Training Images from Process-Imitating Methods. Math Geosci 46(2):241–260

Deutsch C, Journel A (1998) GSLIB: Geostatistical Software Library and Users Guide, 2nd Ed. Oxford University Press

Englert A (2003) Measurement, Estimation and Modelling of Groundwater Flow Velocity at Krauthausen Test Site. Dissertation, RWTH Aachen

Gómez-Hernández JJ, Wen Xh (1998) To be or not to be multi-Gaussian? A reflection on stochastic hydrogeology. Adv Water Resour 21(1):47–61

Guardiano FB, Srivastava RM (1993) Multivariate Geostatistics: Beyond Bivariate Moments. In: Soares, A (ed) Geostatistics Tróia '92: Volume 1, Springer Netherlands, Dordrecht, pp 133–144

Gueting N, Klotzsche A, van der Kruk J, Vanderborght J, Vereecken H, Englert A (2015) Imaging and characterization of facies heterogeneity in an alluvial aquifer using GPR full-waveform inversion and cone penetration tests. J Hydrol 524:680–695

Gueting N, Vienken T, Klotzsche A, van der Kruk J, Vanderborght J, Caers J, Vereecken H, Englert A (2017) High resolution aquifer characterization using crosshole GPR full-waveform tomography: Comparison with direct-push and tracer test data. Water Resour Res 53, DOI 10.1002/2016WR019498

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

Huysmans M, Dassargues A (2009) Application of multiple-point geostatistics on modelling groundwater flow and transport in a cross-bedded aquifer (Belgium). Hydrogeol J 17(8):1901–1911

Huysmans M, Dassargues A (2012) Modeling the effect of clay drapes on pumping test response in a cross-bedded aquifer using multiple-point geostatistics. J Hydrol 450-451:159–167

Kessler T, Comunian A, Oriani F (2013) Modeling Fine-Scale Geological Heterogeneity-Examples of Sand Lenses in Tills. Ground Water 51(5):692–705

Klotzsche A, van der Kruk J, Meles GA, Doetsch J, Maurer H, Linde N (2010) Full-waveform inversion of cross-hole ground-penetrating radar data to characterize a gravel aquifer close to the Thur River, Switzerland. Near Surf Geophys 8:635–649

Klotzsche A, van der Kruk J, Linde N, Doetsch J, Vereecken H (2013) 3-D characterization of high-permeability zones in a gravel aquifer using 2-D crosshole GPR full-waveform inversion and waveguide detection. Geophys J Int 195(2):932–944

Le Coz M, Genthon P, Adler PM (2011) Multiple-Point Statistics for Modeling Facies Heterogeneities in a Porous Medium: The Komadugu-Yobe Alluvium, Lake Chad Basin. Math Geosci 43(7):861

Li W, Englert A, Cirpka OA, Vanderborght J, Vereecken H (2007) Two-dimensional characterization of hydraulic heterogeneity by multiple pumping tests. Water Resour Res 43, DOI 10.1029/2006WR005333

Li W, Englert A, Cirpka OA, Vereecken H (2008) Three-dimensional geostatistical inversion of flowmeter and pumping test data. Ground Water 46(2):193–201

Maharaja A (2008) TiGenerator: Object-based training image generator. Comput Geosci 34(12):1753–1761

Mariethoz G, Caers J (2014) Multiple-point Geostatistics: Stochastic Modeling with Training Images. Wiley-Blackwell

Mariethoz G, Renard P (2010) Reconstruction of incomplete data sets or images using direct sampling. Math Geosci 42(3):245–268

Mariethoz G, Renard P, Straubhaar J (2010) The Direct Sampling method to perform multiple-point geostatistical simulations. Water Resour Res 46(11), DOI 10.1029/2008WR007621

Müller K, Vanderborght J, Englert A, Kemna A, Huisman JA, Rings J, Vereecken H (2010) Imaging and characterization of solute transport during two tracer tests in a shallow aquifer using electrical resistivity tomography and multilevel groundwater samplers. Water Resour Res 46(3), DOI 10.1029/2008WR007595

Ortiz JM, Deutsch CV (2004) Indicator simulation accounting for multiple-point statistics. Mathematical Geology 36(5):545–565

Pickel A, Frechette JD, Comunian A, Weissmann GS (2015) Building a training image with Digital Outcrop Models. J Hydrol 531:53–61

