Parameterization of training images for aquifer 3D facies modeling integrating geological interpretations and statistical inference

Sanjeev Kumar Jha[1],*, Alessandro Comunian[1,2], Gregoire Mariethoz[1,3,4], Bryce F. J. Kelly[1]

[1]The University of New South Wales, Connected Waters Initiative Research Center, National Center for Groundwater Research and Training, Sydney, Australia.

[2]Dipartimento di Scienze della Terra "A.Desio", Università degli Studi di Milano, Italy.

[3]ETH Zurich, Geological Institute, Zurich, Switzerland.

[4]University of Lausanne, Institute of Earth Surface Dynamics, Lausanne, Switzerland.

Email s.jha@unsw.edu.au.

*corresponding author

Abstract

We develop a stochastic approach to construct channelized 3D geological models constrained to borehole measurements as well as geological interpretation. The methodology is based on simple 2D geologist-provided sketches of fluvial depositional elements, which are extruded in the 3rd dimension. Multiple-point geostatistics (MPS) is used to impair horizontal variability to the structures by introducing geometrical transformation parameters. The sketches provided by the geologist are used as elementary training images, whose statistical information is expanded through randomized transformations. We demonstrate the applicability of the approach by applying it to modeling a fluvial valley filling sequence in the Maules Creek catchment, Australia. The facies models are constrained to borehole logs, spatial information borrowed from an analogue and local orientations derived from the

present-day stream networks. The connectivity in the 3D facies models is evaluated using statistical measures and transport simulations. Comparison with a statistically equivalent variogram-based model shows that our approach is more suited for building 3D facies models that contain structures specific to the channelized environment and which have a significant influence on the transport processes.

Introduction

High-resolution 3D geological models are important tools for the detailed interpretation of subsurface structures [*Raiber et al.*, 2012], for determining the hydrological interaction between the surface water bodies and aquifers, or for modeling groundwater contaminant transport [*Wycisk et al.*, 2009] and density-driven flow [*Kerrou and Renard*, 2010]. Geological models can be constructed by integrating borehole facies logs with geophysical data sets, such as seismic cross-sections, structural data from geological maps and topographical information from digital elevation models [*Maxelon and Mancktelow*, 2005; *Maxelon et al.*, 2009; *Wang and Huang*, 2011; *Wu et al.*, 2005]. In this paper, we focus on the construction of stochastic facies models of fluvial valley filling sediments based on borehole data. There are two dominant approaches to stochastic facies modeling: (i) cell-based geostatistical modeling; and (ii) object-based stochastic modeling. Although being widely used, object-based methods [*Deutsch and Wang*, 1996; *Holden et al.*, 1998] can be difficult to condition to a large number of well log data. Furthermore, in the case of channelized structures, object-based models often generate long and continuous channels, whereas a typical alluvium aquifer is not as continuous, presenting channels interrupted by erosion and redeposition processes and by the formation of oxbows. Moreover, object-based models are restricted to a limited number of geometric shapes, and therefore may not be able

to reproduce certain geological settings [*Comunian et al.*, 2011; *Feyen and Caers*, 2006; *Michael et al.*, 2010].

Sequential indicator simulation (SIS), part of the kriging algorithm family, is the most popular cell-based geostatistical method for categorical variables [*Deutsch*, 2002]. It builds on variogram models that conveniently allow describing the spatial structure with a limited number of parameters [*Emery and Parra*, 2013]. However, the convenience achieved by these parametric approaches comes at the price of a limited representation of complex heterogeneity [*Schlüter and Vogel*, 2011; *Western et al.*, 2001; *Zinn and Harvey*, 2003]. A classical justification is that in percolation theory, breakthrough can be achieved in random 3D media even with low percolation thresholds [*Allard et al.*, 2011], therefore it has been argued that a detailed characterization of the connectivity is not critical, and that natural spatial heterogeneity can be adequately represented with a small number of parameters. This argument has however been strongly challenged by geologists because a realistic representation of geological structures can behave very differently than idealized random media [*Neuweiler and Vogel*, 2007; *Zinn and Harvey*, 2003], with for example very thin elongated hydraulically conductive features can result in lateral and vertical hydraulic connectivity and preferential flow within a sedimentary sequence even if its overall proportion is much lower than the percolation threshold [*Huysmans and Dassargues*, 2012]. Moreover, a practical difficulty is that it can be very difficult to model the horizontal variogram characteristics based only on sparse borehole data, resulting in considerable uncertainty on the spatial structure.

Multiple-point geostatistics (MPS) was introduced to overcome the limitations of traditional geostatistical approaches [*Caers and Zhang*, 2005; *Guardiano and Srivastava*, 1993] and has been proven to be a powerful tool for modeling subsurface heterogeneities [*Zhou et al.*, 2012]. The key concept of MPS is to use a training image (TI), which constitutes

a prior description of the architectural features expected in a given sedimentological setting. In 2D studies, the training images are borrowed from available outcrop analogues [*Bayer et al.*, 2011] or remote sensing images [*Ge and Bai*, 2011; *He et al.*, 2013; *Jha et al.*, 2013; *Le Coz et al.*, 2011; *Wojcik et al.*, 2009]. In the absence of extensive remote sensing information, method have been proposed to select an appropriate TI from a group of potential TIs by comparing the spatial statistics of local conditioning data with those of the proposed TIs [*Boisvert et al.*, 2007; *de Iaco and Maggio*, 2011; *Perez et al.*, 2014], or by integrating indirect data through an inverse problem [*Khodabakhshi and Jafarpour*, 2013; *Suzuki and Caers*, 2006]. However, in most geological applications there are insufficient data (only borehole logs located far apart) to infer an appropriate TI that is representative of the area of interest. In 3D, the selection of the TI is more challenging, let alone the construction of a set of 3D TIs to choose from. *Le Coz et al.* [2011] used a set of two-dimensional training images and derived the necessary statistics separately in the horizontal and vertical directions for modeling the facies heterogeneity. *Comunian et al.* [2012] developed a workflow to obtain 3D MPS simulations by constructing the 3D model slice-by-slice using horizontal and vertical 2D training images. The assumption is that the 2D vertical and horizontal training images are fully representative of the 3D heterogeneity along any plane.

Although conceptually simple, manually building a 3D TI based on geologists' and reservoir modelers' prior knowledge would be a lengthy geomodeling exercise. [*Comunian et al.*, 2014] used process-imitating methods to simulate 3D TIs for MPS simulations of an alluvium basin and reported that finding suitable parameters for reproducing the statistical features of available data set is not a straightforward task. *Mariethoz and Kelly* [2011] introduced the concept of elementary training images, which are simple and easy to create. Transform-invariant distances are used to extend the possible structures that can be produced from such a training image. This is achieved by adjusting geometrical transformation

parameters that define the search geometry. Modifying the rotation and affinity parameters enables the expansion of the suite of available patterns within a TI. When both rotation and affinity are applied simultaneously, the MPS simulation is able to generate shapes that resemble connected and elongated geological structures, while honoring the original sparse data set. In this study, we explore in detail how a parameterized training image can be used to obtain a 3D facies model in a fluvial valley filling sequence, which is a practical innovation over *Mariethoz and Kelly* [2011] and an alternative to[*Comunian et al.*, 2014]. This study demonstrates the applicability of the approach for real-world cases and provides methodological guidelines by (i) investigating ways to obtain elementary training images based on geologists' expertise; (ii) inferring the transformation parameters from an analogue; and (iii) combining information from different sources (boreholes, analogues, local orientations).

Methodology

Our methodology for generating a 3D facies model is based on the Direct Sampling (DS) algorithm, a multiple-point simulation tool that uses a training image, a simulation grid and conditioning data as input. Comprehensive details on the DS algorithm are presented in *Mariethoz et al.* [2010]. A brief description of the DS algorithm is presented in Section A.1 of Appendix A. Section A.2 of Appendix A summarizes how rotation and affinity transformations are applied during the DS simulation to create new patterns not available in the training image. Provided the training image represents general characteristics of a given geological setting (facies proportions, elementary architectural shapes), it is possible to build very different models that respect these essential characteristics. Furthermore, it is also possible to include non-stationarity in the orientation of the geological structures by

employing locally varying transformation parameters in the simulation domain (see Section A.3 of Appendix A).

2.1 From 2D to 3D training image using extrusion and affine transformation

The methodology presented in this section is illustrated through a simple example. A geological sketch depicting the position of the channel is shown in Figure 1a. The 2D sketch can be drawn manually using an image editing software (e.g., Microsoft Paintbrush) or using any commercial software for computer-aided design (CAD). A 3D training image is created from this 2D sketch by extruding the cross-sectional shapes in the third dimension (Figure 1b). This is accomplished by stacking horizontally the 2D training image. As a result, the features that are drawn in 2D become elongated bodies. If the 2D sketch is made of simple shapes that can be represented analytically, such as in Figure 1a, one can also use an object-based model to obtain the 3D training image (e.g. TiGenerator, [*Maharaja and Journel*, 2005]). The extrusion approach requires more user input, but can create extruded channels training image from any non-parametric channel shape. The resulting 3D training image, which has straight channels, is not yet an accurate representation of geological structures, but it represents a notion of the horizontal connectivity that is expected in channelized formations. Since naturally occurring channels are seldom straight, the transform-invariant distances will later be used in the simulation to impair a meandering behavior to these structures.

Figure 1c to 1f represent the output from a single realization of DS with different transformation parameters. In Figures 1c and 1d, rotation tolerance of $\varepsilon = 45^o$ and 80° around the Z-axis have been used. The rotated channels result in new geological structures as shown in Figures 1c and 1d. In Figures 1e and 1f, affinity parameters of 2 and 0.5 are used in the Z-direction resulting into connectivity structures that are thicker and thinner in the resulting

simulation. By using both rotation around the Z-axis and affinity in the Z-direction, various kinds of geological structures can be obtained.

2.2 Workflow for obtaining 3D facies models from elementary sketches

Our workflow for generating 3D facies models consists of four main steps: (i) selection; (ii) extrusion; (iii) inference; and (iv) modeling. Figure 2 shows the workflow illustrating these steps.

Step I (selection): A set of 2D sketches is drawn, which are candidate training images provided by one or several geologists. The proportion of each facies, known from the boreholes, is kept identical in all sketches. The transition probabilities between facies are compared between the hand-drawn sketches and that of borehole logs. The transition probability is a conditional probability of having a type of facies (0 or 1) at any point given that a similar or different facies (0 or 1) is present at another location separated by a lag distance [*Carle and Fogg*, 1996]. This comparison, along with geological interpretation and expertise, is used for determining which sketch to select as most representative of the formation to be modeled. It is worth pointing out that instead of transition probability, an indicator variogram could have been used at this stage to select the training image.

Step II (extrusion): The selected sketch is extruded in the 3rd dimension to form a 3D training image. An alternative to the extrusion could be used to produce a 3D elementary TI using an object-based model, provided that the shape of the channels can be accurately described by objects parameters.

Step III (inference): Unconditional DS simulations are performed by changing the orientation and affinity parameters. The method of transform-invariant distances creates new patterns. Among these new patterns, by interaction and overlapping, cross-cutting of channels

are created. Supplementary Figure S1 demonstrates that the transform-invariant distances create cross-cutting channels even when the training image is only made of parallel channels.

Step IV (modeling): Conditional multiple-point simulations are generated with the selected training image (or an object-based training image equivalent when applicable) and the inferred transformation parameters.

