3D multiple-point statistics simulation using 2D training images

A. Comunian *, P. Renard, J. Straubhaar

Centre of Hydrogeology and Geothermics (CHYN), University of Neuchâtel, Rue Emile-Argand 11, CP 158, CH - 2009 Neuchâtel, Switzerland

ARTICLE INFO

ABSTRACT

One of the main issues in the application of multiple-point statistics (MPS) to the simulation of three-dimensional (3D) blocks is the lack of a suitable 3D training image. In this work, we compare three methods of overcoming this issue using information coming from bidimensional (2D) training images. One approach is based on the aggregation of probabilities. The other approaches are novel. One relies on merging the lists obtained using the *impala* algorithm from diverse 2D training images, creating a list of compatible data events that is then used for the MPS simulation. The other (s2Dcd) is based on sequential simulations of 2D slices constrained by the conditioning data computed at the previous simulation steps. These three methods are tested on the reproduction of two 3D images that are used as references, and on a real case study where two training images of sedimentary structures are considered. The tests show that it is possible to obtain 3D MPS simulations with at least two 2D training images. The simulations obtained, in particular those obtained with the s2Dcd method, are close to the references, according to a number of comparison criteria. The CPU time required to simulate with the method s2Dcd is from two to four orders of magnitude smaller than the one required by a MPS simulation performed using a 3D training image, while the results obtained are comparable. This computational efficiency and the possibility of using MPS for 3D simulation without the need for a 3D training image facilitates the inclusion of MPS in Monte Carlo, uncertainty evaluation, and stochastic inverse problems frameworks.

1. Introduction

Multiple-point statistics (MPS, Guardiano and Srivastava, 1993; Strebelle, 2002) is a geostatistical simulation technique that makes it possible to reproduce complex geological patterns and to provide more realistic results than standard geostatistical techniques. Recent implementations of the MPS technique make it possible to include auxiliary information efficiently (Chugunova and Hu, 2008; Straubhaar et al., 2011) and to solve the computational issues related to the simulation of large grids in terms of RAM and CPU time (Straubhaar et al., 2011). Still, an important barrier to applying the MPS simulation technique remains, the lack of training images, especially for three-dimensional (3D) situations. A suitable training image (TI), i.e., a conceptual statistical model of the geology that has to be simulated, is the fundamental requirement for MPS simulation. Finding a 3D TI is often challenging. This is claimed, for example, in the work of Huysmans and Dassargues (2010), where two 2D training images are deduced from two quarry walls in the Brussels Sands

formation (Belgium), or in the work of Le Coz et al. (in press), where only one horizontal training image is available from detailed aerial/satellite images of the Komadougou Yobé alluvium (Niger).

The lack of a 3D TI can be faced using different approaches. One possibility is to build the 3D TI using object-based or process-based techniques (for a review of the methods see, for example, Koltermann and Gorelick, 1996; de Marsily et al., 2005; the description of some object-based and process-based methods can be found in the works of Maharaja, 2008; Pyrcz and Strebelle, 2008; Abrahamsen et al., 2007; Pyrcz et al., 2005). Often, the best solution is to rely on these explicit spatial models (Journel and Zhang, 2006). However, in many cases it can be difficult to apply these techniques. For example, with an object-based technique it is challenging to reproduce all the kind of geological shapes that can be represented by an outcrop mapped at high resolution, or to account for nonstationarity. To overcome this difficulty, one can, for example, mix different techniques and use a hierarchical approach (Comunian et al., 2011), but the simulation framework can become complex. The complexity of the algorithms arises in many cases mainly from conditioning the simulations to the observation and taking into account the nonstationarity rather than from the technique itself (see, e.g., Michael et al., 2010).

---

* Corresponding author.

E-mail addresses: alessandro.comunian@unine.ch (A. Comunian),
philippe.renard@unine.ch (P. Renard), julien.straubhaar@unine.ch (J. Straubhaar).

Another possibility, which is explored in this work, is to use the statistical information obtained from several 2D training images. Along this line of research, Okabe and Blunt (2004, 2007) aggregated with a linear pooling formula the information coming from a 2D thin section of a carbonate rock in order to reproduce the micropores at the submicrometer scale. In this way, they improved the estimation of the porosity of a sample of carbonate rock where the 3D porosity at the macroscale was investigated using micro computed tomography imaging.

The statistical information coming from variograms and 2D training images was used by Caers (2006) to simulate geological structures at different scales. Caers (2006) aggregates the statistical information using the tau model (Journel, 2002) and obtains encouraging results.

This paper is an attempt to investigate in more detail the problems studied by Okabe and Blunt (2004, 2007) and Caers (2006) and to propose new approaches to the *interesting research topic* (Hu and Chugunova, 2008) of 3D multiple-point simulation using 2D training images. In Section 2, the terminology used to describe the multiple-point statistics is introduced; Section 3 contains a description of two novel methods and the description of the method based on the aggregation of probabilities. The three methods are then tested and compared using different benchmarks; when a 3D training image is accessible, a direct comparison of the results with a reference using the criteria depicted in Section 4 is made. The simulation methods are tested and compared to the simulation of a micro computed tomography image (Section 5), a sedimentary environment observed in the Brussels sands (Section 6), and a fluvioglacial aquifer analog (Section 7). The last two sections contain results and discussions.

The approaches presented here are illustrated using a 3D simulation as target and 2D training images as sources of information. However, they can be extended to a more generic framework where the target simulation has dimension m, and the information is collected from sources of dimension $n < m$. For example, these techniques could be applied in the framework of spatiotemporal models, where the target simulation can have $m = 4$.

2. Background and terminology

This section reports the definitions and the notation needed to describe the methodologies illustrated in the following. The MPS method and its implementation using the *impala* algorithm (Straubhaar et al., 2011) are explained using the same notation as the paper of Straubhaar et al. (2011).

The MPS method is a sequential simulation technique. All the nodes are simulated in a sequential order, and the same algorithm is applied repeatedly while accounting from the previously simulated nodes (Strebelle, 2000). At a given iteration, one has to consider the current node to be simulated. Its location is denoted u (a vector of coordinates). A *search template* is then defined. It consists of N lag vectors $h_1, \ldots, h_N$ in a 2D or 3D space. The search template τ of size N referred to a node location u is defined as a set of locations centered on u:

$$\tau(u) = \{u + h_1, \ldots, u + h_N\}. \tag{1}$$

A *data template* is then associated to a search template. If s is a function that associates the value of a facies code to a spatial location, then a data template d of size N at the location u is the vector of the facies codes at all the neighboring nodes of u:

$$d(u) = \{s(u + h_1), \ldots, s(u + h_N)\}. \tag{2}$$

Once a search template is defined, and before the sequential simulation algorithm is started, the first step of many multiple-point simulation algorithms is the construction of a catalog made of all the data templates contained in the training image. The catalog contains information about the frequency of occurrence of a facies code at the central node of a data template.

Two main storage paradigms exist for the catalog. The catalog can be organized using a *three* structure, as in the *snesim* algorithm (Strebelle, 2002), or using a *list* structure, as in the *impala* algorithm (Straubhaar et al., 2011). The methodology depicted in this paper could be adapted to both storage paradigms, but it can be implemented easily with the lists paradigm, which is described here.

A list $\mathcal{L}$ is a collection of list elements L, where each element is composed of two vectors (c, d). The vector $d = (s_1, \ldots, s_N)$ contains one of the facies configurations that are observed in the training image scanned with the search template τ. Given a training image composed of M different facies codes, the vector $c = (c_1, \ldots, c_M)$ contains the counters for the occurrences of each facies at the center of the search template (see Straubhaar et al., 2011, for more details).

3. Methodology

In order to face the lack of a complete 3D training image when a 3D simulation is required, we propose and compare three methods. The first is based on probability aggregation methods; the second is based on 2D simulation constrained by the conditioning data computed at a previous 2D simulation step; the third is based on the mixing of compatible data events of the lists extracted from different 2D training images.

Note that in the following we assume that we have access to at least two different training images oriented along perpendicular directions. When only one training image is accessible, we assume that this same image can be used to depict the required features along another direction too.
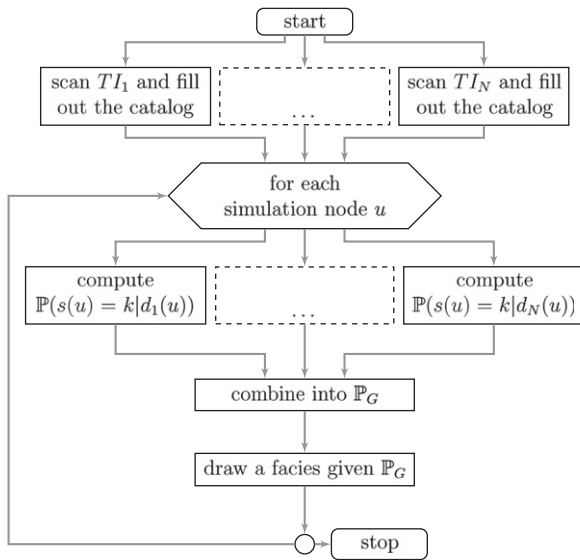
3.1. Probability aggregation

The first method is based on the aggregation of probability, which was previously explored by Okabe and Blunt (2004, 2007) and Caers (2006). In this paper, we apply this concept using two aggregation methods, we test it more extensively than the above authors using synthetic and real case studies, and we check how the results obtained depend on the weighting factors.

The principle of the method, illustrated in Fig. 1, is the following. Since a complete 3D multiple-point conditional probability distribution function (cpdf) is not accessible, an approximation is obtained by aggregating the probabilities that are computed from several sources of information having a lower dimensionality (i.e., 2D training images, or variograms).

Consider, for example, a case where N bidimensional training images $TI_1, \ldots, TI_N$ are accessible. These images can be, for example, three training images having their normal vectors oriented along the coordinates of a 3D Cartesian coordinate system (Fig. 2b, d and e). The first steps of a MPS simulation performed with this approach proceed separately for each training image like a standard simulation. The training images $TI_1, \ldots, TI_N$ are scanned separately using corresponding search templates $\tau_1, \ldots, \tau_N$. For example, the search templates can have the shapes illustrated in Fig. 3a–c. The node locations h_i that compose each data template are selected to lie in the plane defined by the corresponding training image. During the scan, one catalog (a *list* in case of the *impala* algorithm) is created for each training image. Then, at a point u of the simulation domain, the conditional probabilities densities $\mathbb{P}(s(u) = k \mid d_1(u)), \ldots, \mathbb{P}(s(u) = k \mid d_N(u))$ of finding the facies k are computed for each training image and for each facies.

Here the terms $d_1,\ldots,d_N$ are the data events found in the simulation grid using the corresponding search templates $\tau_1,\ldots,\tau_N$. In the following, the cpdf's obtained from the different training images will be denoted for the sake of brevity $\mathbb{P}_1,\ldots,\mathbb{P}_N$.

Fig. 1. The procedure for performing 3D MPS simulations aggregating the probabilities from 2D training images. The dashed blocks represent operations that are performed for each considered training image.

At this step, the different cpdf's are aggregated into a global probability term $\mathbb{P}_G$, which is an approximation of the conditional probability

$$\mathbb{P}_G(u) \approx \mathbb{P}(s(u)=k \mid d_1(u),\ldots,d_N(u)). \tag{3}$$

$\mathbb{P}_G$ is used to draw a given facies at the point u, then the simulation proceeds sequentially using the same approach.

How to aggregate the different cpdf's coming from the different training images? Comunian (2011) and Allard et al. (in preparation) propose a critical overview of the methods that can be used to aggregate probability terms. Here two methods that represents two different categories of aggregation methods are tested: the *linear pooling formula* and Bordley's, 1982 *formula* (or the *tau model*, Journel, 2002). For the sake of brevity, in the following we will often refer to the two methods using the acronyms paLin (probability aggregation using a Linear pooling formula) and paBor (probability aggregation using Bordley's pooling formula).