Renard P, Allard D (2013) Connectivity metrics for subsurface flow and transport. Adv Water Resour 51:168–196

Straubhaar J, Renard P, Mariethoz G, Froidevaux R, Besson O (2011) An improved parallel multiple-point algorithm using a list approach. Math Geosci 43(3):305–328

Strebelle S (2002) Conditional simulation of complex geological structures using multiple-point statistics. Mathematical Geology 34(1):1–22

Tan X, Tahmasebi P, Caers J (2014) Comparing training-image based algorithms using an analysis of distance. Math Geosci 46(2):149–169

Vereecken H, Döring U, Hardelauf H, Jaekel U, Hashagen U, Neuendorf O, Schwarze H, Seidemann R (2000) Analysis of solute transport in a heterogeneous aquifer: the Krauthausen field experiment. J Contam Hydrol 45:329–358

Western AW, Böschl G, Grayson RB (2001) Toward capturing hydrologically significant connectivity in spatial patterns. Water Resour Res 37(1):83–97

Table 1 Porosity n and log hydraulic conductivity Y of the GPR facies according to different methods (CPT = Cone Penetration Test; GPR = Ground Penetrating Radar; DPST = Direct-Push Slug Test; DPIL = Direct-Push Injection Log; FM = Flowmeter; GSD = Grain Size Distribution). Details on the methods and their application at the Krauthausen site can be found in Gueting et al. (2017)

Facies	n_{CPT}	n_{GPR}	Y_{DPST}	Y_{DPIL}	Y_{FM}	Y_{GSD}
1 (sand)	0.27	0.29	-3.65	-3.62	-2.81	-3.78
2 (sandy gravel)	0.24	0.25	-3.48	-3.60	-2.50	-3.09
3 (gravel)	0.17	0.18	-3.18	-3.42	n.a.	-2.14

Table 2 Facies proportions for the conditioning data and for the simulated three-dimensional models shown in Figs. 4(a), 7(a) and 10(a)

Data / Simulation result	Facies 1	Facies 2	Facies 3	Deviation[a]
GPR conditioning data (reference)	36 %	50 %	14 %	-
Three-dimensional reconstruction approach	31 %	62 %	7 %	24 %
Sequential two-dimensional approach	38 %	51 %	11 %	6 %
Combined approach	38 %	49 %	13 %	4 %

[a]Total deviation from reference (i.e., the sum of deviation over all facies)

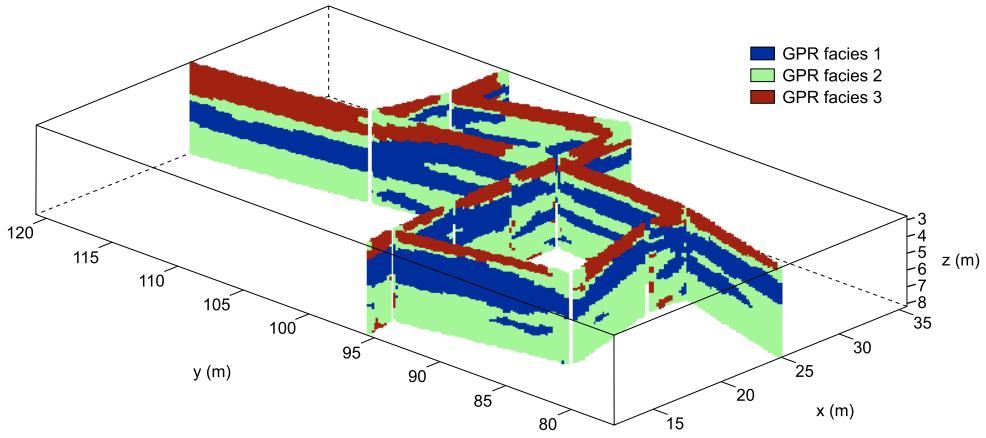

Fig. 1 GPR derived facies distribution along 15 individual two-dimensional cross-borehole planes. Details on the inversion of the GPR data and the classification into facies are given by Gueting et al. (2017)

Fig. 2 Two-dimensional training images along (**a**) yz direction, and (**b**) xz direction, used for the sequential two-dimensional approach