2.3 Additional remarks

The extruded sketches (step II) could be in some cases obtained using object-based models. However, parameterizing the objects based on sparse boreholes can be challenging, even if analogues are available. For example, to obtain meandering channels in TiGenerator, the one needs to define: a) Type of geobodies (Sinusoid, Ellipsoid, half Ellipsoid, Cuboid, or user defined), number of geobodies, number of facies and their proportions. b) Depending upon the type of geobodies selected, the channel properties have to be defined. For example, in the case of sinusoids, length, width, thickness, orientation, amplitude and wavelength needs to be provided as input. Shapes of higher complexity can have more parameters. c) Interaction rules have to define whether channels can cross each other and erode or overlap. In our approach, only the borehole data and geological expertise is used through 2D sketches, without extensive parameterization.

The proposed workflow has the advantage of dealing with any complexity in the shapes of 2D TIs. For example, if the expected geology consists of sinuous channels with low-permeable (or high-permeable) discontinuous sheets at the bottom of each channel, our methodology would be very easy to put in place. To demonstrate the approach, a 2D TI is shown in the Figure 3a. The 2D TI is extruded and a rotation tolerance (ε) of 60^0 is applied to obtain a realization shown in Figure 3b. Obtaining such a model using an object-based

approach would require a high level of complexity of the objects, such as for example in the FLUVSIM code [*Deutsch and Tran*, 2002].

For step III, as an alternative to the sensitivity analysis, an optimization algorithm or a Markov chain Monte Carlo (MCMC) sampler could be used to determine the transformation parameters and their uncertainty.

Case study

3.1. Study area and available data

The proposed workflow is applied to the Maules Creek catchment located in western New South Wales, Australia (Figure 4). It is a sub-catchment of the Namoi catchment that has a surface area of 1500 km^2..The upper and lower reaches of the creek are largely ephemeral and the middle reach is perennial. Surface water and groundwater interactions have been extensively studied at Maules Creek [*Andersen and Acworth*, 2009; *Giambastiani et al.*, 2012; *McCallum et al.*, 2012]. The geology of the catchment is complex. The modern day sediment distribution within the catchment is controlled by the westerly following waters coming from the Nandewar Range via a series of small creeks (including Bibbla, Horsearm, Middle, and Maules Creeks) and on the western side the northerly flowing Namoi River. Based on the palaeovalley shape, a similar combination of westerly flowing creeks feeding into a major northerly flowing river existed in the past. Based on this information, in our modeling approach we assume that the paleo-orientation of the system is identical to the orientation that is observed today.

It is common throughout the Namoi to divide the valley filling sequence into the unconfined Narrabri Formation (approximately 30 m thick) and an underlying semi-confined Gunnedah Formation [*Gates*, 1980; *Williams et al.*, 1987]. However, within the Maules Creek catchment there is little hydrological evidence to support such a division. Within the

9

valley filling sequence there is a clear trend from a wet, high-energy Neogene environment at depth, to the modern semi-arid environment. Over time there has been a trend towards decreasing sand and gravel deposition [*Kelly et al.*, 2014]. There is an increase in the clay and silt content with the onset of aridity at the Pliocene/Pleistocene boundary (located approximately 30 to 40 meters below the ground surface in the vicinity of the Namoi River). The upper 6 meters of the sedimentary sequence consist predominantly of Vertosols. These soils have formed from a combination of fluvial and wind-blown clays and silts. There are 402 lithological logs available in the catchment, where lithofacies have been divided into two classes: clay and silt (low hydraulic conductivity, hereafter called clay) and sand and gravel (high hydraulic conductivity, hereafter called sand). The digital elevation model and locations of boreholes are shown in Figures 4b and 4c.

3.2. Selection of 2D hand-drawn sketches

The 2D sketch is a representation of sedimentary architectural features, which are deemed present according to a geological interpretation of the fluvial setting. This knowledge is established from field mapping or collated case studies [*Gibling*, 2006]. For this paper, 10 sketches with 60×60 pixels of size 20m × 1m in X and Z directions are considered, as shown in Figure 5. These sketches are saved as binary images with black and white colors representing clay and sand facies. Different types of in-filled channel and sheet bodies are generated to represent the expected sand elements. TI-1 and TI-6 consist of wide and deep channels; TI-2 to TI-4, TI-7 and TI-10 contain a mixture of wide, deep and shallow channels; most of the sand bodies in TI-5 are shallow and wide; TI-8 and TI-9 represent narrow and deep sand filled channels. At this stage, the 2D sketches are drawn considering the fraction of facies in the borehole data, which is 50% sand and 50% clay. The global

proportion of sand is kept identical in all sketches. In this example, facies 0 is clay and silt and the facies 1 represents sand or gravel.

To sort the most relevant scenarios for the case study considered, the vertical transition probabilities (TP) of these sketches are compared with that of borehole data. The resulting plots are shown in Figure 6. The TP represent how one type of facies is in contact with other type. The TP of clay over clay is represented as TP-00 and TP of sand over sand is represented as TP-11. The plots of TP-00 and TP-11 corresponding to selected training images are shown in Figures 6a and 6b respectively. The plots of TP01 and TP10 are not shown because they are the complement of TP00 and TP11. From the TP curves, it is clear that TI-7 and TI-9 have similar TPs as that obtained in the borehole data. This comparison allows considering the opinion of an expert geologist to complement statistical measures. This point will be further developed in the following sections. We select sketches TI-7 and TI-9 for further analysis. These will be referred to as training images in the following sections.

3.3 TI extrusion

3D training images consisting of 60×50×60 pixels in X, Y and Z-directions respectively are created from previously selected 2D sketches (TI-7 and TI-9 consisting of 60 × 60 pixels) by extruding them in the Y-direction. The resulting structures (Figure 7) show a high degree of connectivity in the Y-direction. The figure has been vertically exaggerated by a factor 5 for better visualization. Figure 7a represents a scenario with fewer wide and deep channels. Figure 7b shows channels that are narrower and deeper with some degree of connection between the different channels. Note that both TIs have similar TP curves (Figure 6) and facies proportions. TI-9 is however selected for further simulation in preference to TI-

7, because it corresponds more the geological setting described in the previous section, in particular the smaller size of the sand bodies is more consistent with field-based observations.

3.4 Inference of the transformation parameters

The transition probabilities discussed in the previous section gave a measure of how facies are arranged in the vertical direction only. The boreholes density is not sufficient to provide information about the horizontal connectivity, or to measure the spatial correlation using variograms. Instead of using direct measurements, an analogue (presented in Figure 8) is borrowed from a fluvial setting analogous to the higher energy/wetter climate that was present when most of the valley filling sequence was deposited within the Maules-Creek catchment. Here the analogue is a representation of the geology that used to infer the transformation parameters such that its horizontal connectivity can be embedded in the final 3D facies models. The connectivity function of *Pardo-Igúzquiza and Dowd* [2003] is computed on the analogue, which is the probability of a cell with a particular facies being connected to another cell of the same facies as a function of their separation distance *h*.

We use unconditional simulation for the inference of the transformation parameters. The unconditional simulations are carried out in a small 3D domain, and used to compare the resulting statistics with the statistics observed in the borehole data and the geological analogue. A sensitivity analysis is performed on a domain of size 60×100×60 grid nodes in the X, Y and Z-directions respectively. The maximum search distances are defined as 30×30×30. Based on the practical guide of *Meerschman et al.* [2013] to performing DS simulation, the maximum number of neighbors used in the simulation is set to 30, the distance threshold is set to 0.1 and maximum fraction of the TI to scan as 0.5. The implementation then find neighborhoods made of a maximum of 30 pixels, as long as they are less distant than the maximum search distance. For the unconditional simulations, the

overall rotation of the patterns α is kept fixed at 0 degrees in X, Y, and Z directions, meaning that the general orientation of the structures is kept identical as in the extruded training image. To impart sinuosity, a rotation tolerance of 120 degrees around the Z axis is allowed. This tolerance value of 120 degrees is selected based on observation of the meander angles in the analogue. It means that even if we use a training image which contains straight channels, the DS algorithm will search for a pattern in the training image with a flexibility on the orientation provided by the tolerance value, i.e. patterns are searched up to a rotation around the vertical axis, as long as this rotation is in the range [-120 120] degrees. Therefore, the training images with straight channels represented in Figure 7 is enriched with patterns rotated in all the directions considered, which can potentially represent meandering channels. The parameters used in unconditional simulation are listed in Table 1.

We tested the effect of affinity by assigning its values in the X and Y directions ranging from 0.1 to 1.0, with a constant value of 1.0 in the Z-direction. The result is that all structures are stretched or expanded by this factor. A sensitivity analysis on the affinity parameters is performed by running DS simulations using the previously selected 3D training image (TI-9 extruded) and by applying 100 combinations for the values of affinity in X and Y directions. For each combination of rotation, affinity in X and Y-directions, DS simulation is run for 5 realizations. We summarize the results of this sensitivity analysis in Figure 9 by presenting the average RMSE values of connectivity in X and Y directions. In Figure 9, the horizontal and vertical axes represent the affinity parameters in X and Y directions, and the color represents the RMSE values. The figures in the first row represent RMSE values corresponding to the connectivity of facies 0 in the X and Y directions respectively. The RMSE values are computed between the average probability of connection obtained from 5 realizations of DS simulations, and the probability of connection observed in the analogue. Similar results for facies 1 are presented in the second row. The figures in the third row show

the average values of RMSE presented in Figures 9a and 9c; and 9b and 9d respectively. Note that the magnitude of the error values shown in color bars is different in each figure.

The objective here is to find a combination of Affinity along X axis and Affinity along Y axis for which the RMSE values are minimal in all the Figures 9a to 9f. By comparing Figures 9e and 9f, it can be concluded that the optimal parameters correspond to a value of affinity 0.3 in X and 0.4 in Y, evident from Figure 9g. The outputs from 2 realizations of DS simulation, using the selected combination of affinity values, are presented in Figure 10. The figure has been vertically exaggerated by a factor 5 for better visualization.

The selected transformation parameters are further validated against the analogue. To this end, the connectivity functions of an unconditional model are computed on 2D horizontal planes. Such 2D horizontal connectivity properties can be compared to the 2D analogue. There are 60 layers in the model resulting in as many connectivity functions in Figure 11. The probability of connection of facies 0 and 1 in X and Y-directions in the analogue (in blue and cyan dots respectively) are also shown. It is observed that the connectivity functions of the analogue are better reproduced in the X direction than in the Y direction. In the Y direction, the probability of connection obtained from the analogue deviates after the lag-distance of 5. We attribute this discrepancy to the non-stationarity of the analogue. In the next section we will see that it can be reduced by providing a locally-varying orientation map in the DS simulation.

3.5 Conditional simulation and results

In this section, we present the details of conditional DS simulation performed over a sub-region of the Maules-Creek catchment. The parameters determined in the sensitivity analysis are used. These parameters are listed in Table 2. The simulation domain is 4.24 km by 6.32 km and divided into 212×316×94 grid nodes, each cell measuring 20×20×1 m. The

same training image as in the sensitivity analysis is used (Figure 7b). Additional information, such as borehole logs and orientation of the existing stream networks in the study area, are used to locally constrain the model simulation. There are 29 boreholes logs with 1540 facies data within the simulation domain. The global orientation of the existing streams in the Maules-Creek catchment is estimated by manually selecting 10 points along the streams and determining the channel direction. Figure 12a shows the study area with the stream network. An inverse-square-distance interpolation is applied based on the 10 digitized points, resulting in the angle map shown in Figure 12c (identical angles along the vertical).