3.1.1. Linear pooling formula

The first aggregation method considered is the linear pooling formula; it aggregates the probability terms using a weighted sum. The method is convex and it does not have the external Bayesianity property, a property that is suitable for the case of sequential simulations (see Comunian, 2011; Allard et al., in preparation, for details). However, the method is appealing because of its flexibility and its simplicity. It is widely used (for example, it was used by Okabe and Blunt, 2004, 2007). With this
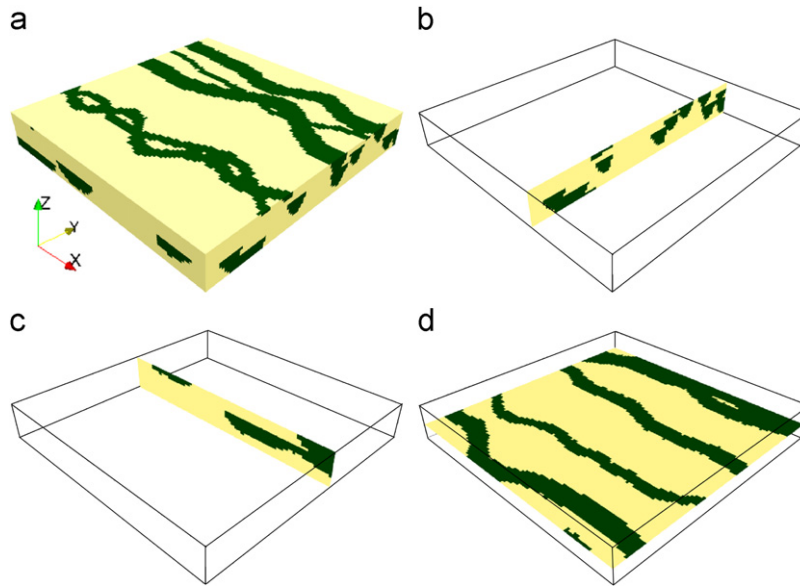
Fig. 2. (a) A complete 3D training image containing channel structures and (b), (c), (d) its slices along different orientation planes. These slices can be used as 2D training images in the proposed approaches.
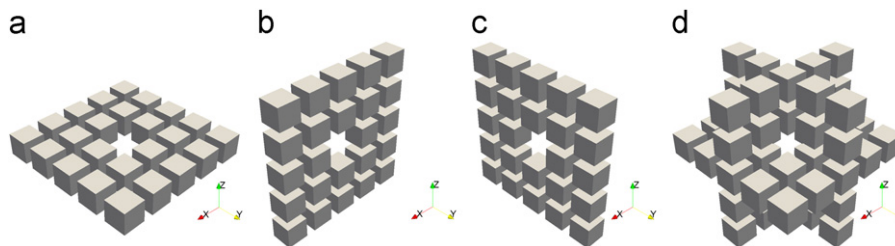
Fig. 3. Examples of bidimensional search templates used to scan training images oriented along different directions ((a), (b), and (c)) and a search template obtained by merging the first three search templates ((d)). (a) $5 \times 5 \times 1$, (b) $5 \times 1 \times 5$, (c) $1 \times 5 \times 5$ and (d) merged.

method, the aggregated probability is given by

$$\mathbb{P}_G(u) = \sum_{i=1}^{N} w_i \mathbb{P}_i \quad \text{with } w_1, \ldots, w_N \in \mathbf{R}^+. \tag{4}$$

To obtain a probability term $\mathbb{P}_G \in [0,1]$, the weights w_i are selected so that their sum S_w equals one.

3.1.2. Bordley's formula

The other aggregation method considered is Bordley's formula, which is expressed in terms of odds:

$$\mathbb{O}_G(u) \propto \mathbb{O}_0^{1-S_w} \prod_{i=1}^{N} \mathbb{O}_i^{w_i} \quad \text{and} \quad S_w = \sum_{i=1}^{N} w_i. \tag{5}$$

The odds $\mathbb{O}_i$ are related to the corresponding probability terms $\mathbb{P}_i$ by

$$\mathbb{O}_i = \frac{\mathbb{P}_i}{1 - \mathbb{P}_i} \quad \text{with } i \in \{G, 0, 1, \ldots, N\}. \tag{6}$$

In Eq. (5), the term $\mathbb{O}_0$ is the odds of the prior probability term $\mathbb{P}_0$. This last term represents prior information that Bordley's formula makes it possible to integrate in the aggregation process; it can, for example, be used to incorporate the target facies proportions derived from a previously performed survey. The weights $w_1, \ldots, w_N \in \mathbf{R}^+$ can be selected so that their sum S_w is greater than, less than, or equal to one. In the last case the contribution of $\mathbb{P}_0$ is ignored. More details about the other cases can be found in Bordley (1982), Comunian (2011) and Allard et al. (in preparation).

Bordley's formula has properties that are desirable in a geostatistical simulation framework. One of these properties is the 0/1 *forcing property* (Allard et al., 2011): if one probability term provides a null value (impossible event), then the aggregated result will be null; the same is valid for the value one (certain event). Other properties of Bordley's formula are external Bayesianity and nonconvexity (for more details see Comunian, 2011; Allard et al., in preparation).

3.1.3. Weights selection

In order to apply the aggregation methods presented, the weighting factors must be defined. Comunian (2011) and Allard et al. (in preparation) review the methods that can be used to determine these weights and apply some methods to the solution of synthetic case studies. Here, three weighting schemes are proposed and tested.

The first weighting scheme considers the weights equal and fixed during the entire simulation process. If there are no elements that makes it possible to prefer one training image rather than another, this is a reasonable choice, which can easily be applied to the linear pooling formula: the aggregation then reduces to the arithmetic mean of the cpdf extracted from each training image. However, this can be more complicated for Bordley's formula, where there is not the restriction $S_w=1$. For this formula, one of the most common approaches is to assume unitary weights, which correspond to the assumption of conditional independence among sources of information (Journel, 2002; Krishnan, 2008). Different weight choices are explored in the following test cases. In the case of Bordley's formula, important differences are expected when $S_w > 1$ or when $S_w < 1$; these two situations correspond to negative or positive values of the exponent of the term $\mathbb{O}_0$. In the case of binary variables, the aggregated probability follows the tendency suggested by the prior term P_0 when $S_w < 1$, and has an opposite tendency when $S_w > 1$. The works of Bordley (1982), Comunian (2011), and Allard et al. (in preparation) provide a detailed description of this topic.

The second weighting scheme assumes that the weights vary during the simulation process. It is known that the weights should not be constant (Krishnan, 2008), because they may depend on the data configuration. However, the complete cpdf (Eq. (3)) not being accessible, one is left with having to provide a reasonable approximation. A possible scheme, when multigrids are employed, is to assign a different weight for each multigrid level, giving greater confidence (with larger weights) to the information that comes from the training images for the simulation of the first multigrid levels; then, when the main structures are formed and for the simulation of the final multigrid levels, more weight is given to the prior terms in order to try to get closer to the target facies proportions. With this weighting scheme, we performed some tests that are not reported here because they did not shown significant improvements in the results obtained.

The last weighting scheme proposed here is novel: weights vary at each step of the simulation process. They are computed at each point u of the simulation path, and are proportional to the size of the data events that are used to compute the cpdf from the different training images. The weights are selected to be proportional to the size of the data event, but of course their sum must respect the condition imposed in the case of the linear pool ($S_w=1$) or for Bordley's formula (S_w should be equal to the sum of the default values of the weights).
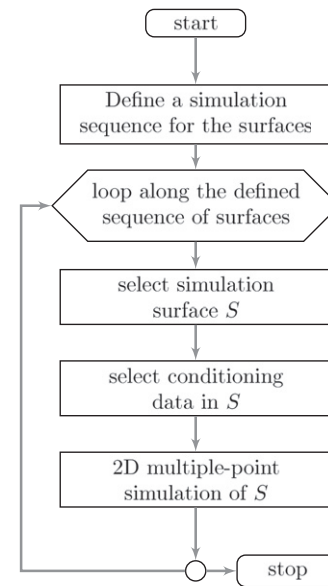
Fig. 4. The procedure for simulating a 3D volume using sequential bidimensional MPS simulations and conditioning data (approach `s2Dcd`).
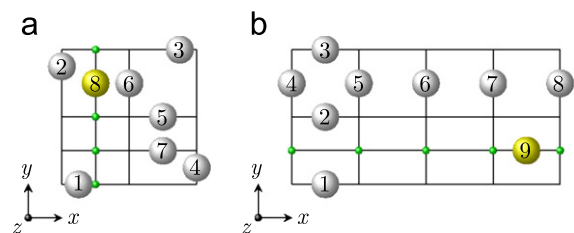
Fig. 5. The top view in the plane xy of two possible sequences of bidimensional surfaces S_i for the `s2Dcd` method when (a) the dimensions x and y of the domain are comparable or (b) $x \gg y$. Only the first eight (a)/nine (b) surfaces are shown. The green dots represent the vertical columns of conditioning data considered for the simulation of the surfaces S_8 (a) and S_9 (b). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.2. Sequential 2D simulations with conditioning data

The second approach proposed here is based on sequential 2D multiple-point simulations and on the use of conditioning data (in brief, approach s2Dcd, sequential 2D conditioning data). The procedure, depicted in Fig. 4, is the following. At the first step, a sequence for the sequential simulation of the 2D surfaces is defined. Then, for each simulation step i, a 2D MPS simulation is performed along a surface S_i using the training image that describes the heterogeneity along the direction of S_i; the conditioning data used are the points contained in $S_i \cap (\bigcup_{j < i} S_j)$ and the points contained in $S_i \cap D$, where D is the set of hard conditioning data provided

Fig. 6. The procedure for performing 3D multiple-point simulation aggregating the probabilities from 2D training images. The dashed blocks represent operations that are performed for each training image considered.

(if available) at the beginning of the simulation. These steps are repeated until all the domain is simulated.

The definition of the sequence of simulation surfaces is crucial for this approach, which relies on the information provided by the conditioning data, and must therefore try to include, at each simulation step, as many conditioning data as possible. At the same time, the geometrical relations among the dimensions of the simulation domain must be considered in the choice of the simulation sequence. If the dimensions of the simulation domain are comparable, then a possible simulation order for the surfaces S_i is illustrated in Fig. 5a. It refers to a square domain where there are no conditioning data at the first simulation step and where only two training images (one perpendicular to the axis x and the other perpendicular to the axis y) are available. When conditioning data are available or when the dimensions of the simulation domain differ significantly, the order of simulation of the 2D surfaces should be customized according to the location of the conditioning data and to the shape of the simulation domain. As a general rule, the simulation order of the 2D surfaces should be selected in order to gradually fill up the simulated domain and to include as many conditioning data as possible. However, different choices can be made according to the correlation length of the structures contained in the training images and/or to the location of the conditioning data.

An example of simulation sequence for a simulation domain with $x \gg y$ is illustrated in Fig. 5b.

3.3. Lists merging

Another method of simulating in 3D using 2D training images, illustrated in Fig. 6, relies on the use of lists. Nevertheless, the method described here for the case of the *impala* algorithm can be adapted to other MPS implementations.