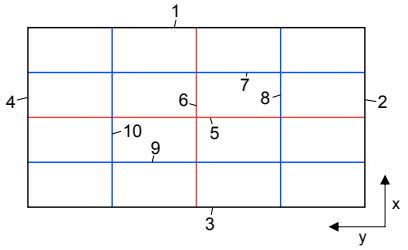

Fig. 3 Simulation sequence for the first 10 slices using the sequential two-dimensional approach (top view). The numbers refer to the order in which individual slices are simulated. Each two-dimensional simulation considers the nodes along the intersection lines with already simulated slices as conditioning data

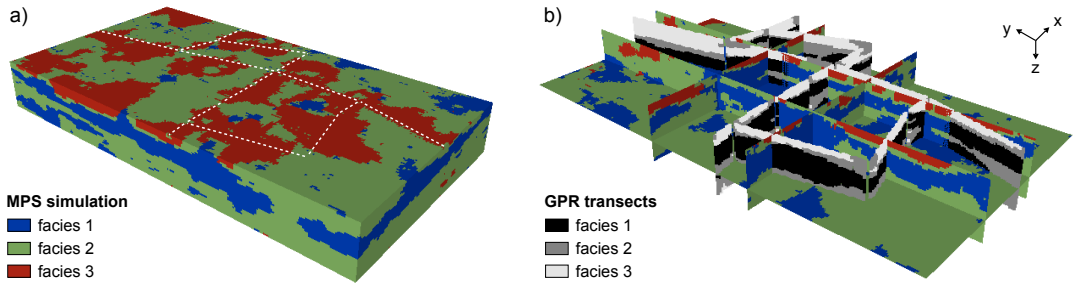

Fig. 4 Result of the three-dimensional reconstruction approach. (**a**) Block view of the simulated three-dimensional volume. White dashed lines show the locations of the GPR transects, which were used as training and conditioning data. (**b**) Cross-sectional slices and comparison with the GPR planes. For clarity, a different color-scale is prescribed to simulated nodes and to GPR informed nodes

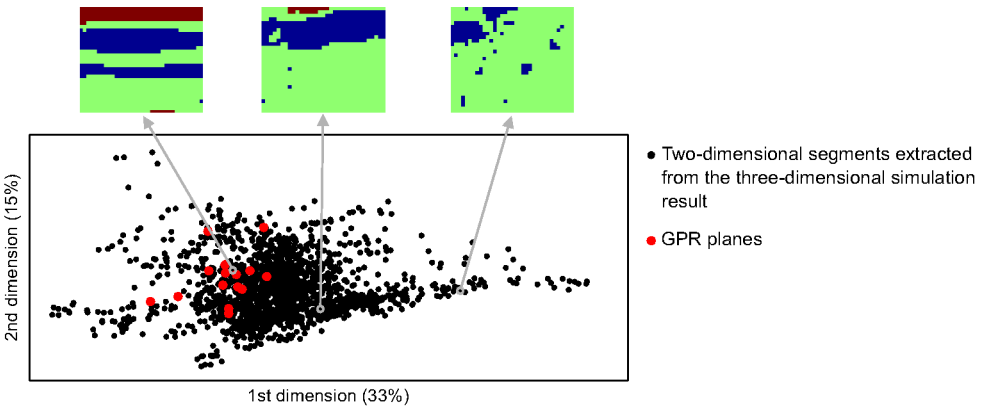

Fig. 5 MDS plot for the three-dimensional reconstruction approach. The distance is the JS-divergence between the MPHs of simulated segments/GPR planes. Percentages in parenthesis quantify the amount of variance explained by a dimension

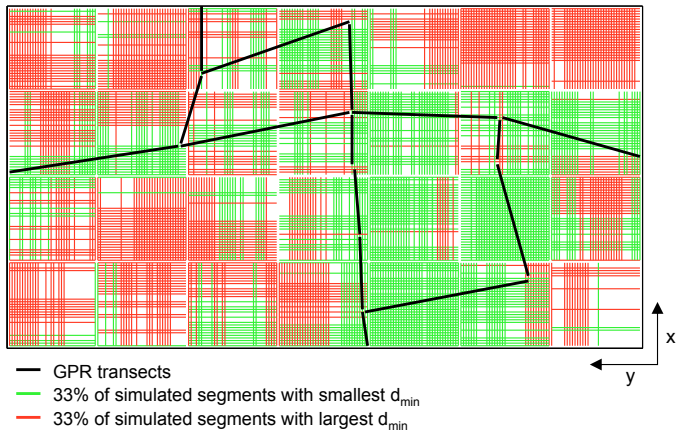

Fig. 6 Spatial distribution of two-dimensional segments with "high quality" (green) and "low quality" (red) derived from the three-dimensional reconstruction approach. As a measure of quality, the minimal distance $d_{\min}$ between the MPH of a simulated segment and the MPH of a GPR plane is used