One realization of the resulting model is shown in Figure 13. The figure has been vertically exaggerated by a factor 10 for better visualization. The simulation output is validated against four criteria: (i) comparing the proportion of facies in the simulation output with that calculated from borehole data; (ii) visual inspection; (iii) comparing connectivity measures of results from DS simulations and those from the analogue; (iv) comparing the flow and transport features with a statistically equivalent Gaussian random field.

The overall proportion of facies is very close to the target of 50% in all realizations. For visual inspection of the results, 2D images showing the meandering structures of six layers (layer 1, 20, 40, 60, 80, 94) are presented in Figure 14. In the iamge, black and grey pixels corresponds to clay and sand respectively. In all the layers sand channels are oriented towards the north-east, and some channels are joining from east to west reflecting that the general orientation set as a conditioning in the simulation is honored. The connectivity functions of each of the 94 horizontal layers in a realization are presented in Figure 15 and compared with the connectivity in the analogue (Figure 8). The comparison indicates that the simulated meandering patterns are similar to those present in the analogue. By comparing Figure 11 and Figure 15, it can be seen that although the same transformation parameters are

used in both cases, additional local information provided in the conditioning data have improved the model results.

For further validation, the transport properties of 3D facies models obtained from DS simulations are compared with a statistically equivalent model obtained with a traditional variogram-based approach. For flow and transport computations, a subdomain of size 212×316×30 is considered, each cell having a size of 20m×20m×1m. Sequential indicator simulation (SIS) is performed using variogram parameters obtained from the results of the DS simulations. The variogram parameters in X, Y and Z directions are listed in Table 3 and one realization of SIS is shown in Figure 16. The Groundwater Finite Element code [*Cornaton*, 2007] is used with fixed head boundary conditions prescribed on the sides of the domain (at Y=0 and Y=6320 m) and a concentration of 1 of a conservative tracer is prescribed at the inlet of the system (at Y=0). Table 4 shows the values used for hydraulic conductivity and porosity of both facies, and dispersivity in X, Y and Z directions. The evolution of the outflow concentration over a period of 200 days is recorded. Figure 17 shows the contaminant distribution after 5, 10, 15, and 30 days, in both DS and SIS models. By comparing Figures 17a and 17e, it is observed that in the DS models a significant concentration has already reached the other side of the domain. In the SIS model however, the contaminant had reached a significant concentration only up to 1000 m in the domain. This observation is also valid for other time intervals, where higher concentrations are systematically observed with the DS model.

The breakthrough curves (BTC), integrated on different 2D planes perpendicular to the flow direction at a distance of Y=1000, 2000, 3000, 4000, 5000 and 6000m, are shown in Figure 18. Each curve in shows the evolution of concentration at one node in a particular section. A broken line is drawn in at day 5 for easy comparison of the arrival times. The importance of modeling heterogeneity using conditional DS simulation is clearly visible by

comparing the BTCs obtained from DS and SIS simulations. Again, the SIS model tends to systematically underestimate the contaminant arrival. The solute generally moves faster in DS models than in the SIS model, and the spread of the BTCs also vary between the different sections. While for the initial parts of the domain (Y=1000 m), the BTCs of DS and SIS are mostly overlapping, marked differences are observed further downstream. SIS shows more tailing in the concentration, which is a typical sign of lack of connectivity throughout the domain, because the contaminant gets trapped in low conductivity zones and takes time to be washed away. These observations, together with the analysis of Figure 17, confirm that SIS does not create the preferential flow paths that are typical of channel-type formations, whereas DS is able to create such preferential flow. This is the case despite the fact that the proportion of the conductive facies (50%), as well as the variographic properties, are identical in both models.

An interesting observation is that the spatial features and the connectivity properties can play an important role, strongly influencing the hydraulic properties even when the proportion of the conductive facies is higher than typical percolation thresholds (i.e. about 0.3 for random binary media). To investigate this further, we run our model with similar settings except that the proportion of sand is now 30% in both DS and SIS models (see Figure S3). The variogram parameters used for SIS shown in Table S1. The same phenomenon of systematic faster contaminant arrival is also observed in this case (Figure S4), indicating that the higher connectivity is not only a percolation effect due to the relatively high proportion of channels in the Maules Creek formation, but it is due to the fundamental connectivity properties that are embedded in the training image through the extrusion process.

Conclusion

In this paper we presented a methodology to generate channel-type alluvial aquifer formations based on simple 2D training images that are extruded in the 3rd dimension. The workflow consists of four steps: (i) selection, (ii) extrusion, (iii) inference and (iv) modeling. A 2D training image is selected based on the data statistics and available expert geological knowledge, representing the facies proportions and the subsurface features that a geologist might expect. Transform-invariant distances are then used with these elementary training images to imprint horizontal variability such as meandering properties. In the final modeling stage, borehole data and overall orientation of the field data are used to constrain the multiple realizations of the 3D facies model.

Our approach allows integrating two levels of information regarding orientation. The first level is a deterministic orientation component, which is inferred from the main orientation of the channels observed on the modeled site. Around this main direction, the transform-invariant distances are used to impose a randomized orientation component to generate sinuous channels. The transform-invariant distance can be controlled by continuous parameters (tolerances on angles and transformations). This allows adjusting the type, shape and spatial behavior of the channels.

Our approach shows potential for building 3D geological models of channelized structures bsed on prior knowledge. The method was applied to the Maules Creek valley filling fluvial sequence in Australia, for which we have moderate borehole control and good geological knowledge of the depositional processes active at the time of aggradation. We demonstrated the value of the outlined workflow by contrasting contaminant movement through 3D facies models built using our new approach and a traditional SIS model. We have demonstrated the need to replicate both the facies proportions and the sediment architecture where estimates of residence time and contaminant breakthrough are required.

Acknowledgements

References

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

Andersen, M. S., and R. Acworth (2009), Stream-aquifer interactions in the Maules Creek catchment, Namoi Valley, New South Wales, Australia, *Hydrogeology Journal*, *17*(8), 2005-2021.

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

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

Caers, J., and T. Zhang (2005), Multiple-point geostatistics: A quantitative vehicle for integrating geologic analogs into multiple reservoir models, *AAPG Memoir*(80), 383-394.

Carle, S. F., and G. E. Fogg (1996), Transition probability-based indicator geostatistics, *Mathematical Geology*, *28*(4), 453-476.

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

Comunian, A., P. Renard, and J. Straubhaar (2012), 3D multiple-point statistics simulation using 2D training images, *Computers & Geosciences*, *40*, 49-65.

Comunian, A., S. K. Jha, B. S. Giambastiani, G. Mariethoz, and B. J. Kelly (2014), Training Images from Process-Imitating Methods, *Math. Geosci.*, *46*(2), 241-260, 10.1007/s11004-013-9505-y.

Cornaton, F. J. (2007), Ground Water : A 3-D Ground Water and Surface Water Flow, Mass Transport and Heat Transfer Finite Element Simulator, 398 pp, University of Neuchatel, Switzerland.

de Iaco, S., and S. Maggio (2011), Validation techniques for geological patterns simulations based on variogram and multiple-point statistics, *Math. Geosci.*, *43*(4), 483-500.

Deutsch, C. V., and L. Wang (1996), Hierarchical object-based stochastic modeling of fluvial reservoirs, *Mathematical Geology*, *28*(7), 857-880.

Deutsch, C. V. (2002), *Geostatistical reservoir modeling*, Oxford: Oxford University Press.

Deutsch, C. V., and T. T. Tran (2002), FLUVSIM: A program for object-based stochastic modeling of fluvial depositional systems, *Computers and Geosciences*, *28*(4), 525-535.

Emery, X., and J. Parra (2013), Integration of crosswell seismic data for simulating porosity in a heterogeneous carbonate aquifer, *Journal of Applied Geophysics*, *98*, 254-264.

Feyen, L., and J. Caers (2006), Quantifying geological uncertainty for flow and transport modeling in multi-modal heterogeneous formations, *Advances in Water Resources*, *29*(6), 912-929.

Gates, G. (1980), The hydrogeology of the unconsolidated sediments in the Mooki Valley, *New South Wales. MSc Thesis, The University of New South Wales, Australia*.

Ge, Y., and H. Bai (2011), Multiple-point simulation-based method for extraction of objects with spatial structure from remotely sensed imagery, *International Journal of Remote Sensing*, *32*(8), 2311-2335, 10.1080/01431161003698278.

Giambastiani, B. M. S., A. McCallum, M. Andersen, B. F. J. Kelly, and R. Acworth (2012), Understanding groundwater processes by representing aquifer heterogeneity in the Maules Creek Catchment, Namoi Valley (New South Wales, Australia), *Hydrogeology Journal*, *20*(6), 1027-1044.

Gibling, M. R. (2006), Width and thickness of fluvial channel bodies and valley fills in the 488 geological record: A literature compilation and classification, *Journal of Sedimentary Research*, *76*(5-6), 731-770.

Guardiano, F. B., and R. M. Srivastava (1993), Multivariate geostatistics: beyond bivariate moments, in *Geostatistics Troia'92*, edited, pp. 133-144, Springer.

He, X., T. O. Sonnenborg, F. Jørgensen, and K. H. Jensen (2013), The effect of training image and secondary data integration with multiple-point geostatistics in groundwater modeling, *Hydrol. Earth Syst. Sci. Discuss.*, *10*(9), 11829-11860, 10.5194/hessd-10-11829-2013.

Holden, L., R. Hauge, Ø. Skare, and A. Skorstad (1998), Modeling of fluvial reservoirs with object models, *Mathematical Geology*, *30*(5), 473-496.

Huysmans, M., and A. Dassargues (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.

Jha, S. K., G. Mariethoz, J. P. Evans, and M. F. McCabe (2013), Demonstration of a geostatistical approach to physically consistent downscaling of climate modeling simulations, *Water Resour. Res.*, *49*(1), 245-259, 10.1029/2012WR012602.

Kelly, B., W. Timms, T. Ralph, B. Giambastiani, A. Comunian, A. McCALLUM, M. Andersen, R. Blakers, R. Acworth, and A. Baker (2014), A reassessment of the Lower Namoi Catchment aquifer architecture and hydraulic connectivity with reference to climate drivers, *Australian Journal of Earth Sciences*, *61*(3), 501-511, http://dx.doi.org/510.1080/08120099.08122014.08900647.

Kerrou, J., and P. Renard (2010), A numerical analysis of dimensionality and heterogeneity effects on advective dispersive seawater intrusion processes, *Hydrogeol. J.*, *18*(1), 55-72.

Khodabakhshi, M., and B. Jafarpour (2013), A Bayesian mixture-modeling approach for flow-conditioned multiple-point statistical facies simulation from uncertain training images, *Water Resour. Res.*, *49*(1), 328-342.

Le Coz, M., P. Genthon, and P. M. Adler (2011), Multiple-point statistics for modeling facies heterogeneities in a porous medium: The Komadugu-Yobe alluvium, Lake Chad basin, *Mathematical Geosciences*, *43*(7), 861-878.

Lee, J., and T. Mukerji (2012), The Stanford VI-E reservoir: A synthetic data set for joint seismic-EM time-lapse monitoring algorithms., *25th Annual Report, Stanford Center for Reservoir Forecasting, Stanford University, Stanford, CA.*

Maharaja, A., and A. G. Journel (2005), Hierarchical simulation of multiple-facies reservoirs using multiple-point geostatistics, paper presented at SPE Annual Technical Conference and Exhibition, Society of Petroleum Engineers, Dallas, Texas, 9-12 October.