Let us illustrate this method with an example (Fig. 7). Consider two 2D binary training images T_{xz} and T_{xy} oriented along the
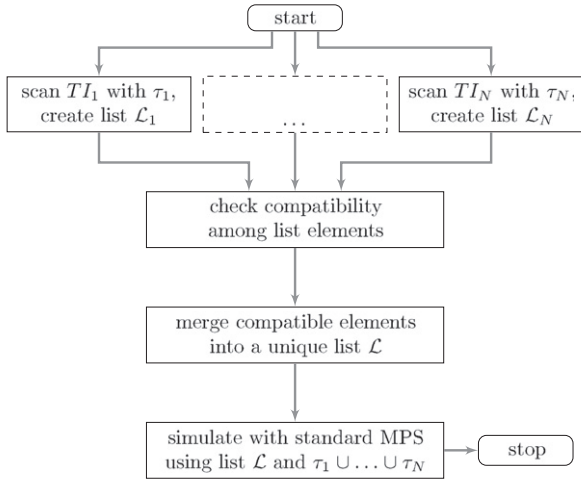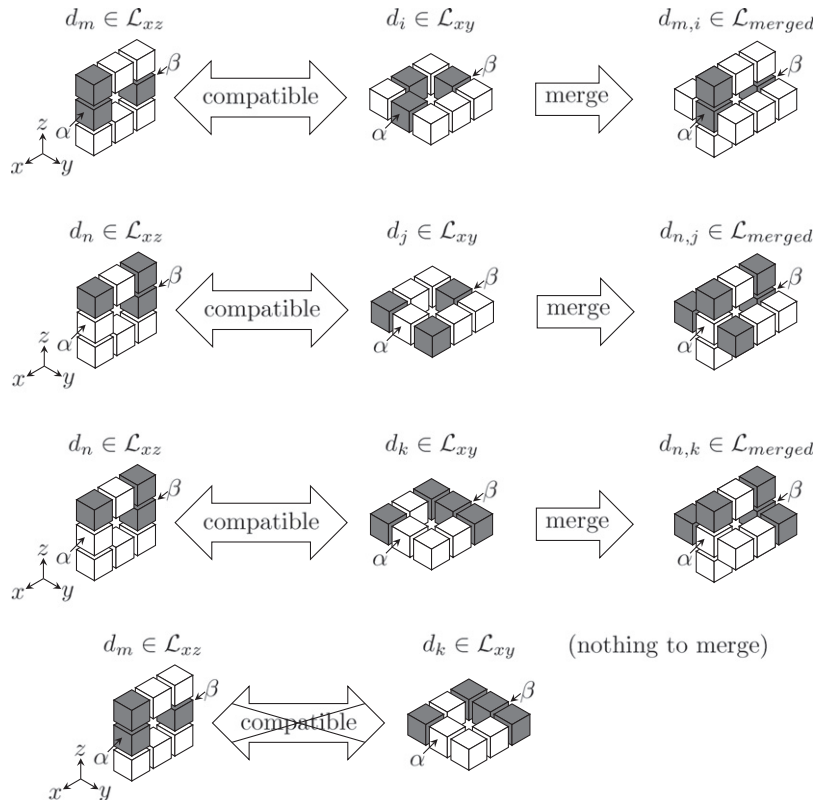
Fig. 7. The data events of the couples d_m and d_i, d_n and d_j, d_n and d_k belonging to the lists $\mathcal{L}_{xz}$ and $\mathcal{L}_{xy}$ are compatible. Therefore, they can be merged into the data events $d_{m,i}$, $d_{n,j}$, and $d_{n,k}$. The data events d_m and d_k are not compatible (different facies in the position α), and they cannot be merged.

planes xz and xy and two corresponding search templates τ_{xy} and τ_{xz} of size 3×3 pixels. For each training image, the corresponding lists $\mathcal{L}_{xz}$ and $\mathcal{L}_{xy}$ are computed using the *impala* algorithm. Each element of the lists, for example, $L_m \in \mathcal{L}_{xz}$, is composed of a vector that contains the configuration of a data event d_m and a vector c_m that contains counters for the occurrences of each facies for the given d_m. Fig. 7 illustrates some situations where the data events d_m and d_n belonging to the list elements L_m, $L_n \in \mathcal{L}_{xz}$ can be combined with data events d_i, d_j, and d_k, belonging to the elements of the list $\mathcal{L}_{xy}$. For example, the data events d_m and d_i are compatible, and a list element of the merged list $\mathcal{L}_{\mathrm{merged}}$ can be created. This last list element contains the data event $d_{m,i}$ obtained by merging d_m and d_i. The compatibility between data events of the lists $\mathcal{L}_{xz}$ and $\mathcal{L}_{xy}$ is checked by looking at the facies at the positions α and β. With the same procedure, the compatibility between all the data events of the two lists is verified and, if it is the case, an element of $\mathcal{L}_{\mathrm{merged}}$ is created. Of course, it can happen that a data event of one list (for example, $d_n \in \mathcal{L}_{xz}$) is compatible with more than one data event of the other list (for example, $d_j, d_k \in \mathcal{L}_{xy}$). In this case, one element of $\mathcal{L}_{\mathrm{merged}}$ is created for each compatible combination. Once the compatibility among the data events belonging to the elements of the two lists is verified and a merged data event is created, another step is needed to fill up the element of $\mathcal{L}_{\mathrm{merged}}$—the computation of the vector of counters. These counters can be computed as a function of the counters of the two combined list elements. Once again, this problem is ill posed because there is no way to be sure that the combination is correct, unless the 3D information is available. Here four

approaches are proposed and tested. For the first one, the counters are computed using the minimum value of the counters available for the compatible data events in each training image. For the other three approaches the mean value, the maximum value, or the sum of the counters is used.

The merged list computed in this way can be used to simulate the 3D domain. At the same time, the size of the merged list can be used as an indicator of the compatibility of the training images considered. Indeed, in some situations it can be difficult to determine if the models provided by two 2D training images are coherent and can depict different but compatible aspect of the same 3D features which must be reproduced. This method offers a quantitative approach to deal with this issue.

4. Comparison criteria

For the test considered when a 3D reference TI is available, direct comparisons among this 3D training image, the simulation obtained with the methods proposed here, and the standard MPS simulation methods that make use of this 3D training image are performed. When a 3D reference is not available, the features of the results are compared with those of the considered training images. The comparisons between the simulations obtained with the proposed methods and the references are based on visual comparisons, the reproduction of facies proportions, the computing time, and the following criteria and parameters:

Number of geobodies: The number of connected components (clusters) composed of the same facies code. Comparing this value computed for the reference image and the one obtained for the simulations makes it possible to check the continuity of the structures that are simulated. However, as we show in the following, this parameter should be considered as indicative only. Moreover, while the interpretation of this parameter is intuitive for the simulations of binary images, it is less intuitive for complex structures and when the number of simulated facies is greater than 2.

Geobodies area (volume): The area (volume) of the connected components of the same facies is computed by counting the number of pixels (voxels) contained in each component. For each simulation, the statistical distribution of the areas (volumes) is compared with that of the reference image.

Connectivity functions: Following the definition of connectivity proposed by Strauffer and Amnon (1994) and Allard (1992) for a binary medium Ω, two points x and y are connected when there exists at least one path completely contained in the permeable component $\Omega_{|\mathrm{perm.}}$ of the medium that joins the two points; in this case the notation $x \leftrightarrow y$ is used. With this notation, a connectivity function between two points at a distance h can be defined as

$$g(h) = \mathbb{P}(x \leftrightarrow x + h \,|\, x \in \Omega_{|\mathrm{perm.}}). \tag{7}$$
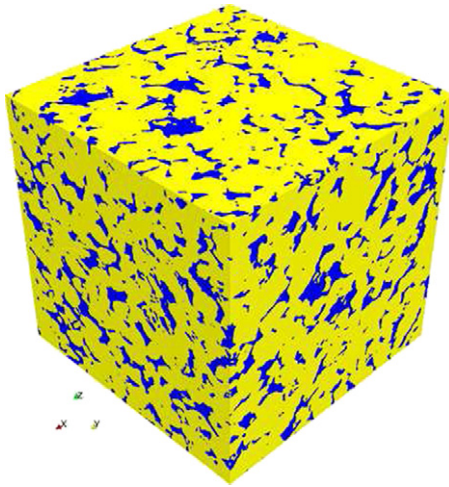
Fig. 8. The micro CT of a sample of Berea sandstone used as a 3D reference. In the image, of size $400 \times 400 \times 400$ voxels, the blue represents the pores and the yellow the matrix. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
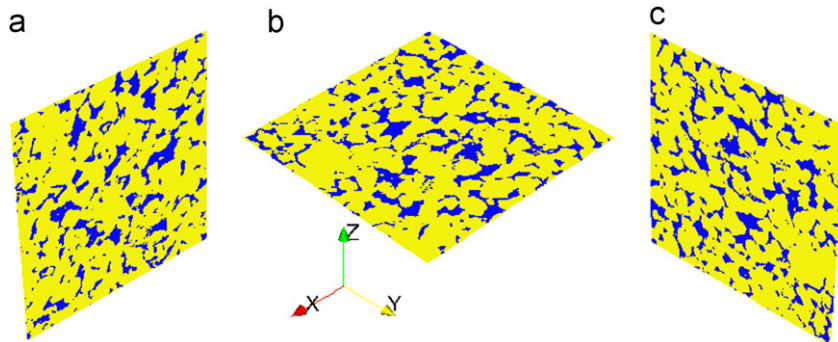
Fig. 9. Training images used along the different directions for the tests on the micro CT image. (a) TI along xz, (b) TI along xy and (c) TI along yz

In this work we compare the connectivity functions computed for the reference image and for the simulations. In the cases considered the images can be considered isotropic and only the connectivity along the direction *x* is computed.

Equivalent hydraulic conductivity tensor: The comparisons among the morphological features of the geobodies can become difficult when the number of facies that are simulated is greater than 2. In this case, we support the comparisons using another criterion, based on the computation of the equivalent hydraulic conductivity tensor **K**. **K** is computed using typical values of the hydraulic conductivity of the geological facies simulated, performing stationary flux simulations with linear boundary conditions along the three different axis directions (Rubin and Gómez-Hernández, 1990; Renard et al., 2001) with the finite elements code Groundwater (Cornaton, 2007). To compare the equivalent permeability of the reference $\mathbf{K}_{ref}$ with that of the simulation $\mathbf{K}_{sim}$, a relative error ε based on the Frobenius norm $\| \ \|$ is computed:

$$\varepsilon = \frac{\|\mathbf{K}_{ref} - \mathbf{K}_{sim}\|}{\|\mathbf{K}_{ref}\|}. \tag{8}$$

When the computational effort allows it, all the above parameters are computed for a minimum of 10 and a maximum of 100 simulations.

In this work, we considered suitable for the comparisons the set of parameters listed above. Note, however, that other parameters could be checked for the comparisons, such as the ones proposed by De Iaco and Maggio (2011), and Boisvert et al. (2010).

Table 1
Comparison of different methods for the simulation of the micro CT image of a Berea sandstone. The cases where the weights are selected at each simulation step proportional to the data event size are indicated with "$\propto d$".

Test	Method	Method details	Time (s)	# TI	# geob.	Porosity (%)
(a)	paLin	$w_{1,2} = 0.5$	1299	2	2916	21.01
(b)	paLin	$w_{1,2} \propto d$	1333	2	2384	20.38
(c)	paBor	$w_{1,2} = 1$	1152	2	592	20.74
(d)	paBor	$w_{1,2} \propto d$	1172	2	519	25.72
(e)	paBor	$w_{1,2,3} = 1$	1689	3	269	23.44
(f)	paBor	$w_{1,2,3} \propto d$	1999	3	227	29.01
(g)	s2Dcd	2 TI	459	2	226	22.10
(h)	s2Dcd	3 TI	434	3	212	19.58
(i)	3D	$\tau_{xy} \cup \tau_{xz}$	72 500	1	775	20.86
(j)	3D	$\tau_{xy} \cup \tau_{xz} \cup \tau_{yz}$	131 234	1	395	21.33
(k)	3D	$7 \times 7 \times 7$	493 414	1	79	22.40
Ref.					44	20.13

5. Micro computed tomography image

The first realistic benchmark used to test the proposed methods is the result of a micro computed tomography (micro CT) on a sample of Berea sandstone (Okabe and Blunt, 2004, 2007). It is a 3D pore-space binary image discretized in $400 \times 400 \times 400$ voxels (Fig. 8). It was selected for a number of reasons. First, the 3D results of the simulation methods proposed can be directly compared with the actual 3D reference. Second, it is sufficiently stationary to be considered as a training image. Finally, it shows that the methods proposed here can be applied to a range of spatial scales, pore-scale included.