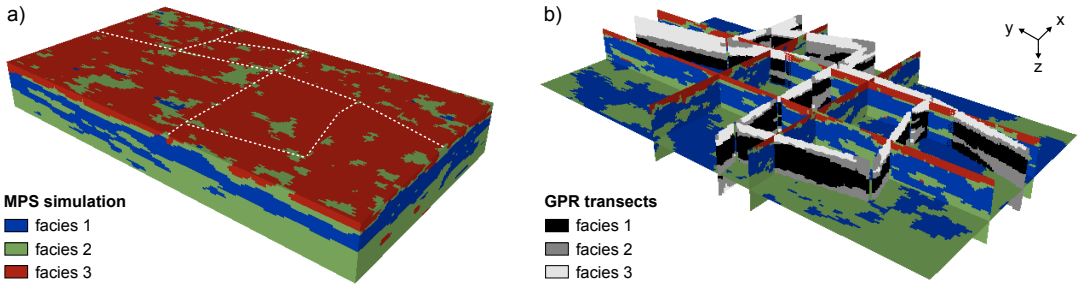

Fig. 7 Result of the sequential two-dimensional approach. (**a**) Block view of the simulated three-dimensional volume. White dashed lines show the locations of the GPR transects, which were used as conditioning data. (**b**) Cross-sectional slices and comparison with the GPR planes. For clarity, a different color-scale is prescribed to simulated nodes and to GPR informed nodes

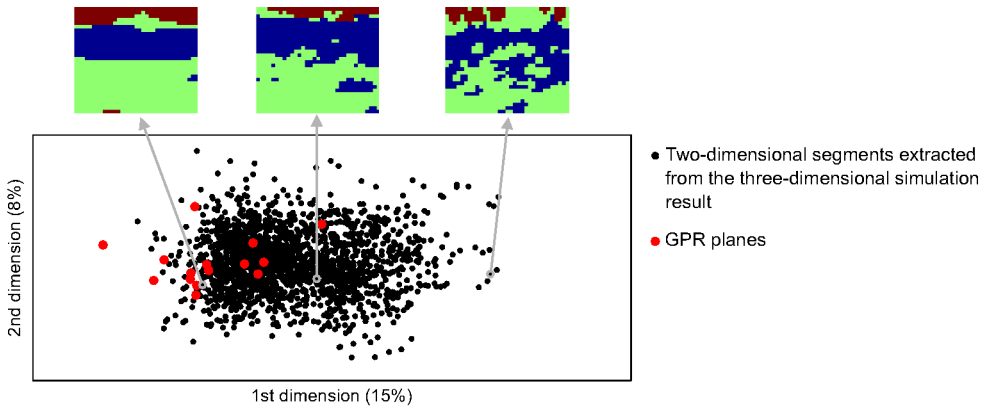

Fig. 8 MDS plot for the sequential two-dimensional approach. The distance is the JS-divergence between the MPHs of simulated slices / GPR planes