Mariethoz, G., P. Renard, and J. Straubhaar (2010), The direct sampling method to perform multiple-point geostatistical simulations, *Water Resources Research*, *46*(11), W11536.

Mariethoz, G., and B. F. J. Kelly (2011), Modeling complex geological structures with elementary training images and transform-invariant distances, *Water Resources Research*, *47*(7), W07527.

Maxelon, M., and N. S. Mancktelow (2005), Three-dimensional geometry and tectonostratigraphy of the Pennine zone, Central Alps, Switzerland and Northern Italy, *Earth-Science Reviews*, *71*(3), 171-227.

Maxelon, M., P. Renard, G. Courrioux, M. Brändli, and N. Mancktelow (2009), A workflow to facilitate three-dimensional geometrical modelling of complex poly-deformed geological units, *Computers & Geosciences*, *35*(3), 644-658.

McCallum, A. M., M. S. Andersen, B. Giambastiani, B. F. J. Kelly, and R. Ian Acworth (2012), River–aquifer interactions in a semi-arid environment stressed by groundwater abstraction, *Hydrological Processes*, *27*(7), 1072-1085.

Meerschman, E., G. Pirot, G. Mariethoz, J. Straubhaar, M. V. Meirvenne, and P. Renard (2013), A practical guide to performing multiple-point statistical simulations with the Direct Sampling algorithm, *Computers & Geosciences*, *52*, 307-324.

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

Neuweiler, I., and H. J. Vogel (2007), Upscaling for unsaturated flow for non-Gaussian heterogeneous porous media, *Water Resour. Res.*, *43*(3).

Pardo-Igúzquiza, E., and P. Dowd (2003), CONNEC3D: a computer program for connectivity analysis of 3D random set models, *Computers & geosciences*, *29*(6), 775-785.

Perez, C., G. Mariethoz, and J. Ortiz (2014), Verifying the High-order consistency of training images with data for multiple-point Geostatistics, *Comp. & Geosci.*, 10.1016/j.cageo.2014.06.001.

Raiber, M., P. A. White, C. J. Daughney, C. Tschritter, P. Davidson, and S. E. Bainbridge (2012), Three-dimensional geological modeling and multivariate statistical analysis of water chemistry data to analyse and visualise aquifer structure and groundwater composition in the Wairau Plain, Marlborough District, New Zealand, *Journal of Hydrology*, *436-437*, 13-34.

Schlüter, S., and H. J. Vogel (2011), On the reconstruction of structural and functional properties in random heterogeneous media, *Adv. Water Resour.*, *34*(2), 314-325.

Suzuki, S., and J. Caers (2006), History matching with an uncertain geological scenario.

Wang, G., and L. Huang (2011), 3D geological modeling for mineral resource assessment of the Tongshan Cu deposit, Heilongjiang Province, China, *Geoscience Frontiers*, *3*(4), 483-491.

Western, A., G. Blöschl, and R. Grayson (2001), Toward capturing hydrologically significant connectivity in spatial patterns, *Water Resour. Res.*, *37*(1), 83-97.

Williams, R. M., N. P. Merrick, and J. B. Ross (1987), Natural and induced recharge in the lower Namoi 561 Valley, New South Wales, paper presented at Symposium on Groundwater 562 Recharge, Mandurah.

Wojcik, R., D. McLaughlin, A. Konings, and D. Entekhabi (2009), Conditioning Stochastic Rainfall Replicates on Remote Sensing Data, *IEEE Trans. Geosc. Rem. Sen.*, *47*(8), 2436-2449.

Wu, Q., H. Xu, and X. Zou (2005), An effective method for 3D geological modeling with multi-source data integration, *Computers & Geosciences*, *31*(1), 35-43.

Wycisk, P., T. Hubert, W. Gossel, and C. Neumann (2009), High-resolution 3D spatial modelling of complex geological structures for an environmental risk assessment of abundant mining and industrial megasites, *Computers & Geosciences*, *35*(1), 165-182.

Zhou, H., J. Gomez-Hernandez, and L. Li (2012), A Pattern Search Based Inverse Method, *Water Resour. Res.*, *48*(2), W03505, doi:10.1029/2011WR011195.

Zinn, B., and C. Harvey (2003), When good statistical models of aquifer heterogeneity go bad: A comparison of flow, dispersion, and mass transfer in connected and multivariate Gaussian hydraulic conductivity fields, *Water Resour. Res.*, *39*(3), WR001146.

Figure captions

Figure 1. Demonstration of the object modelling methodology using different transform-invariant distance parameters. R refers to rotation around Z-axis and A refers to affinity in Z-direction. (a) 2D sketch. (b) 3D training image obtained from extruding 2D sketch. (c-f) correspond to single realizations obtained from Direct Sampling (DS) with transformation parameters. The vertical aspect ratio in the 2D training image and 3D models are different to enhance the visualization in the 3D models.

Figure 2. Workflow used in 3D geological facies modeling using DS. It consists of four parts: Selection, Extrusion, Inference and Modeling.

Figure 3. Demonstration of the high level of structural complexity that can be result from our workflow. (a) 2D sketch that is then extruded and used for simulation. (b) One resulting realization based on transform-invariant distances. 2D sketch is taken from [*Lee and Mukerji*, 2012].

Figure 4. (a) Map showing the location of study area. (b) Digital elevation model; and (c) location of borehole data in the Maules Creek catchment.

Figure 5. Possible training images, drawn with an image editing software (Microsoft Paintbrush), showing various types of sand channels in colour. The fraction of colored pixels in each image is 50%. The fraction of facies 1 in the borehole data is 0.503. In the heading, TI indicates training image followed by a reference number. The number of grids in TI are 60×60 with size 20m×1m. Training images have been exaggerated for better visualization.

Figure 6. Comparison of probability of transition of facies clay to clay (00 in a) and sand to sand (11 in b) in all Tis and borehole data. The transition probability of TI-7 and TI-9 are closest to the borehole data.

Figure 7. 3D training images obtained by extruding selected 2D sketches from TI-7 and TI-9. The number of grids in the 3D training images are 60×50×60 with sizes 20m×20m×1m in X, Y and Z-directions respectively. The sketches and training images have been exaggerated for better visualization.

Figure 8. An image of a meandering river system. This is used as an analogue for comparing the connectivity of DS outputs. (Source: http://www.fas.org/irp/imint/docs/rst/Sect17/Sect17_4.html)

Figure 9. Sensitivity analysis of the value of the affinity in X and Y directions. a) RMSE of the connectivity function of facies 0 (RMSE_0) in X-direction; b) RMSE of the connectivity function of facies 0 in Y direction; c) RMSE of the connectivity function of facies 1 (RMSE_1) in X-direction; d) RMSE of the connectivity function of facies 1 in Y-direction; e) average RMSE of the connectivity function of both the facies (Mean_RMSE_01) in X-direction; f) average RMSE of the connectivity function of both the facies in Y-direction; g) average RMSE of the connectivity function of both the facies in both X and Y-directions (Mean_RMSE_01_XY). RMSE refers to root-mean-square-error between the DS output and the analogue. The colour bar shows the RMSE values.

Figure 10. Realizations of unconditional DS simulation using Affinity along X-direction as 0.3, and Affinity along Y-direction is 0.4. The figure has been exaggerated by 5 times in the vertical direction for better visualization. The number of grids in the simulation are $60 \times 100 \times 60$ with sizes $20m \times 20m \times 1m$ in X, Y, and Z-directions respectively.

Figure 11. Comparison of connectivity plots for an unconditional realization of DS simulation and analogue in (a) X and (b) Y-directions. The dots indicate connectivity of facies 0 and 1 in the analogue.

Figure 12. Map showing the existing streams in the Maules-Creek catchment; a) Image of the study area; b) enlarged view with streams delineated in green color; c) Interpolated map used in the conditional simulation. The angles are referenced from east = 0°.

Figure 13. A conditional realization of the Maules-Creek catchment 3D facies model. The figure has been exaggerated by 10 times in the vertical direction for better visualization. In the simulation, borehole data and channel orientation have been used as conditioning data. The number of grids in the simulation are $212 \times 316 \times 94$ with sizes $20m \times 20m \times 1m$ in X, Y, and Z-directions respectively.

Figure 14. 2D images of various layers of a realization of the Maules-Creek catchment 3D facies model obtained from conditional DS simulation.

Figure 15. Connectivity plots of various layers of a conditional realization of the Maules-Creek catchment 3D facies model obtained from conditional DS simulation and analogue in (a) X- and (b) Y-directions. The dots indicate connectivity of facies 0 and 1 in the analogue.

Figure 16. A conditional realization obtained from SIS. The figure has been exaggerated by 10 times in the vertical direction for better visualization. In the simulation, borehole data have been used as conditioning data. The number of grids in the simulation is 212×316×30 with sizes 20m×20m×1m in X, Y, and Z-directions respectively.

Figure 17. The relative concentration in the 3D simulation obtained from DS (top row) and SIS (bottom row) after (a) 5 days; (b) 10 days; (c) 15 days; and (d) 30 days.

Figure 18. Comparison of breakthrough curves at each point at six XZ planes of 3D simulation obtained from DS simulations and SIS at a distance (a) 1,000m; (b) 2,000m; (c) 3,000m; (d) 4,000m; (e) 5,000m; and (f) 6,000m in y direction away from the injection section.

Figure S 1. Demonstration of resulting cross-cutting of channels by applying transformation parameters. (a) A training image with straight channels; (b) and (c) show realizations of unconditional DS simulation with different rotation tolerance.

Figure S 2. Samples of 3D TI obtained from DS simulations (a) to (d) and from Object-based simulations (e) to (h). In both approaches similar parameters are used.

Figure S 3. A realization of unconditional SIS (a) and DS simulations (b). The figures have been exaggerated by 10 times in the vertical direction for better visualization. In the simulation, no conditioning data have been used. The number of grids in the simulation is 212×316×30 with sizes 20m×20m×1m in X, Y, and Z-directions respectively.

Figure S 4. Comparison of breakthrough curves at each point at six XZ planes of 3D simulation obtained from unconditional DS simulations and SIS at a distance (a) 1,000m; (b)

2,000m; (c) 3,000m; (d) 4,000m; (e) 5,000m; and (f) 6,000m in y direction away from the injection section.

Figure A 1. Rotation and affinity transformation applied to Nx. (a) Original neighborhood Nx. (b-e) correspond to resulting transformed patterns. R refers to rotation around the Z-axis and A refers to affinity along the X-direction. Adapted from *Mariethoz and Kelly* [2011].

Table captions

Table 1. Values of parameters used in unconditional DS simulations

Table 2. Values of parameters used in conditional DS simulations

Table 3. Variogram parameters used in SIS.

Table 4. Values of flow and transport parameters used in the Groundwater code.

Table S 1. Variogram parameters used in SIS with 30% sand.