The data set extracted from the reference image is composed of a 3D training image of dimension $200 \times 200 \times 200$, smaller than the original one in order to allow faster simulation for the methods that require a full 3D training image; a subset of the original image of dimension $100 \times 100 \times 100$, which is of the same size as the 3D simulations obtained with the methods under examination, which can be used for direct comparisons with the simulations; three bidimensional training images obtained as slices 400×400 of the reference image along the planes *xy*, *xz*, and *yz*.

5.1. Preliminary bidimensional calibration

General guidelines for how to set up the parameters of a multiple-point simulation exist (Liu, 2006). However, these parameters strongly depend on the considered training image and should be adapted to the specific case. To look directly on 3D simulations with a trial and error procedure for a suitable parameter set requires higher computation and interpretation efforts than on 2D simulations. For this reason, the procedure used to select the number of multigrid levels and the size of the search template was performed only on 2D cases. This preliminary step is performed for all the case studies considered in this work.

### 5.2. Choose the weights for Bordley's formula

The application of the two proposed aggregation formulae requires the definition of weighting factors. For the case of the micro CT image, there is no element that makes it possible to prefer one of the three 2D training images extracted as sections along the planes *xy*, *xz*, and *yz* to another. The 3D image can be considered isotropic. As a consequence, the same weight is associated with each considered training image. While in the
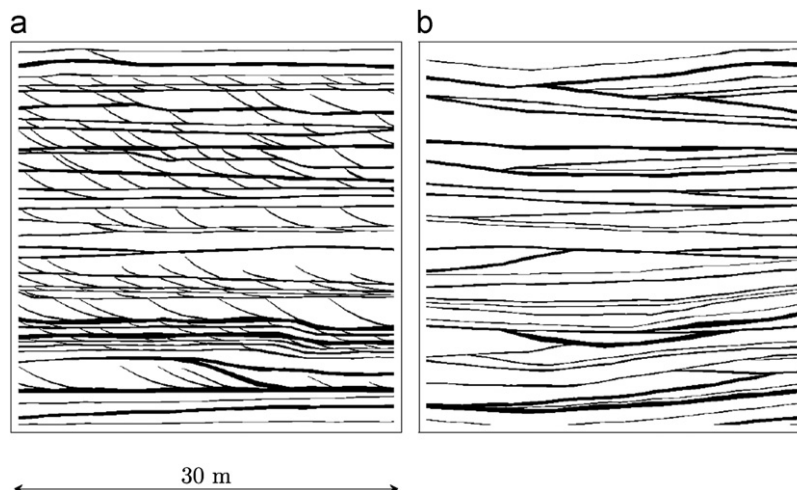
Fig. 10. The training images considered for the Brussels Sands sedimentary environment (a) along the plane *xz* and (b) along the plane *yz*. Clay-rich bottom set are in white, while mud drapes are in black.

case of the linear pooling formula this reduces the aggregation process to a simple arithmetic mean of the probability terms, in the case of Bordley's formula more freedom is left in the choice of
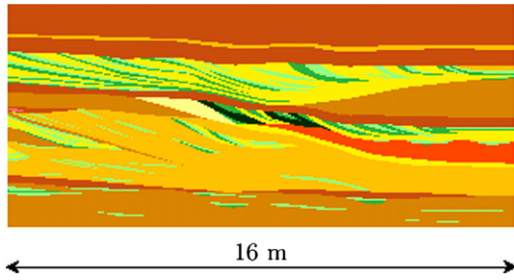

Fig. 11. One of the outcrops of the Herten aquifer analog dataset. It is used as a training image in both the *xz* and *yz* directions for the simulation with the methods `paBor` and `s2Dcd`.

the weights, and of their sum S_w in particular (for more details see Bordley, 1982; Comunian, 2011; Allard et al., in preparation). In our case, having access to the properties of the target 3D results, we performed a trial-and-error procedure in order to determine which set of equal weights w_i provides the result closer to the reference image.

The `paBor` method is applied to 11 different values of the weights $w_{1,2}$ ranging from 0.1 to 2.0; two training images parallel to the *xz* and to the *xy* are used. For each parameter set, 100 realizations are obtained.

5.3. Methods comparison

The number of multigrid levels and the size of the data template used for the following tests are defined by the trial-and-error procedure discussed in Section 5.1. The training images used in the following are illustrated in Fig. 9.
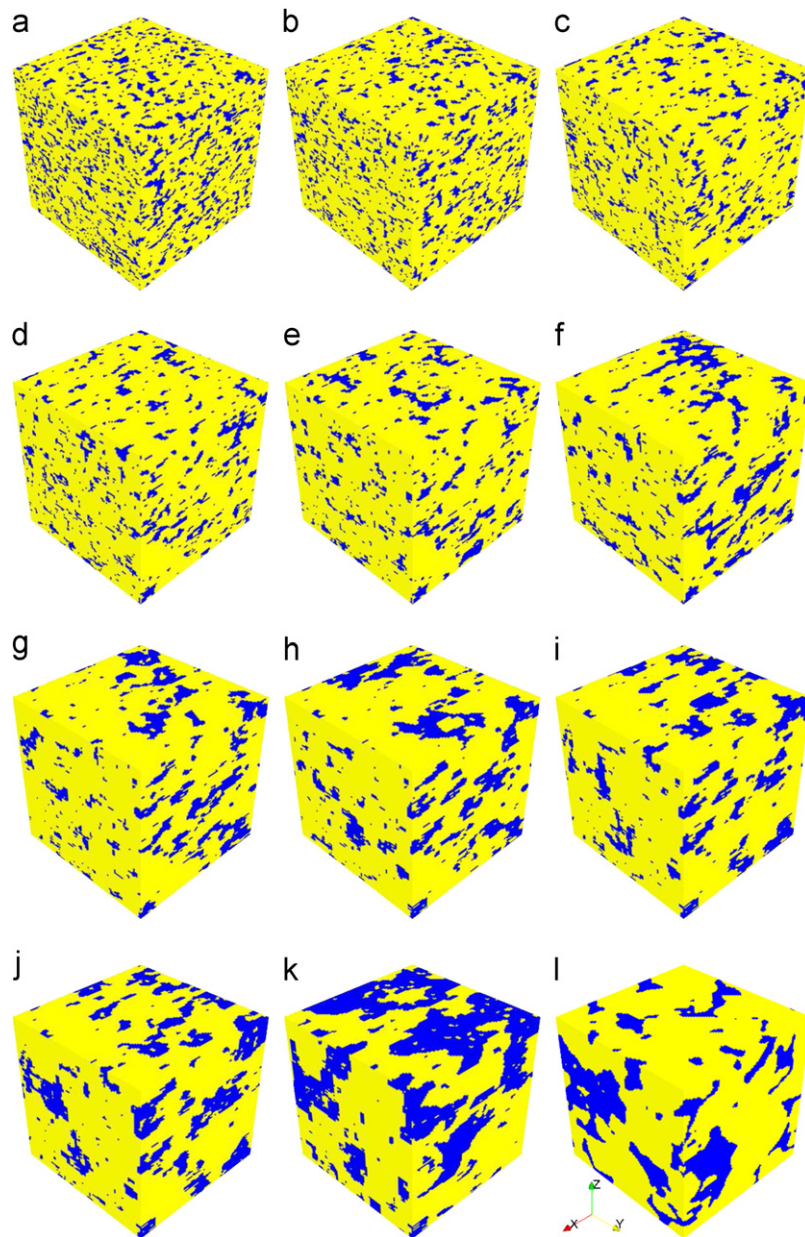
Fig. 12. (a)–(k) Simulations of the micro CT image obtained by the `paBor` method with different values of $w_{1,2}$ and (l) a subset of the original image of the same size of the simulations, for direct comparisons. The two training images used for the simulations are parallel to the planes *xy* and *xz*. (a) $w_{1,2} = 0.1$, (b) $w_{1,2} = 0.2$, (c) $w_{1,2} = 0.3$, (d) $w_{1,2} = 0.4$, (e) $w_{1,2} = 0.5$, (f) $w_{1,2} = 0.6$, (g) $w_{1,2} = 0.7$, (h) $w_{1,2} = 0.8$, (i) $w_{1,2} = 0.9$, (j) $w_{1,2} = 1.0$, (k) $w_{1,2} = 2.0$ and (l) reference.

For the method based on probability aggregation, both the linear formula and Bordley's formula are tested.

The first is tested with two training images parallel to the planes *xy* and *xz* and with weights $w_{1,2} = 0.5$ (test (a), Table 1) or proportional to the size of the data event and such that $S_w = 1$ (test (b), Table 1).

Bordley's aggregation formula is tested using two training images parallel to the planes *xy* and *xz* (tests (c) and (d), Table 1) and three training images parallel to the planes *xy*, *xz*, and *yz* (tests (e) and (f) Table 1). The weights used are $w_{1,2} = 1.0$, $w_{1,2,3} = 1.0$ and weights proportional to the size of the data event with $S_w = \sum_i w_i$ and $w_i = 1$ for each i; the considered prior probability terms are the porosity and $1-$ porosity.

The approach `s2Dcd` is tested using two training images parallel to the planes *xy* and *xz* (test (g), Table 1) or three training images parallel to the planes *xy*, *xz*, and *yz* (test (h), Table 1).

The results obtained with the method based on the merging of lists are not shown here because their quality is worse than that
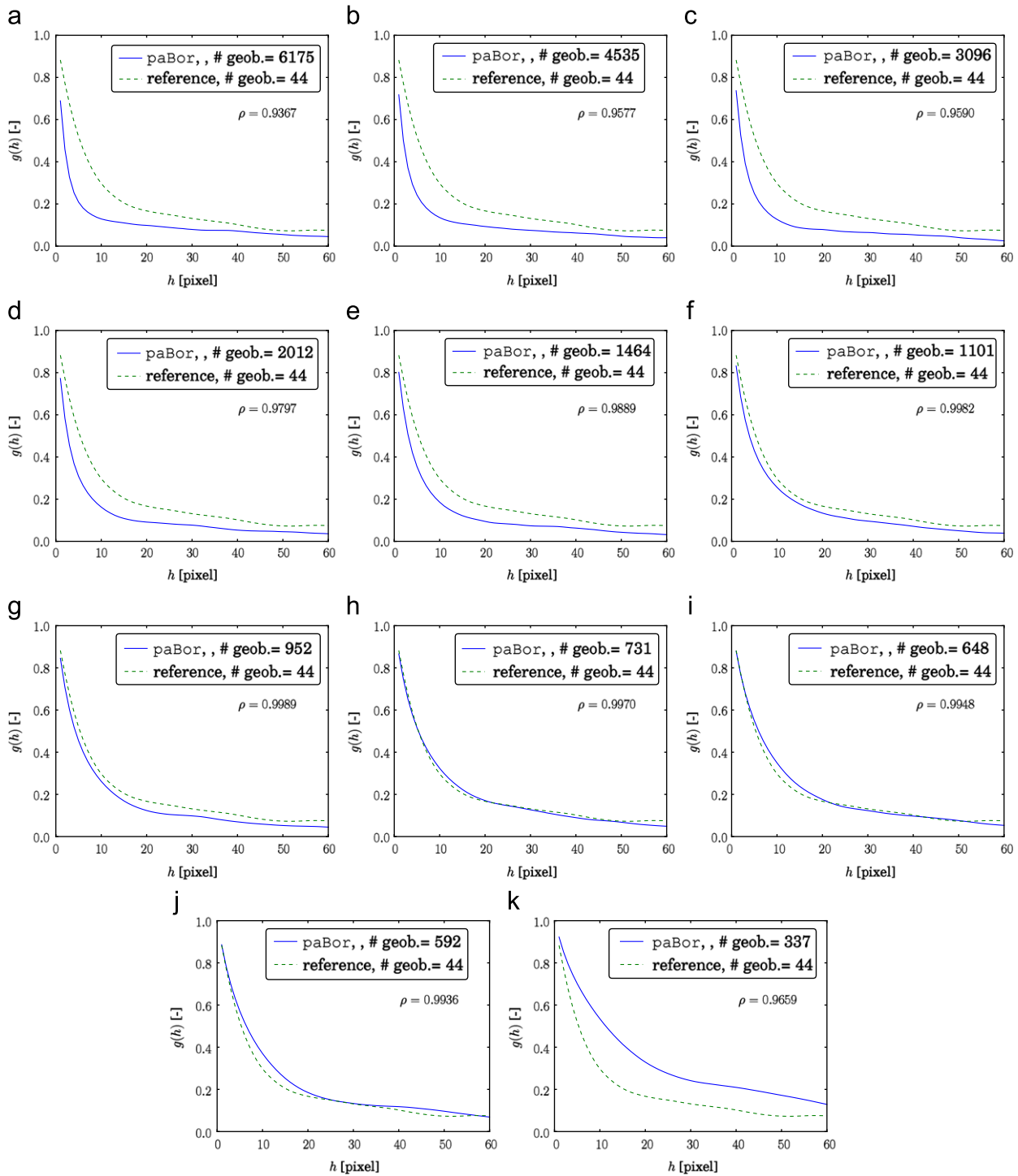
Fig. 13. Connectivity functions computed along the coordinate *x* for the 3D simulations obtained with the `paBor` method and different values of $w_{1,2}$, which correspond to the simulations shown in Fig. 12a– 12k. The dashed line represents the reference; the correlation coefficient ρ between the reference and the computed curve is reported. (a) $w_{1,2} = 0.1$, (b) $w_{1,2} = 0.2$, (c) $w_{1,2} = 0.3$, (d) $w_{1,2} = 0.4$, (e) $w_{1,2} = 0.5$, (f) $w_{1,2} = 0.6$, (g) $w_{1,2} = 0.7$, (h) $w_{1,2} = 0.8$, (i) $w_{1,2} = 0.9$, (j) $w_{1,2} = 1.0$, (k) $w_{1,2} = 2.0$.