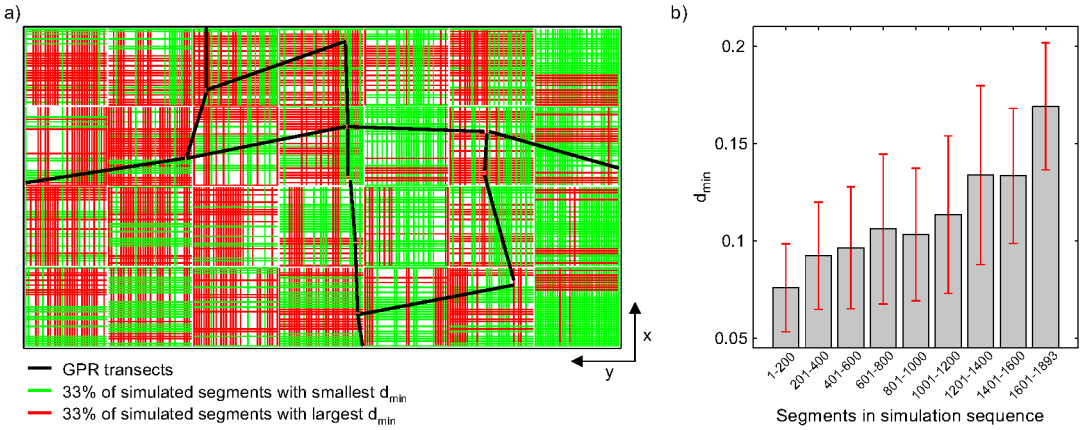

Fig. 9 (**a**) Spatial distribution of two-dimensional segments with "high quality" (green) and "low quality" (red) for the sequential two-dimensional approach. (**b**) Quality of the segments as function of the simulation sequence. Red lines represent two times the standard deviation

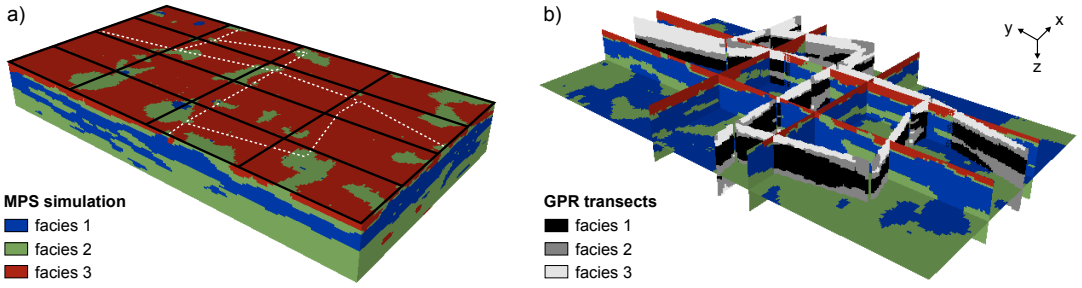

Fig. 10 Result of the combined approach. (**a**) Block view of the simulated three-dimensional volume. White dashed lines indicate the GPR transects. Black lines indicate the slices simulated with the sequential two-dimensional approach before switching to the three-dimensional reconstruction approach. (**b**) Cross-sectional slices and comparison with the GPR planes. For clarity, a different color-scale is prescribed to simulated nodes and to GPR informed nodes

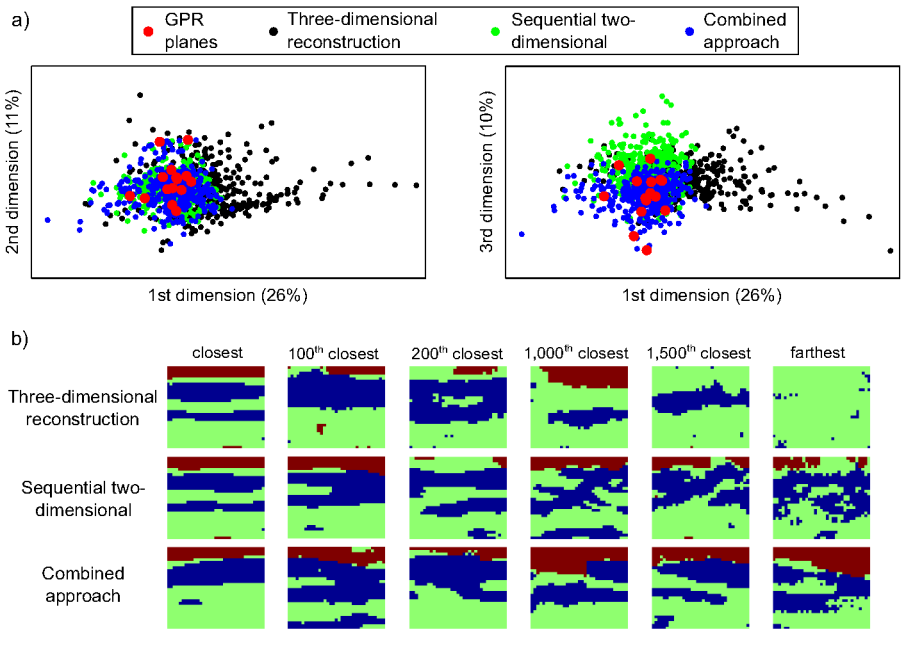

Fig. 11 (**a**) MDS map projections along the first three dimensions. For clarity, only every fifth segment extracted from each model is plotted. (**b**) Comparison of individual segments (with different $d_{\min}$) from the different models