A.1 Description of the DS algorithm

In DS simulation, for each simulation grid node, the training image is sampled to find a representative of the expected pattern conditioned to the values of other neighboring nodes. The algorithm can be applied to both categorical and continuous variables. For categorical variables, the algorithm works as follows: let $Z(\mathbf{x})$ be a spatially correlated variable of interest; $\mathbf{x}$ is a location in the simulation grid, and $\mathbf{y}$ is a location in the training image; both $\mathbf{x}$ and $\mathbf{y}$ are n-dimensional vectors containing the coordinates of a point in the Cartesian coordinate system. The simulation begins at a random location $\mathbf{x}$. A neighborhood is defined around $\mathbf{x}$ with $\mathbf{N}$ being the ensemble of n closest informed voxels, defined as data event $\mathbf{N_x} = [Z(\mathbf{x}+\mathbf{h_1}),...,Z(\mathbf{x}+\mathbf{h_n})]$; where $\mathbf{h_1},...,\mathbf{h_n}$ are the lags between the location $\mathbf{x}$ and its neighbors. The aim of the simulation is to find a voxel with a similar neighborhood $\mathbf{N_y} = [Z(\mathbf{y}+\mathbf{h_1}), ... , Z(\mathbf{y}+\mathbf{h_n})]$ in the training image. This is accomplished by defining a mismatch or distance between the data events $\mathbf{N}_x$ and $\mathbf{N}_y$ as $d[\mathbf{N_x},\mathbf{N_y}]$, which is defined as

$$d\left[\mathbf{N}_x,\ \mathbf{N}_y\right] = \frac{1}{n}\sum_{i=1}^{n}a_i\ \in [0\ ,1], \qquad \text{where } a_i = \begin{bmatrix} 0 & \text{if } Z(x_i) = Z(y_i) \\ 1 & \text{if } Z(x_i) \neq Z(y_i) \end{bmatrix}. \qquad (1)$$

Once d is under an acceptable threshold t, the value of the node in the training image $Z(\mathbf{y})$ is assigned to the location $\mathbf{x}$ in the simulation grid. In this study, the method is applied for binary variables only but Eq. (1) is valid for any number of facies.

A.2 Application of transform-invariant distances

The application of transform-invariant distances allows one to calculate $d\left[T_p\left(\mathbf{N_x}\right), \mathbf{N_y}\right]$, where T_p represents the geometric transformation applied to the neighborhood of each simulation grid using two parameters, rotation and affinity. For the rotation transformation, a random value of the rotation angle is chosen in the range $[a, b]$, and the neighborhood $\mathbf{N_x}$ is rotated by that angle centered on the location $\mathbf{x}$. Similarly for the affinity transformation, the affinity ratio is randomly chosen and used to expand or contract the neighborhood. Figure A1 illustrates a few examples of rotation and affinity transformations applied to a simple neighborhood $\mathbf{N_x}$. The flexibility in the selection of the rotation and affinity parameters, within a defined range, allows to considered in the simulation a wider range of patterns than is available in the training image. The rotation and affinity transformations can be used simultaneously to obtain more diverse patterns. It should be noted here that the transformation parameters are continuous, allowing one to incur gradual changes in the spatial structure obtained in the simulations.

The simulation continues by sampling another random value of the transformation parameter in the range $[a, b]$ and calculating the distance d between a neighborhood $\mathbf{N_y}$ in the training image and a transformed neighborhood $\mathbf{N_x}$. When $d < t$, the transformed pattern $T_p\left(\mathbf{N_x}\right)$ is considered equivalent to $\mathbf{N_y}$ and the corresponding value $Z(\mathbf{y})$ is assigned to $Z(\mathbf{x})$. Besides such constraint on the patterns (textural constraints), local constraints are considered when there are conditioning data available.

A.3 Non-stationary transformation parameters

In the case of known orientation of the structures at discrete locations $\mathbf{x}_i$, the range of rotation angles $[a, b]$ can be obtained by assigning the median value α and a tolerance ε leading to parameters varying between $a = \alpha(\mathbf{x}_i) - \varepsilon(\mathbf{x}_i)$, and $a = \alpha(\mathbf{x}_i) + \varepsilon(\mathbf{x}_i)$. Since the parameters $\alpha(\mathbf{x})$ and $\varepsilon(\mathbf{x})$ can vary spatially, they can be interpolated based on the locations where the orientation is known. In this case we use an inverse square distance interpolation of the angles to obtain the median orientation $\alpha(\mathbf{x})$:

$$\alpha(\mathbf{x}) = \sum_i \frac{w_i(\mathbf{x})\alpha_i}{\sum_j w_j(\mathbf{x})} , \tag{2}$$

where $w_i = \dfrac{1}{\|\mathbf{x} - \mathbf{x}_i\|^2}$.

The tolerance $\varepsilon(\mathbf{x})$ is determined as a function of the distance to the closest point where the orientation is known.

$$\varepsilon(\mathbf{x}) = \frac{\arg\min\left(\|\mathbf{x} - \mathbf{x}_i\|\right)}{S} \varepsilon_{\max} \tag{3}$$

where and S is the domain size.

a) 2D Training image
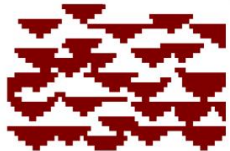
b) N_x
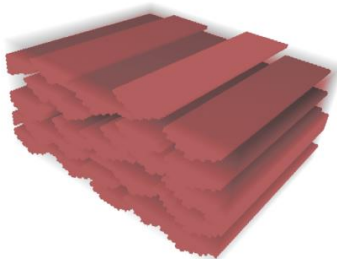
c) $R_{45}A_1(N_x)$
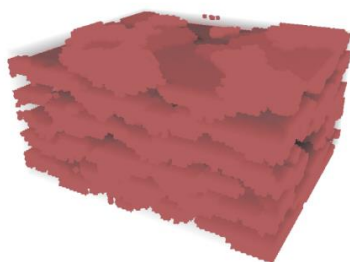
d) $R_{80}A_1(N_x)$

e) $R_0A_2(N_x)$

f) $R_0A_{0.5}(N_x)$

a)

b)

Boundary (3)

Channel (2)

Point bar (1)

Flood plain (0)

a)

Elevation (m)
- -432 - 105
- 106 - 225
- 226 - 354
- 355 - 492
- 493 - 646
- 647 - 818
- 819 - 1,002
- 1,003 - 1,192
- 1,193 - 1,401
- 1,402 - 1,643
- 1,644 - 1,913
- 1,914 - 2,204
- 2,205 - 2,228

WA: Western Australia
NT: Northern Territory
QLD: Queensland
NSW: New South Wales
SA: South Australia
VIC: Victoria

b)

c)