of the results obtained with the other methods. However, an application to check the compatibility among training images based on this method is illustrated in Section 9.

The availability of a 3D training image makes it possible to perform MPS simulations that use it to compare the results with the ones obtained with the methods that use 2D training images only. Here the simulations with a 3D training image (tests (i), (j), and (k), Table 1) are performed using diverse data templates: (k) a complete $7 \times 7 \times 7$ template; (i) a template obtained by merging two bidimensional templates of dimensions $7 \times 7 \times 1$ and $7 \times 1 \times 7$; (j) a template obtained by merging three data templates $7 \times 7 \times 1$, $7 \times 1 \times 7$, and $1 \times 7 \times 7$ (see, for example, Fig. 3). Incomplete 3D search templates are considered in tests (i) and (j) in order to compare the results of the standard MPS technique and the ones tested in this work using search templates of comparable size.

The results obtained with the setup described in this section are discussed in Section 8.2.

6. Reproduction of sedimentary structures

The methods `paBor` and `s2Dcd` are tested using two training images that represent some typical sedimentary structures mapped in the Brussels Sands by Huysmans and Dassargues (2010). In the training images, one can distinguish a matrix of clay-rich bottom sets and distinct mud drapes (Fig. 10).

For this case study, the results are compared visually and from the point of view of equivalent conductivity. An equivalent conductivity tensor is computed in 2D for one of the training images (the one oriented along the plane xz, Fig. 10a). Another equivalent conductivity tensor is computed in 3D for each of the 10 simulations performed for the method `paBor` and for the 10 simulations with the method `s2Dcd`.

To compute the relative error ε on K we use as K_{ref} the 2×2 tensor computed for the training image, and as K_{sim} the four elements of the 3×3 tensors computed for the simulations that correspond to the four elements of K_{ref}. In order to observe a remarkable anisotropy in the equivalent conductivity tensor, for the flow simulations a contrast between the two facies of 100 is used in place of the real values of conductivity measured in the field (Huysmans et al., 2008).

7. The Herten aquifer analog

The `s2Dcd` and `paBor` methods are tested on the reproduction of the hydrofacies distribution of a high-resolution fluvioglacial aquifer analog also. The Herten aquifer analog is described and studied in detail by Bayer et al. (2011). The part of the data set used here consists of six parallel outcrops (oriented along the plane xz in the reference system adopted here) where 10 hydrofacies were recognized and mapped during the excavation of a quarry. A recent study (Comunian et al., 2011) made it possible to obtain a detailed and realistic 3D reconstruction of the entire volume with a hierarchical geostatistical simulation approach; the results of that study are used here as a reference.

The data set is used as conditioning data and training images as well. In our test, we used one of the profiles (Fig. 11) as a 2D training image for both the plane xz and the plane yz. This is a strong hypothesis that can be debatable, but for the application of the proposed methods, we decided to rely only on field data, and no outcrop was available along the perpendicular direction.

For this data set, we test the method based on the aggregation of probabilities using the Bordley's formula and weighting factors

$w_{1,2} = 1$. The method `s2Dcd` is tested with and without the use of the coordinate z as an auxiliary variable.

8. Results and discussion

8.1. Micro CT image and `paBor` weights

The tests performed with the method `paBor` on the micro CT image using different weighting factors suggest that the best set of weights that can be used in this case is $w_{1,2} = 1$. While the 3D representations (Fig. 12) and the connectivity curves (Fig. 13) show that results close to the reference can be obtained with $w_{1,2} = 0.8$, 0.9, or 1.0, the pores' proportions (Fig. 14) show that the best weights set in this case is $w_{1,2} = 1$.

Note that, as expected, the orientations along which the shapes contained in the reference images are reproduced better than the orientations along which a training image is used (Fig. 12). The information about the volume of the geobodies (Fig. 15) did not provide significant insight for the selection of the most appropriate set of weights.
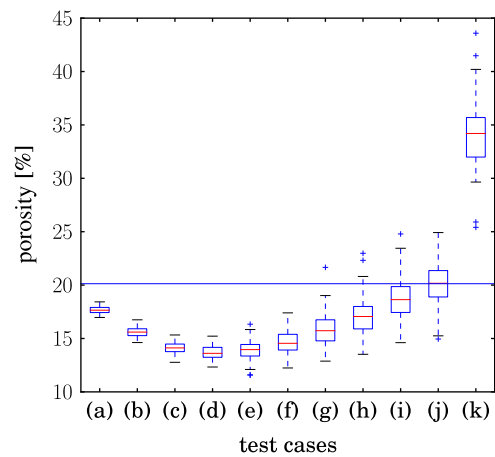
Fig. 14. Proportions of pores for simulations of the micro CT image obtained for different values of the weights $w_{1,2}$ and the method `paBor`. Each test case (boxplots from (a) to (k)) represents a set of 100 simulations performed with a different random seeds and the same weights $w_{1,2}$. The value of the weights grows from 0.1 (a) to 2.0 (k). The horizontal line represents the porosity of the reference image.
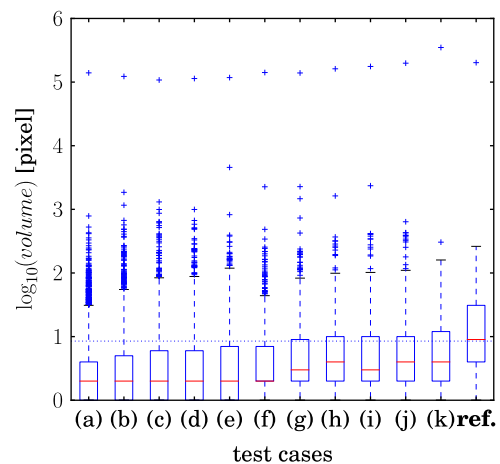
Fig. 15. The $\log_{10}$ of the volume of the pores computed on the simulations of the micro CT image shown in Fig. 12a– 12k. The simulations are obtained with the `paBor` method and with a different value of the weights $w_{1,2}$ for each test case. The weights vary from 0.1 (a) to 2.0 (k). The dashed line represents the median of the volume of the geobodies contained in the reference image.

The number of geobodies follows a trend opposite to the other criteria. The best results are obtained with the set of weights $w_{1,2} = 2$: a mean number on the 100 simulation of 318 geobodies, while 608 geobodies are observed on average for $w_{1,2} = 1$, with a target number in the reference image of 44 geobodies. This is related to the overestimation of the pore space obtained with $w_{1,2} = 2$ (Fig. 14). Thus, the results that are closer to the reference image according to the majority of the comparison criteria are obtained with $w_{1,2} = 1$.

The issue raised by comparison between the results for $w_{1,2} = 1$ and 2 shows that the best parameters set cannot be selected by considering only one parameter (i.e., the number of geobodies); instead, a number of parameters that entails different aspects and descriptions of the results must be considered.

For the following comparisons with other methods, the `paBor` method is used with weights $w_i = 1$ for each i.

8.2. Methods comparison on the micro CT image

The method that better reproduces the features of the reference training image is the `s2Dcd`. The visual inspection of the simulations (Fig. 16), the connectivity functions (Fig. 17), and the data of Table 1 provides strong support for this. Note that the number of geobodies obtained with this method is closer to that of the reference training image than the one obtained by the standard MPS simulation approach but with search templates reduced to a merge of two or three 2D search templates with different orientations (Table 1, methods (i) and (j)). Moreover, the computational effort of the method `s2Dcd` is more than two orders of magnitude smaller than for methods (i) and (j). In general, this is true for all the aggregation methods tested in this case study: the simulation time required by the methods `paLin`, `paBor`, and `s2Dcd` is from two to three orders of
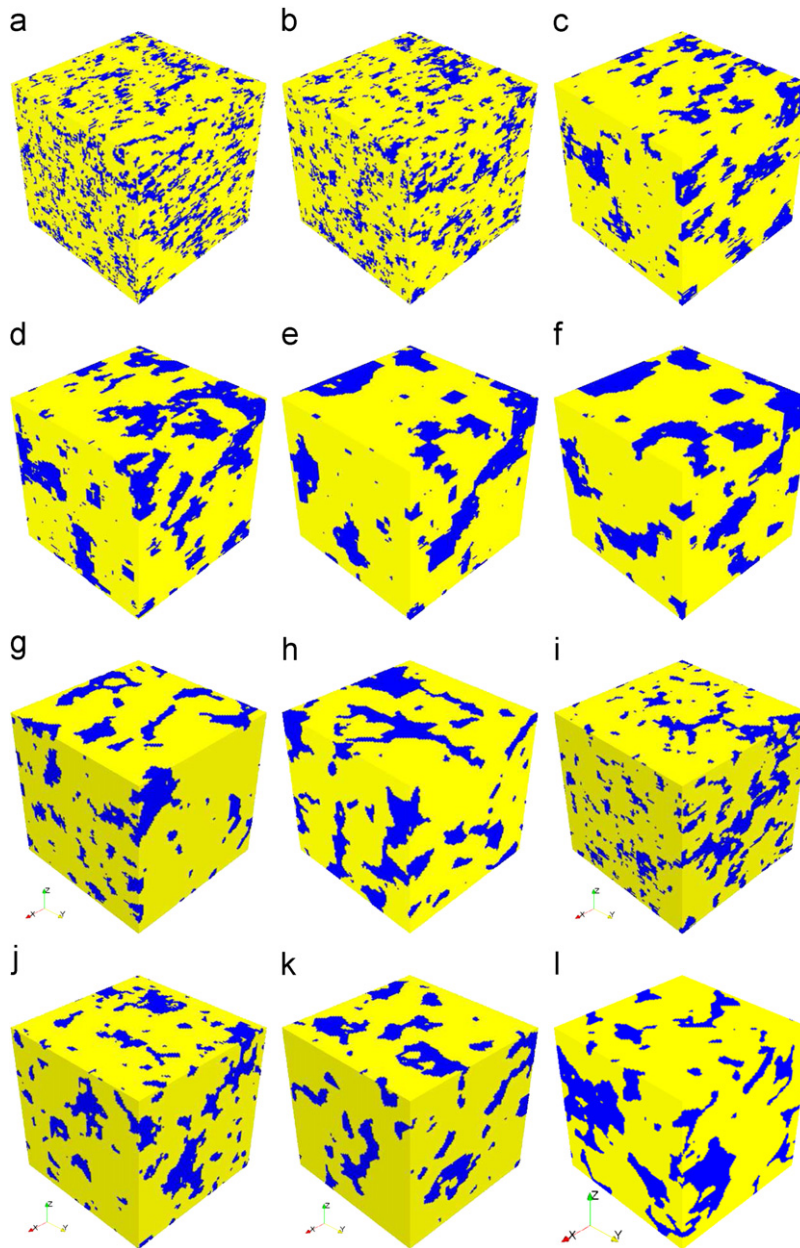
Fig. 16. Simulations obtained by different methods for the micro CT image. When two 2D TI are used, they are oriented along the planes xy and xz. The methods that adopt weights proportional to the size of the data event and with $S_w = \sum_i w_i$ with $w_i = 1$ are denoted by "$\propto d$." The figure (l) is considered as a reference and it is used as TI for the simulations (i), (j), and (k). (a) `paLin`, $w_i = 0.5$, 2 TI, (b) `paLin`, $w_i \propto d$, 2 TI, (c) `paBor`, $w_i = 1.0$, 2 TI, (d) `paBor`, $w_i \propto d$, 2 TI, (e) `paBor`, $w_i = 1.0$, 3 TI, (f) `paBor`, $w_i \propto d$, 3 TI, (g) `s2Dcd`, 2 TI, (h) `s2Dcd`, 3 TI, (i) 3D TI, $\tau_{xy} \cup \tau_{xz}$, (j) 3D TI, $\tau_{xy} \cup \tau_{xz} \cup \tau_{yz}$, (k) 3D TI, $\tau = 7 \times 7 \times 7$ and (l) reference.

magnitude smaller than that required by a standard multiple-point simulation with a complete 3D training image.