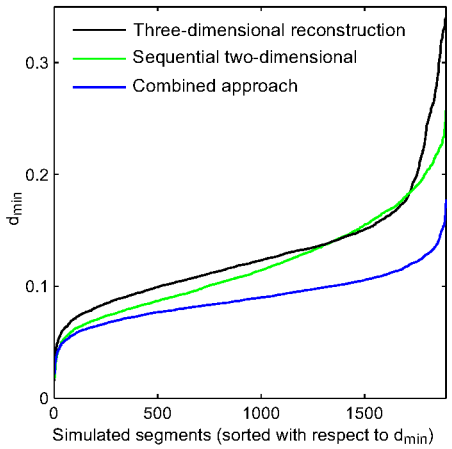

Fig. 12 Characteristic distribution curves for $d_{\min}$ obtained from the different approaches

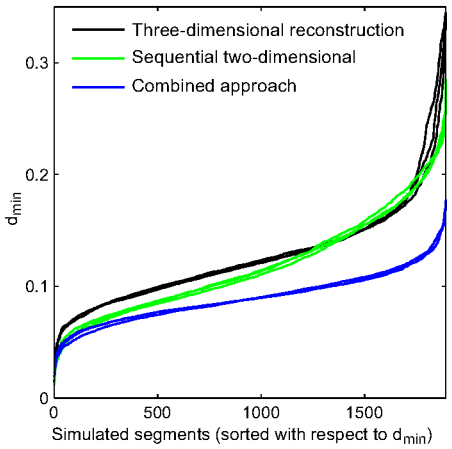

Fig. 13 Distribution of $d_{\min}$ for three different realizations obtained from each approach

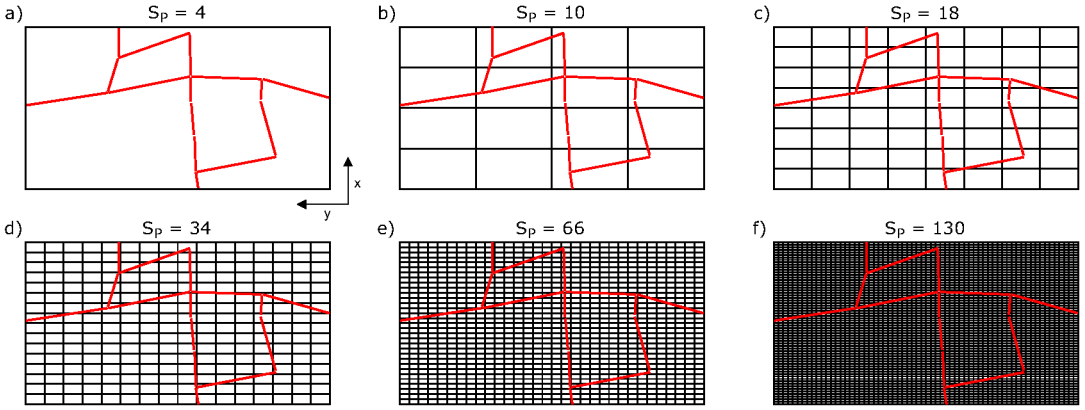

Fig. 14 (**a-f**) Different switching points. S_P refers to the number of slices (black lines) simulated with the sequential two-dimensional approach. The conditioning GPR planes are represented by red lines

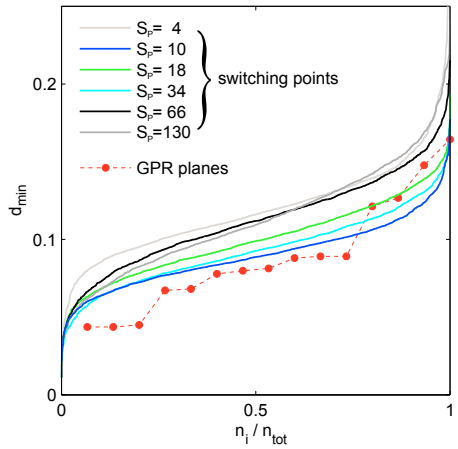

Fig. 15 Effect of different switching points on simulation results and comparison with the (reference) distribution of $d_{\min}$ derived from the GPR images