Sand & Gravel
Clay & Silt

TI-1   TI-2   TI-3   TI-4   TI-5

TI-6   TI-7   TI-8   TI-9   TI-10

a) **TI-7**

b) **TI-9**
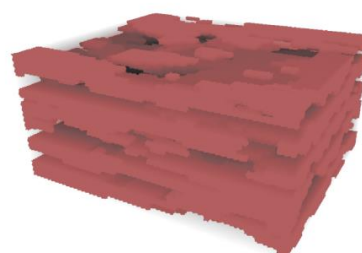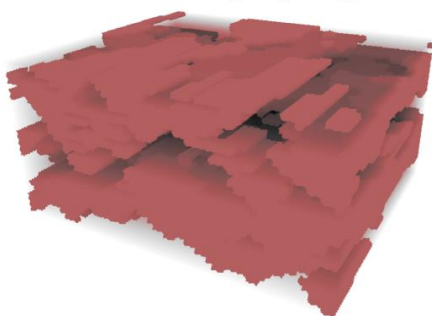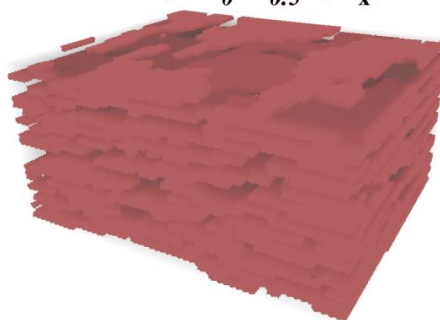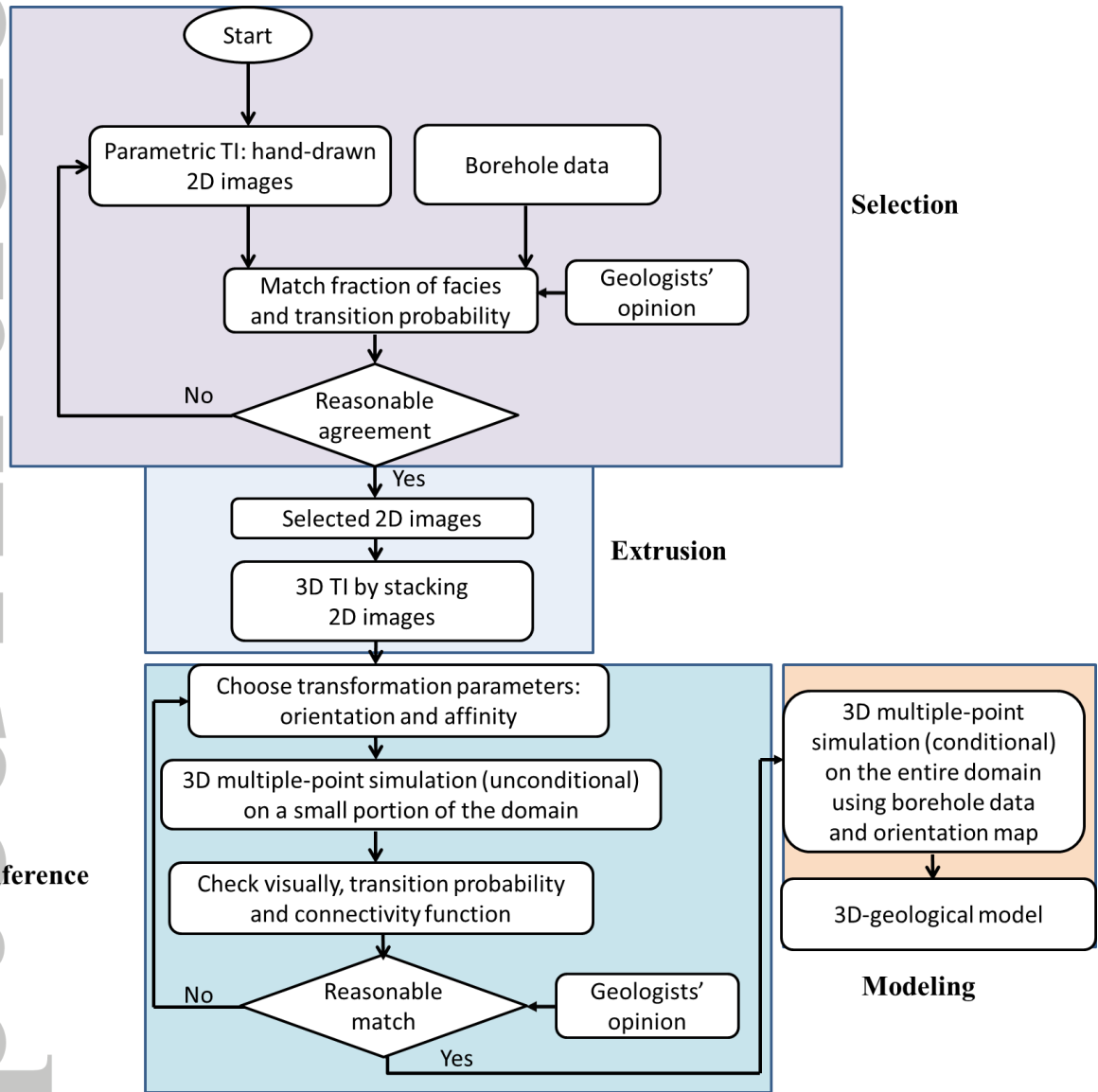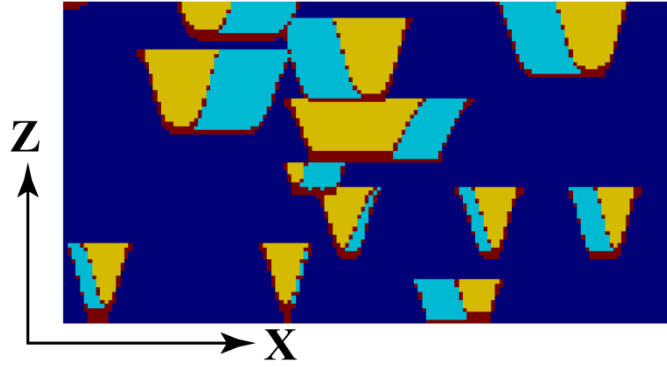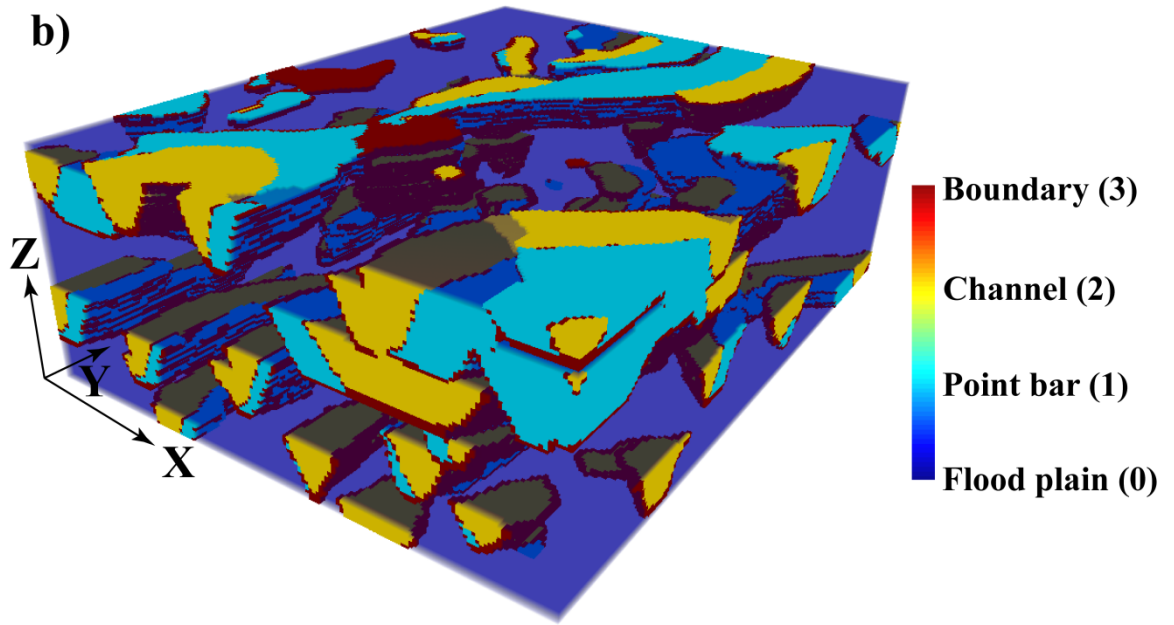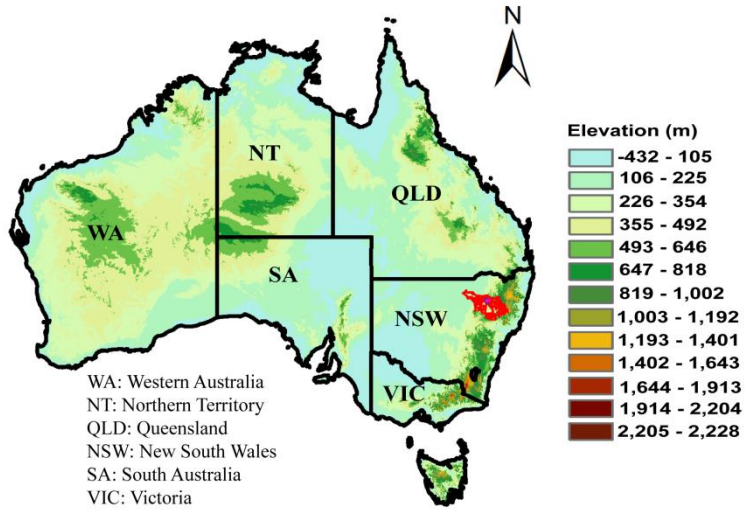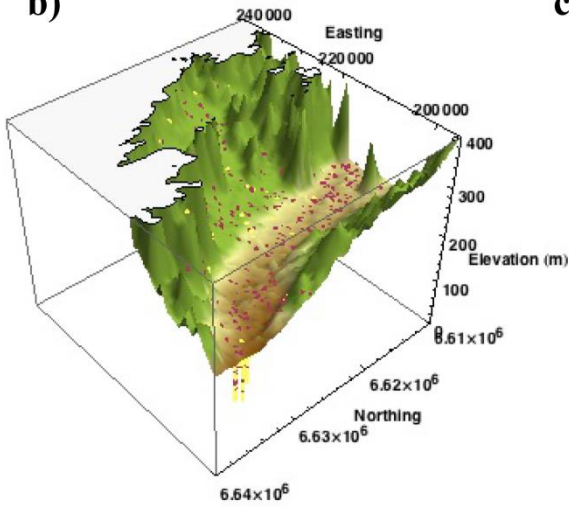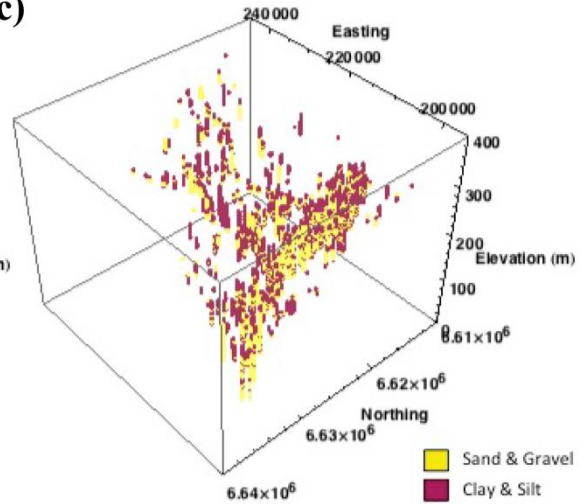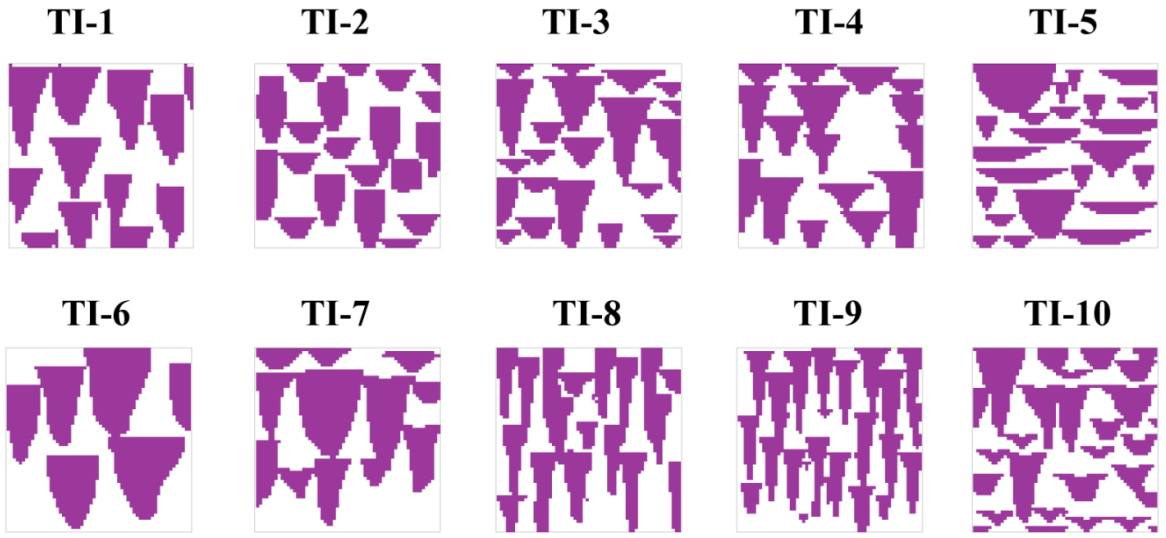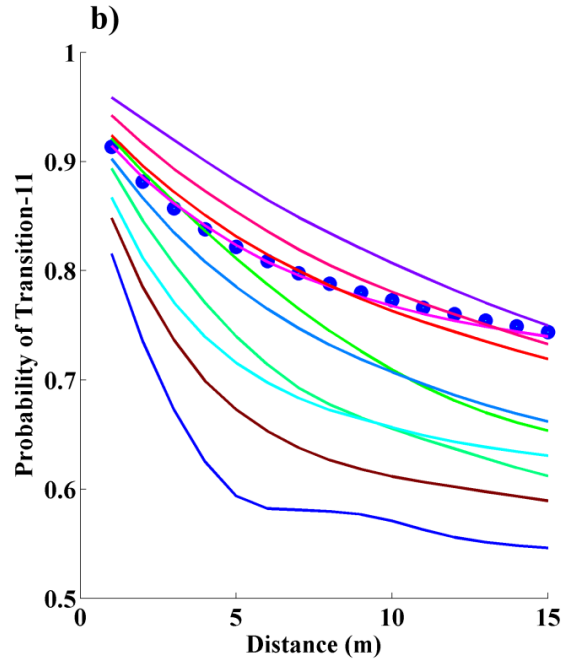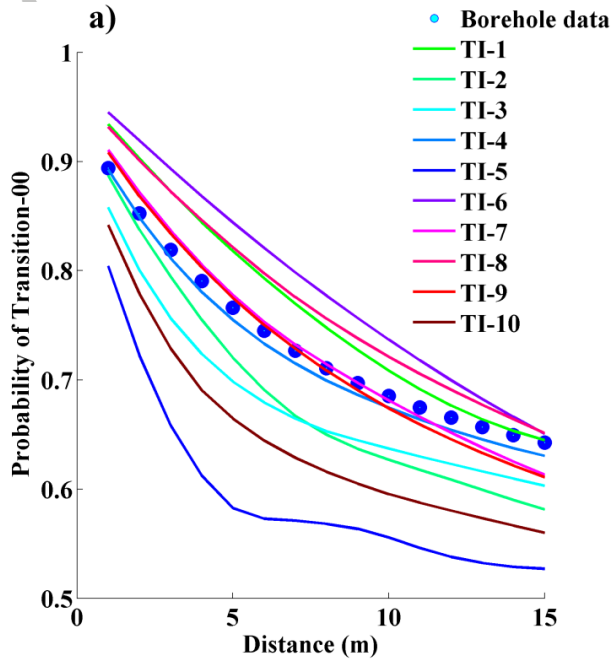
Selected 2D sketches

3D training image

Z

Y

X

RMSE of the connectivity function along X

RMSE of the connectivity function along Y

Realization 1

Y Z X

Realization 2

a)

b)

110⁰

-128⁰

105⁰

c)