The comparisons between the linear aggregation method (paLin, tests (a) and (b)) and Bordley's formula (paBor, tests (c), (d), (e), (f)) show that with the second method it is possible to obtain results that are more satisfactory in terms of visual aspect (Fig. 16), connectivity function (Fig. 17), and number of geobodies

(Table 1). Therefore in this situation Bordley's aggregation formula is more suitable than the linear formula.

Of course, the results obtained using three training images are more realistic than the ones obtained using only two training images (Fig. 16, Table 1). Moreover, in the case of the s2Dcd, using three training images does not entail a greater global computational effort as it does for the methods based on the
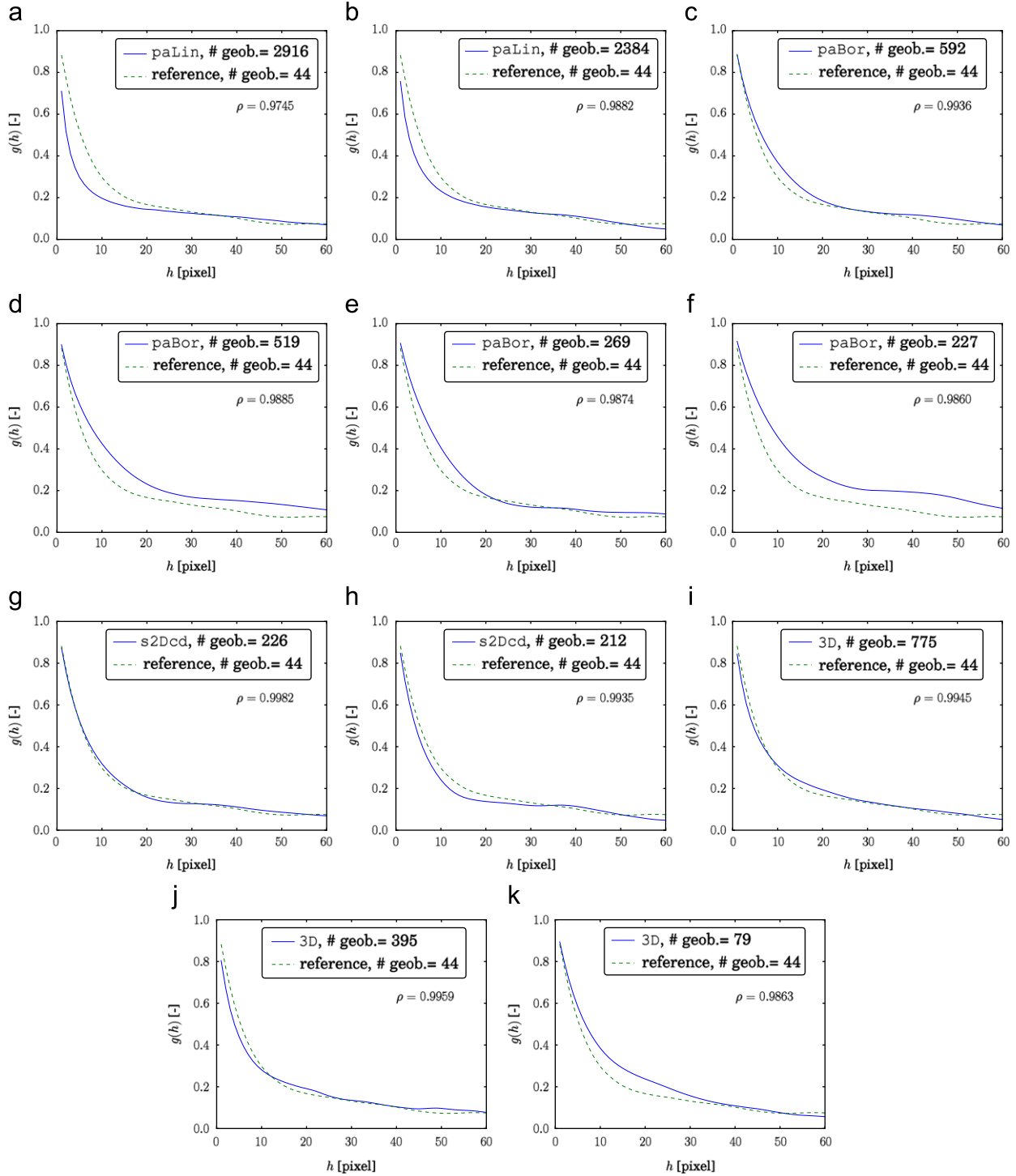
Fig. 17. Connectivity functions along the x-axis for the 3D tests performed for the micro CT image. The dashed curve is the function computed on the reference training image. The correlation coefficient between the curves and the number of geobodies is reported. (a) paLin, $w_i = 0.5$, 2 TI, (b) paLin, $w_i \propto d$, 2 TI, (c) paBor, $w_i = 1.0$, 2 TI, (d) paBor, $w_i \propto d$, 2 TI, (e) paBor, $w_i = 1.0$, 3 TI, (f) paBor, $w_i \propto d$, 3 TI, (g) s2Dcd, 2 TI, (h) s2Dcd, 3 TI, (i) 3D TI, $\tau_{xy} \cup \tau_{xz}$, (j) 3D TI, $\tau_{xy} \cup \tau_{xz} \cup \tau_{yz}$ and (k) 3D TI, $\tau = 7 \times 7 \times 7$.

aggregation of probability. Indeed, even if more simulation steps are required for the `s2Dcd` method when three training images are used, the simulation time decreases faster than in the case where only two training images are used, because of the faster increase of the number of conditioning data. This can result in total computing times that are smaller when three training images are used (Table 1).

Selecting the weights as proportional to the size of the data event did not improve the results noticeably. When this option is used (Table 1, tests (b), (d), and (f)), the number of geobodies is slightly closer to the target number of geobodies, but at the same time the proportions get farther from the target proportions. The observation concerning the set of weights $w_{1,2} = 2$ made in the previous section support this hypothesis.

Some tests are performed using the method based on list merging for compatible data events. They are not reported because the quality of the simulations was not comparable with that of the simulations obtained with the other methods. Another drawback of this method is that when there is high compatibility among the lists that are merged, the size of the merged list becomes unacceptable.

This problem can be faced in different ways. One possibility is to increase the size of the search templates; this entails a reduction in the number of compatible data events. Another possibility is to develop some criteria to reduce the size of the lists in such a way that the final result is not sensibly altered and the information content is preserved.

An additional issue related to this method is how to compute the occurrences counters in the merged list as a function of the counters of the departure lists.

Despite its drawbacks, this method reveals its usefulness when it is necessary to check the compatibility among 2D training images that are selected to represent different aspect of a 3D training image. Section 9 is devoted to this last aspect.

8.3. Brussels sand sedimentary structures

With the method `s2Dcd` it is possible to obtain credible 3D simulations of the sedimentary structures depicted along two different sections by the training images in Fig. 10. These simulations contain less noise than the ones obtained with the method `paBor` (Fig. 18). Note that the slices in Fig. 18 are not selected looking for the best or the worst situation, but trying to depict an equilibrated outline.

The better simulation quality that is observed for the method `s2Dcd` in Fig. 18 is confirmed by the relative error ε (Fig. 19). This last is computed by comparing the equivalent conductivity tensor of the training image of Fig. 10a and the concerning components of the tensors computed for the 3D simulations. The cluster of simulations performed with the method `s2Dcd` has a lower relative error than the corresponding cluster of simulations obtained with the `paBor` method.

8.4. The Herten aquifer analog

The results obtained for the Herten aquifer analog are compared visually (Fig. 20). The ones obtained with the `paBor` method contain a noisy component that is less evident in the simulations obtained with the other methods. With the method `s2Dcd` the results obtained considering or not the coordinate z as an auxiliary variable (Fig. 20e and f, respectively) are comparable; they make it possible to obtain features that are very close to the ones of the reference image (Fig. 20b).

To compare the features of the simulated image in this case study using the same criteria as used for the micro CT images can be tedious. For example, the connectivity function can be computed for each of the 10 hydrofacies used in the simulations, and the same for the number of geobodies and the proportions,
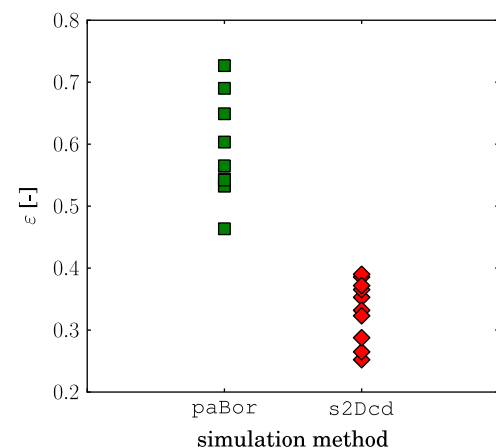

Fig. 19. The relative error ε on **K** computed for 10 simulations performed with the `paBor` method and 10 simulations with the `s2Dcd` method. The reference conductivity tensor is computed on the training image of Fig. 10a.
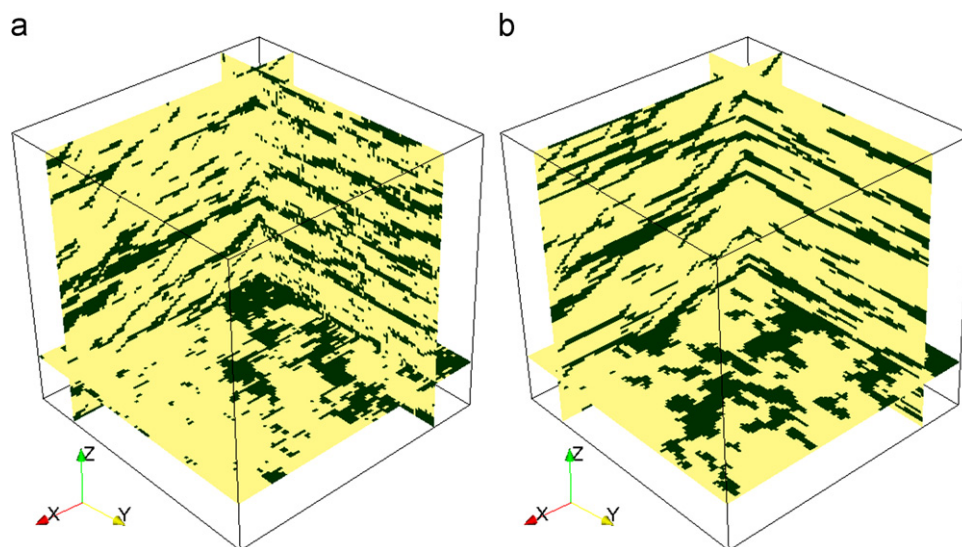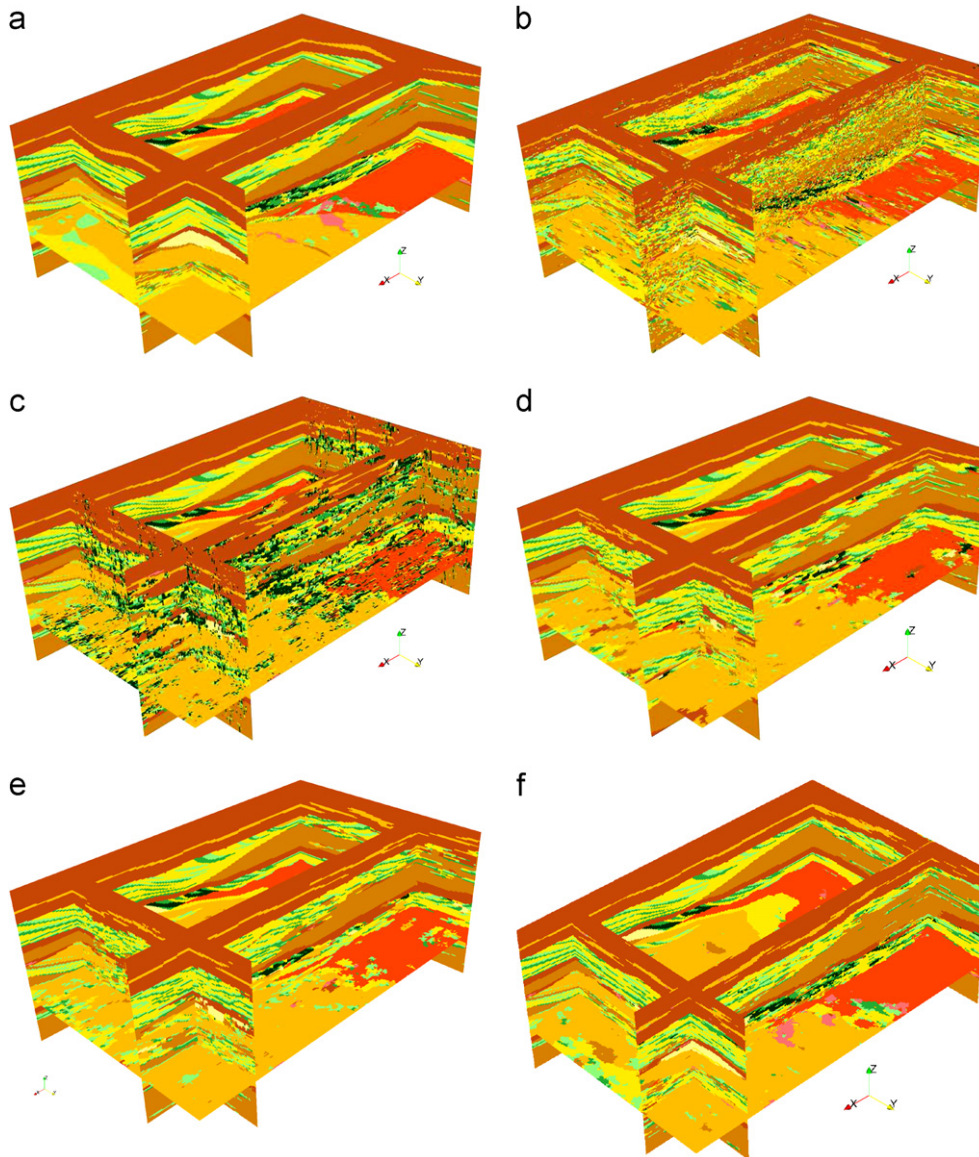

Fig. 18. Simulations of the Brussels Sands environment (training images in Fig. 10) obtained using (a) the `paBor` method and (b) the `s2Dcd` method.

Fig. 20. The simulations obtained with the different methods for the Herten aquifer analog: (a) hierarchical technique (considered as reference, Comunian et al., 2011); (b) transition probability/Markov chain approach (Maji and Sudicky, 2008); (c) `paBor` approach; (d) `s2Dcd` with no auxiliary variables; (e) `s2Dcd` with the coordinate *z* as auxiliary variable; (f) with the use of (a) as 3D training image. The background section along *xz* contains one of the six sections used as conditioning data, while the other section is placed in between two parallel sections of conditioning data.

but these results cannot be compared as easily as for a binary image. For this reason, another criterion is used. The comparisons are made by computing for each simulation an equivalent conductivity tensor $\mathbf{K}$ and its distance in terms of the Frobenius norm to the equivalent conductivity of the reference image. This distance ε, normalized by the value of the Frobenius norm of the reference, provides a global parameter that allows quantitative comparison of the characteristics of the simulation results in term of hydraulic parameters (Fig. 21).

The distance from the reference image makes it possible to distinguish a cluster of simulations performed with the `paBor` method, while the results obtained with the method `s2Dcd` with and without the auxiliary variable *z* have the same distance from the reference, five times smaller than the distance obtained with the `paBor` method. The distance of the methods `s2Dcd` is comparable with that of the simulation method TProgs (transition probability/Markov chain, TP/MC, Carle, 1996; Maji and Sudicky, 2008). The method that makes it possible to obtain a conductivity tensor closer to the reference is the one that uses it as a 3D training image.

The computation times, obtained with an Intel Xeon 3.2-GHz workstation with 2 GB of RAM, are reported in Table 2. The simulation times required for simulations with the methods `paBor` and `s2Dcd` are comparable (2.1 and 3.5 h respectively, with two 2D search templates 7×7), while the time required for a simulation that uses a 3D training image is one order of magnitude greater, using a search template $3 \times 3 \times 3$. If a template of size $7 \times 7 \times 7$ is used for the simulation with the 3D training image, then the computing times are more that two orders of magnitude greater than the times required by the `paBor` and `s2Dcd` methods.

Note that the computation times reported here and in Table 2 correspond to serial implementations of the codes for the MPS simulation. Of course, the parallel implementations of the MPS would have provided smaller computational times (Straubhaar et al., 2011), but some of the simulation tools developed for this study are serial only, for the moment. Therefore, the simulations are performed using serial implementations, in order to have a fair comparison.
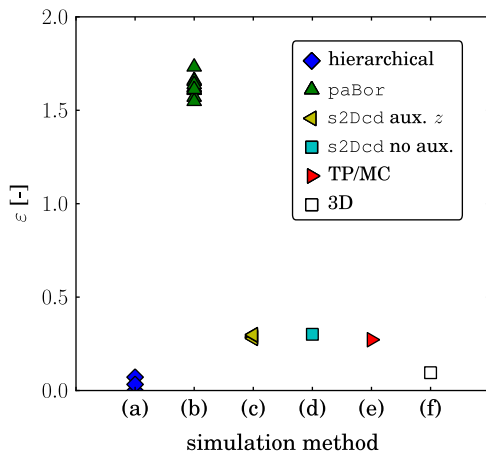
9. Check compatibility among training images

In the examples treated in this chapter, the 2D training images that are used are compatible with each other because they are either from the same 3D image, or observed in the field from a single geological environment. However, in practice, the user is left with providing two or more independent training images, and it may happen that they are incompatible.

Here we discuss how the technique to merge the lists proposed in Section 3.3 can be used to test their compatibility.

Here the tests are performed on two training images. One training image is the training image containing channels (Fig. 22h), introduced by Strebelle (2002), considered as oriented

along the xz plane. This image is associated with the list $\mathcal{L}_1$. The other training image, oriented along the yz plane, is chosen among the different images obtained by eroding or dilating the image used along the plane xz (Fig. 22, list $\mathcal{L}_2$).
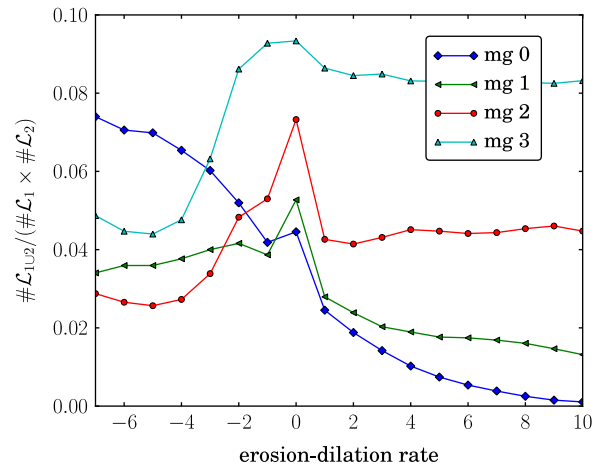
We then define an indicator of compatibility between the training images as the ratio between the size of the merged list $\mathcal{L}_{\text{merged}}$ and the size that the list would have in case all the data events were compatible, that is, $\#\mathcal{L}_1 \times \#\mathcal{L}_2$ (Fig. 23). The size of $\mathcal{L}_{\text{merged}}$ alone cannot provide a clear indication, because it depends both on the sizes of $\mathcal{L}_2$ and $\mathcal{L}_1$, and on the number of combined compatible data events.

Fig. 23 shows, for different multigrid levels, the ratio

$$\#\mathcal{L}_{\text{merged}}/(\#\mathcal{L}_1 \times \#\mathcal{L}_2). \qquad (9)$$

For each multigrid level, a maximum of this ratio is observed in correspondence to the abscissa 0, that is, for the value of erosion/dilation which corresponds to the departure image, when the training images used along the two directions are the same. The maximum is less evident for the first multigrid level (Fig. 23 mg 0) because for the given data template and the given training image the number of data events is too small.

Even if this feature should be confirmed by other case studies, the results shown here indicate that this criterion could be used as a comparison criterion to check the compatibility of 2D training images to be used for a joint simulation of a 3D block.

Fig. 21. The relative error in K computed using as reference one realization of the Herten aquifer analog obtained with a hierarchical technique (Comunian et al., 2011) for (a) 2 other realizations obtained by the hierarchical approach; (b) 10 realizations obtained with `paBor`; (c) 10 realizations obtained with `s2Dcd` and z as auxiliary variable; (d) 1 realization obtained with `s2Dcd` and no auxiliary variable; (e) 1 realization obtained with a TP/MC approach (Maji and Sudicky, 2008); (f) 1 realization obtained using the reference image as TI.

Table 2
Computation times for the different methods used to simulate the Herten aquifer analog.