Number of grids in Y-direction

Number of grids in X-direction
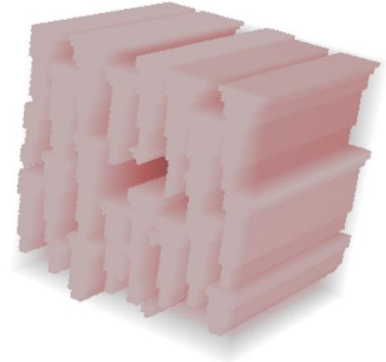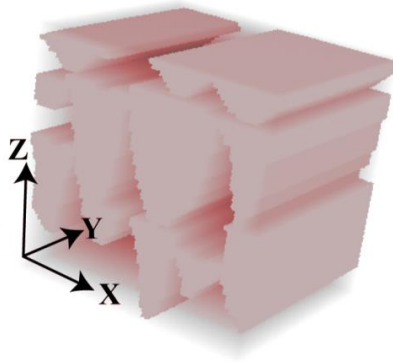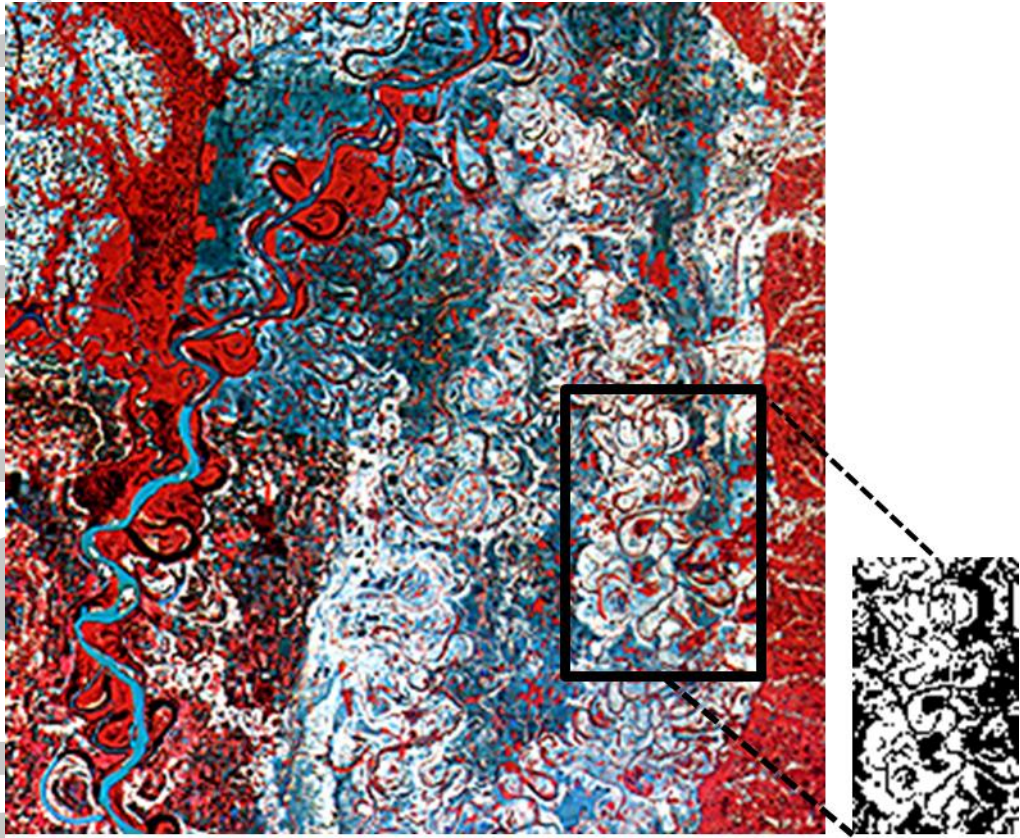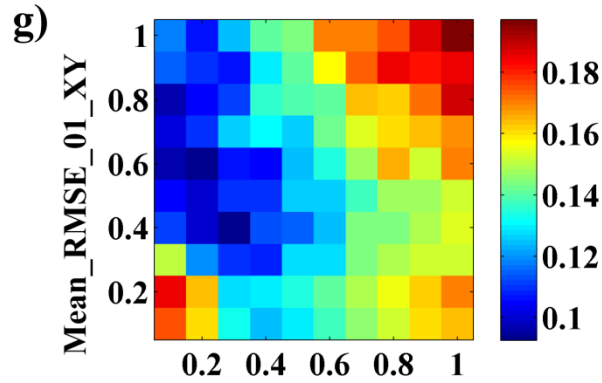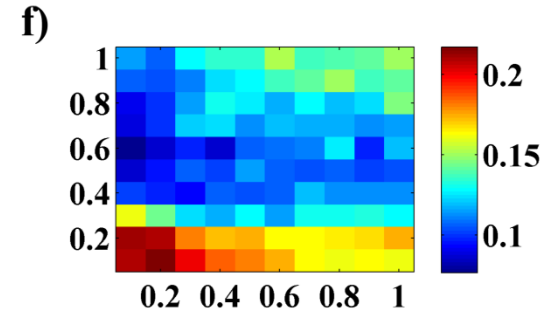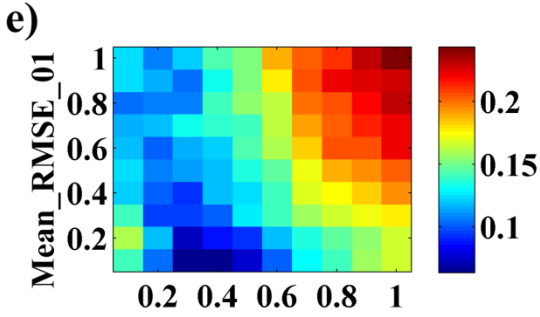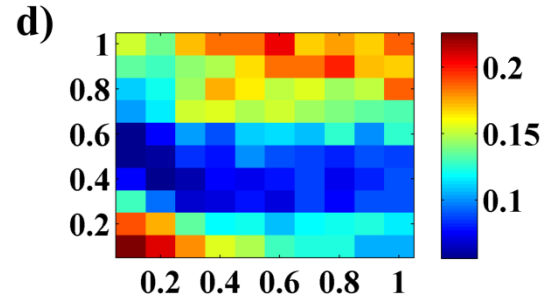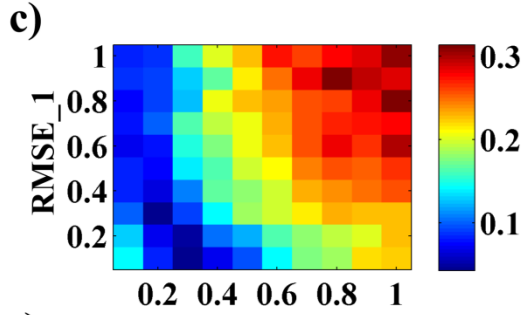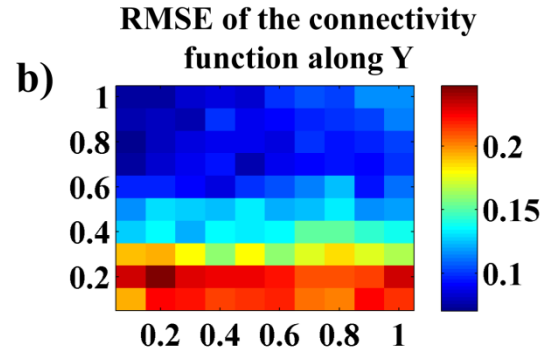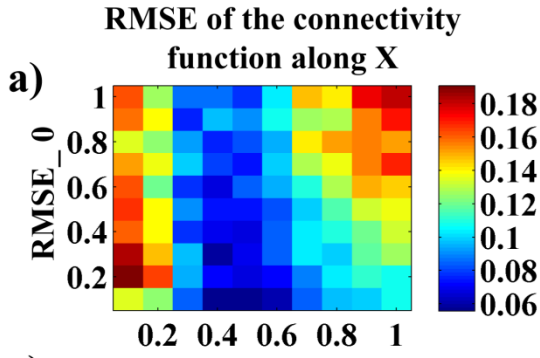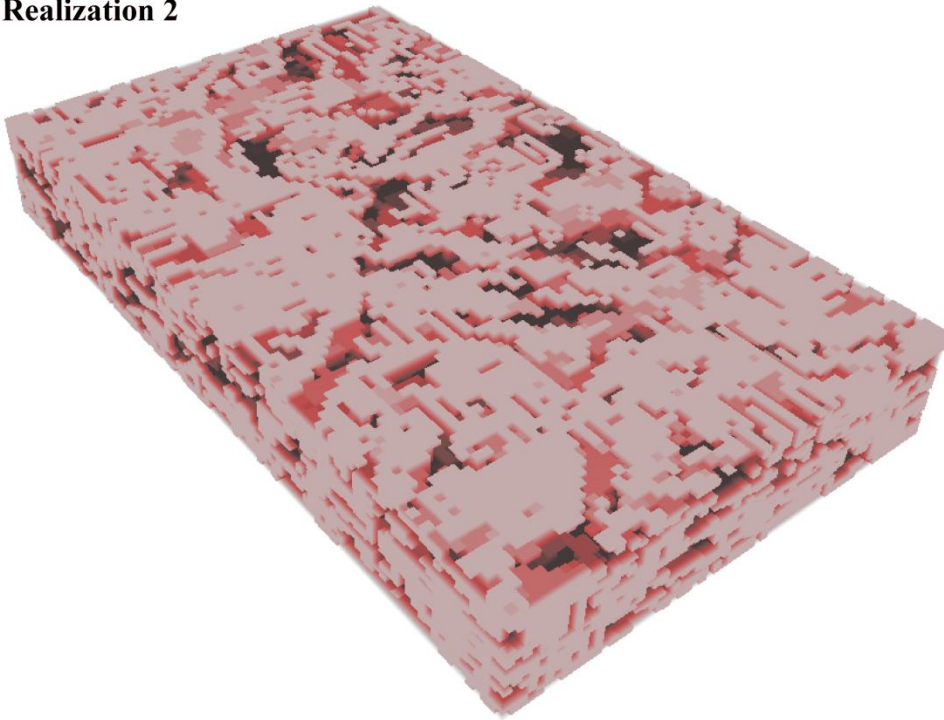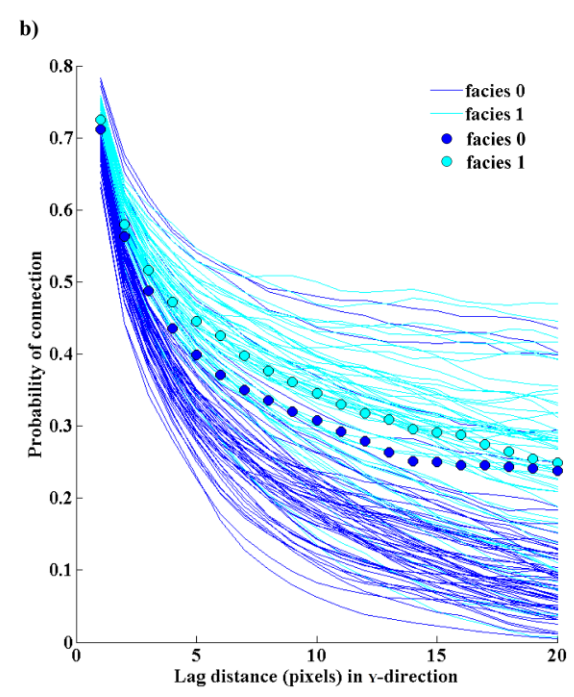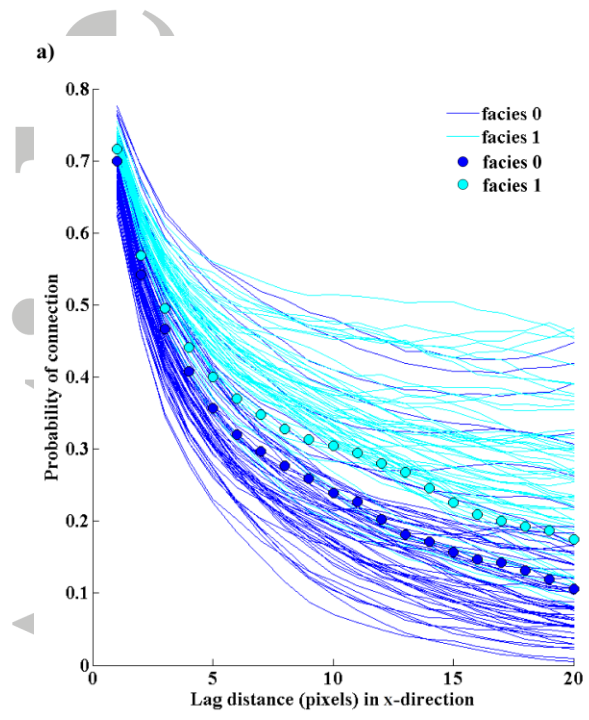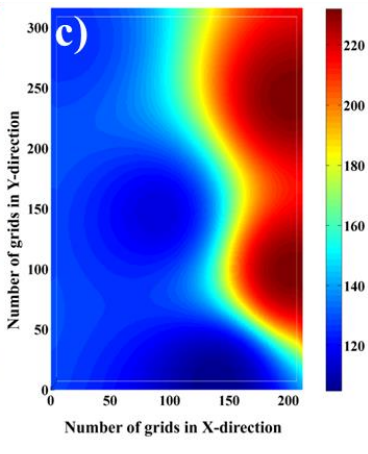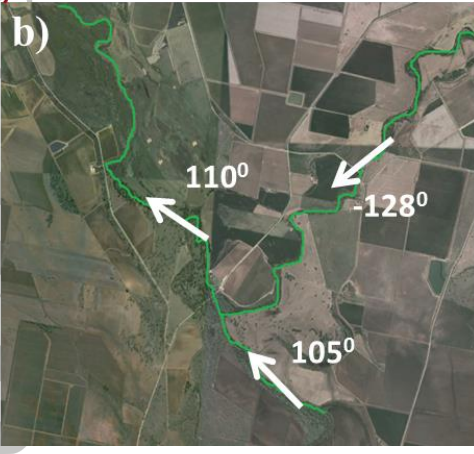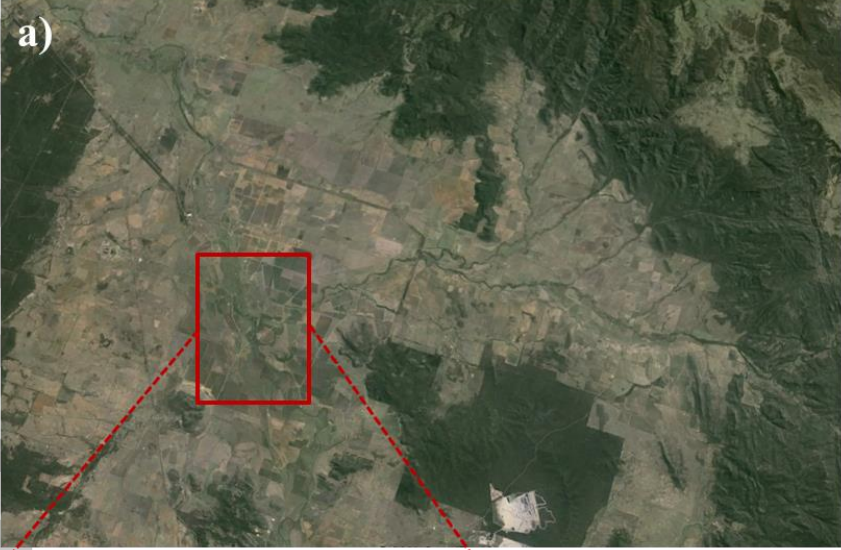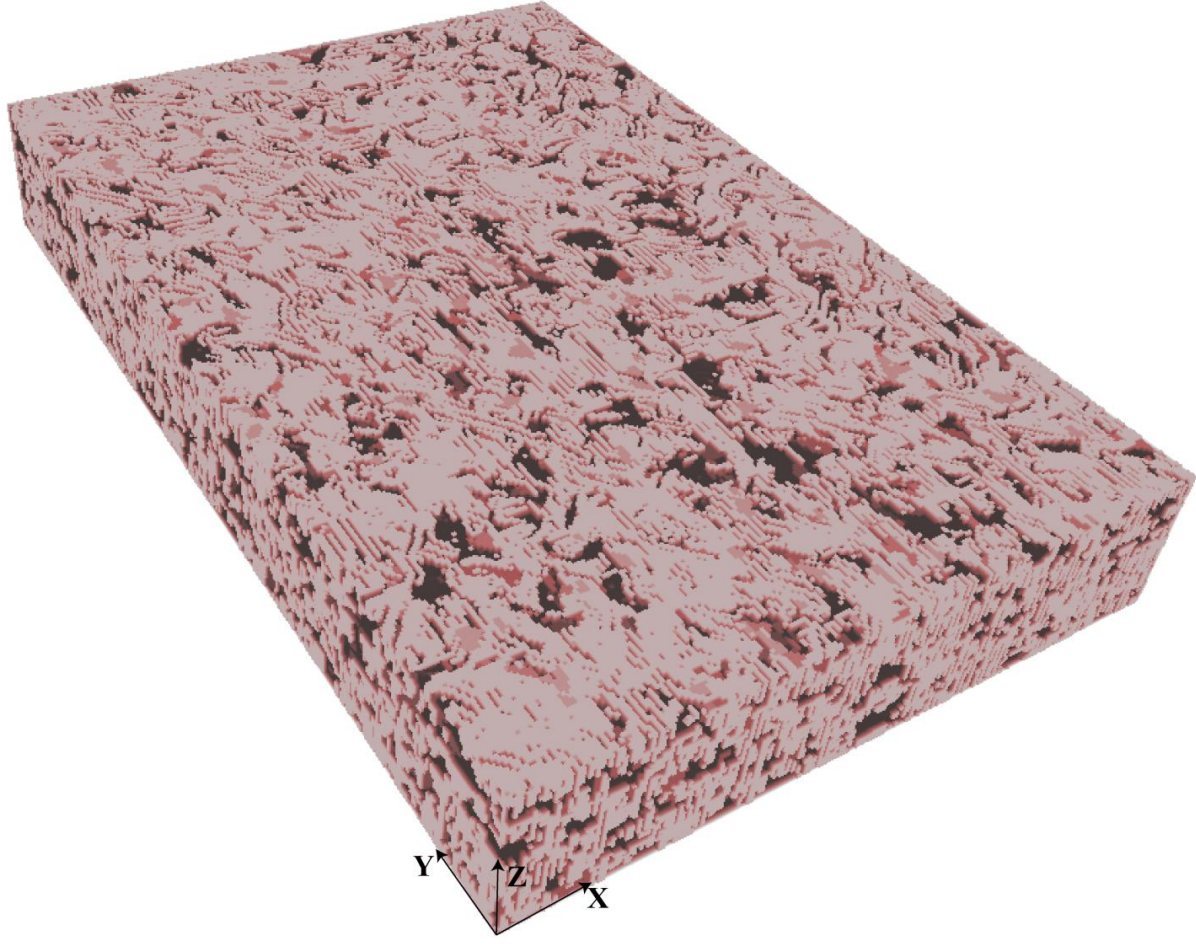
Z = 1
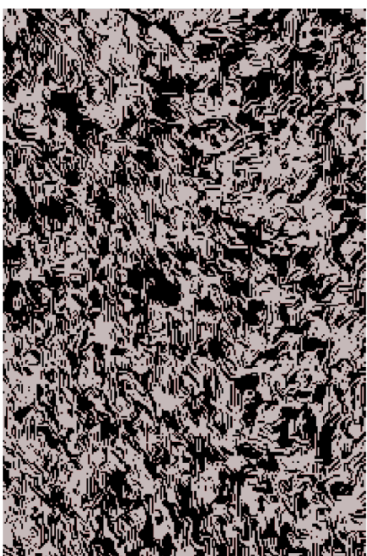
Z = 20