Data template	Multigrid levels	Method	Time (h)
$7 \times 1 \times 7$, $1 \times 7 \times 7$	3	paBor	2.1
$7 \times 1 \times 7$, $1 \times 7 \times 7$	3	s2Dcd	3.5
$3 \times 1 \times 3$, $1 \times 3 \times 3$	6	s2Dcd	0.2
$3 \times 3 \times 3$	6	3D	22.9
$7 \times 7 \times 7$	3	3D	> 216.0

Fig. 23. The ratio between the size of the merged list $\mathcal{L}_{1\cup2}$ and the size of the merged list under the hypothesis that all the elements of the two merged lists $\mathcal{L}_1$ and $\mathcal{L}_2$ are compatible. The list $\mathcal{L}_1$ is computed from the training image of Fig. 22h, while the list $\mathcal{L}_2$ is obtained from the same image eroded or dilated (Fig. 22). The sizes of the merged lists are computed for four different multigrid levels.
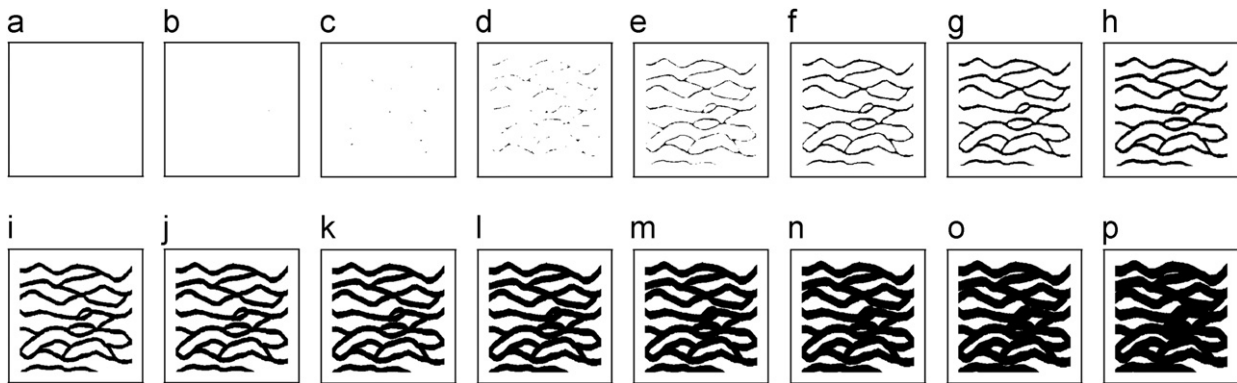
Fig. 22. The reference training image used along the plane xz (h). The images (a)–(g) and (i)–(p) are respectively obtained by eroding ($-$) or dilating ($+$) the reference training image (h) by a number of successive iterations (shown in the subcaptions). (a) -7, (b) -6, (c) -5, (d) -4, (e) -3, (f) -2, (g) -1, (h) ref., (i) $+1$, (j) $+2$, (k) $+3$, (l) $+4$, (m) $+5$, (n) $+6$, (o) $+7$ and (p) $+8$.

Another application could be in synergy with other methodologies (for example, Boisvert et al., 2007) aimed at the selection of TI from comparisons with data statistics.

10. Conclusion

This work demonstrates that even without the use of a 3D training image, it is possible to obtain 3D simulations that reproduce the main characteristics of simulations obtained with standard multiple-point statistics (standard in the sense that they use a complete 3D training image).

One of the novel methods proposed here (s2Dcd) produces simulations that are closer to the reference 3D image than methods based on the aggregation of probability, according to most of the evaluation criteria considered here. Moreover, for the application of the s2Dcd method, it is not necessary to define additional parameters such as the weighting factors required for the application of the paBor method.

The other novel method proposed, based on the merging of lists of compatible data events, does not make it possible to obtain realistic simulation results. However, it is based on a principle that is useful and easy to use for determining when two (or more) 2D training images can be considered compatible for the joint description of 3D geological structures.

An important aspect of the methods proposed here is their computational efficiency. Even if, of course, these methods cannot reach the same accuracy in the reproduction of complex geological patterns of MPS technique with a 3D training image, the computation time is from two to four orders of magnitude smaller than for the methods using a full 3D image. Therefore, with these methods, it is possible to approach problems that are prohibitive for the standard MPS simulation in terms of CPU time required. For example, with the same computational effort, with these methods it is possible to select search templates bigger than the ones allowed by the standard MPS. This computational efficiency makes it possible to include MPS in Monte Carlo and uncertainty evaluation simulation frameworks, or in inverse problems that require a high number of realizations performed within a reasonable computation time.

Moreover, with the method s2Dcd, it is possible to face simulation problems that would normally require huge RAM resources, because the simulation process is reduced to a sequence of 2D MPS simulations. This made it possible, for example, to simulate domains of about 200 million voxels, using two 2D training images of sand lenses from a clay till outcrop in Denmark mapped by Kessler et al. (2010).

Still, the s2Dcd method has some limitations. Even if the overall quality of the results presented here is good, some artifacts can be observed. For example, looking the 2D slices along their simulation sequence, a gradual deterioration of their quality can be noticed. While at the beginning the 2D simulations can easily be constrained by the available conditioning data, the constraints can become too strong when many conditioning data have been simulated. The strength of the constraints is due to the fact that the 2D simulations are obtained by ignoring lateral correlation (the only information used in the conditioning data on the same simulation surface), and therefore some incoherent features can arise during the simulation procedure. These drawbacks can be reduced by selecting a suitable simulation sequence of the surfaces, but further investigations are required to minimize them.

The methodologies presented here are based on the assumption that the 3D model can be represented, along some directions, by one or more models with a smaller dimension. Of course this assumption should always be checked and considered according to the model's goals and aims.

Acknowledgments

References

Abrahamsen, P., Fjellvoll, B., Hauge, R., 2007. Process based stochastic modelling of deep marine reservoirs. In: Petroleum Geostatistics September.

Allard, D., 1992. On the connectivity of two random set models: The truncated gaussian and the boolean. In: Geostatistics Troia '92. Kluwer, Dordrecht, The Netherlands, pp. 467–478.

Allard, D., D'Or, D., Froidevaux, R., 2011. An efficient maximum entropy approach for categorical variable prediction. European Journal of Soil Science 62 (3), 381–393. doi:10.1111/j.1365-2389.2011.01362.x.

Allard, D., Comunian, A., Renard, P. Probability aggregation methods in geoscience: a comparative study, in preparation.

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

Boisvert, J., Pyrcz, M., Deutsch, C., 2010. Multiple point metrics to assess categorical variable models. Natural Resources Research 19, 165–175.

Boisvert, J.B., Pyrcz, M.J., Deutsch, C.V., 2007. Multiple-point statistics for training image selection. Natural Resources Research 16 (4), 313–321.

Bordley, R.F., 1982. A multiplicative formula for aggregating probability assessments. Management Science 28, 1137–1148. doi:10.1287/mnsc.28.10.1137.

Caers, J., 2006. A General Algorithm for Building 3D Spatial Laws from Lower Dimensional Structural Information. Technical Report, Stanford Center for Reservoir Forecasting.

Carle, S., 1996. A Transition Probability-Based Approach to Geostatistical Characterization of Hydrostratigraphic Architecture. Ph.D. Thesis, Department of Land, Air and Water Resources, University of California Davis.

Chugunova, T.L., Hu, L.Y., 2008. Multiple-point simulations constrained by continuous auxiliary data. Mathematical Geosciences 40 (2), 133–146.

Comunian, A., 2011. Probability Aggregation Methods and Multiple-point Statistics for 3D Modeling of Aquifer Heterogeneity from 2D Training Images. Ph.D. Thesis, University of Neuchâtel, January.

Comunian, A., Renard, P., Straubhaar, J., Bayer, P., 2011. Three-dimensional high resolution fluvio-glacial aquifer analog. Part 2. Geostatistical modeling. Journal of Hydrology 405 (1–2), 10–23.

Cornaton, F., 2007. Ground Water: A 3-D Ground Water Flow and Transport Finite Element Simulator. Reference Manual. Technical Report, University of Neuchâtel, Switzerland.

De Iaco, S., Maggio, S., 2011. Validation techniques for geological patterns simulations based on variogram and multiple-point statistics. Mathematical Geosciences 43, 483–500.

de Marsily, G., Delay, F., Goncalves, J., Renard, P., Teles, V., Violette, S., 2005. Dealing with spatial heterogeneity. Hydrogeology Journal 13 (1), 161–183.

Guardiano, F.B., Srivastava, R.M., 1993. Multivariate geostatistics: Beyond bivariate moments. In: Soares, A. (Ed.), Geostatistics: Troia '92, vol. 1. Kluwer, Dordrecht, The Netherlands, pp. 133–144.

Hu, L.Y., Chugunova, T.L., 2008. Multiple-point geostatistics for modeling subsurface heterogeneity: A comprehensive review. Water Resources Research 44 (W11413), 14.

Huysmans, M., Dassargues, A., 2010. Application of multiple-point geostatistics on modelling groundwater flow and transport in a cross-bedded aquifer. In: Atkinson, P.M.M., Lloyd, C.D.D. (Eds.), geoENV VII. Geostatistics for Environmental Applications. Quantitative Geology and Geostatistics, vol. 16. Springer, Netherlands, pp. 139–150.

Huysmans, M., Peeters, L., Moermans, G., Dassargues, A., 2008. Relating small-scale sedimentary structures and permeability in a cross-bedded aquifer. Journal of Hydrology 361 (1–2), 41–51.

Journel, A., 2002. Combining knowledge from diverse sources: An alternative to traditional data independence hypotheses. Mathematical Geology 34, 573–596.

Journel, A., Zhang, T., 2006. The necessity of a multiple-point prior model. Mathematical Geology 38 (5), 591–610.

Kessler, T., Klint, K.E.S., Renard, P., Nilsson, B., Bjerg, P.L., 2010. Geostatistical description of geological heterogeneity in clayey till as input for improved characterization of contaminated sites. In: GQ10 Groundwater Quality

Management in a Rapidly Changing World. Proceedings of the 7th International Groundwater Quality Conference, Zurich, 13–18 June 2010.

Koltermann, C.E., Gorelick, S.M., 1996. Heterogeneity in sedimentary deposits: A review of structure-imitating, process-imitating, and descriptive approaches. Water Resources Research 32 (9), 2617–2658.

Krishnan, S., 2008. The tau model for data redundancy and information combination in earth sciences: Theory and application. Mathematical Geosciences 40, 705–727.

Le Coz, M., Genthon, P., Adler, P.M. Multiple-point statistics for modeling facies heterogeneities in a porous medium: The Komadugu–Yobe alluvium, Lake Chad basin. Mathematical Geosciences, in press. doi:10.1007/s11004-011-9353-6.

Liu, Y., 2006. Using the Snesim program for multiple-point statistical simulation. Computers & Geosciences 32 (10), 1544–1563.

Maharaja, A., 2008. Tigenerator: Object-based training image generator. Computers Geosciences 34 (12), 1753–1761.

Maji, R., Sudicky, E., 2008. Influence of mass transfer characteristics for DNAPL source depletion and contaminant flux in a highly characterized glaciofluvial aquifer. Journal of Contaminant Hydrology 102 (1–2), 105–119.

Michael, H.A., Li, H., Boucher, A., Sun, T., Caers, J., Gorelick, S.M., 2010. Combining geologic-process models and geostatistics for conditional simulation of 3-D subsurface heterogeneity. Water Resources Research 46 (5), W05527.

Okabe, H., Blunt, M.J., 2004. Prediction of permeability for porous media reconstructed using multiple-point statistics. Physical Review 70 (6).

Okabe, H., Blunt, M.J., 2007. Pore space reconstruction of vuggy carbonates using microtomography and multiple-point statistics. Water Resources Research 43. doi:10.1029/2006WR005680.

Pyrcz, M., Strebelle, S., 2008. Event-based geostatistical modeling. In: Ortiz, J., Emery, X. (Eds.), Geostatistics Santiago 2008. Springer, Netherlands.

Pyrcz, M.J., Catuneanu, O., Deutsch, C.V., 2005. Stochastic surface-based modeling of turbidite lobes. AAPG Bulletin 89 (2), 177–191.

Renard, P., Genty, A., Stauffer, F., 2001. Laboratory determination of the full permeability tensor. Journal of Geophysical Research 106 (B11), 26,443–26,452.

Rubin, Y., Gómez-Hernández, J.J., 1990. A stochastic approach to the problem of upscaling of conductivity in disordered media: Theory and unconditional numerical simulations. Water Resources Research 26, 691–701.

Straubhaar, J., Renard, P., Mariethoz, G., Froidevaux, R., Besson, O., 2011. An improved parallel multiple-point algorithm using a list approach. Mathematical Geosciences 43 (3), 305–328. doi:10.1007/s11004-011-9328-7.

Strauffer, D., Amnon, A., 1994. Introduction to Percolation Theory. Taylor and Francis.

Strebelle, S., 2000. Sequential Simulation Drawing Structure from Training Images. Ph.D. Thesis, Stanford University, Stanford, CA, USA.

Strebelle, S., 2002. Conditional simulation of complex geological structures using multiple-point statistics. Mathematical Geology 34, 1–21.