Z = 40

Z = 60
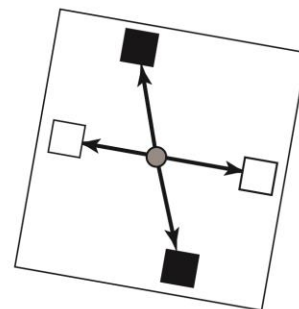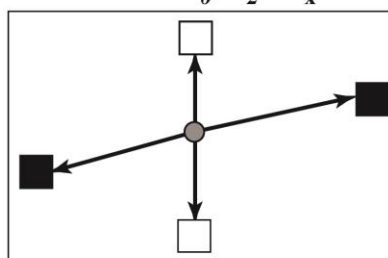
Z = 80

Z = 94
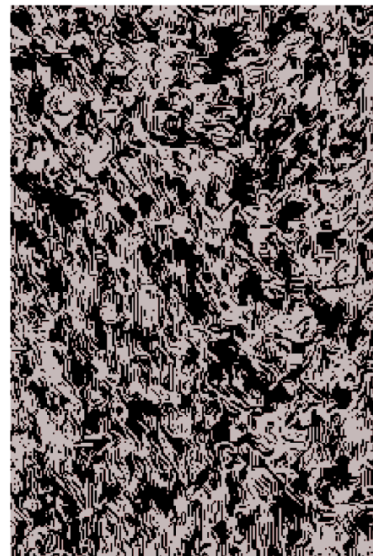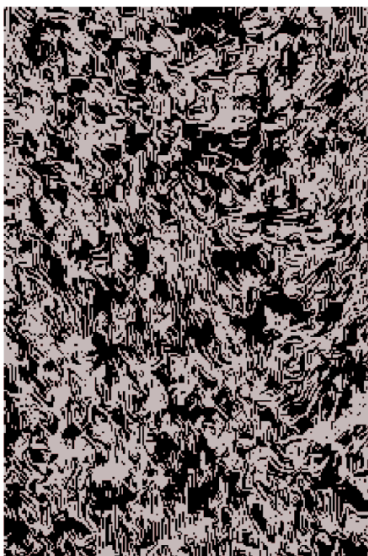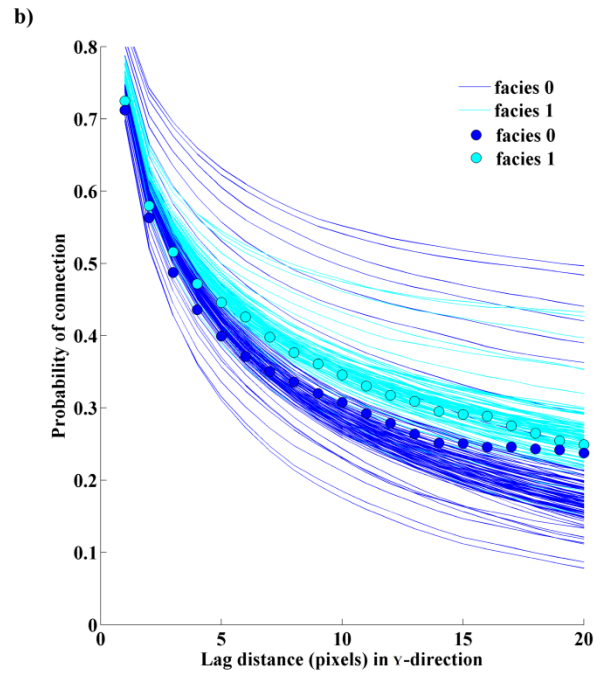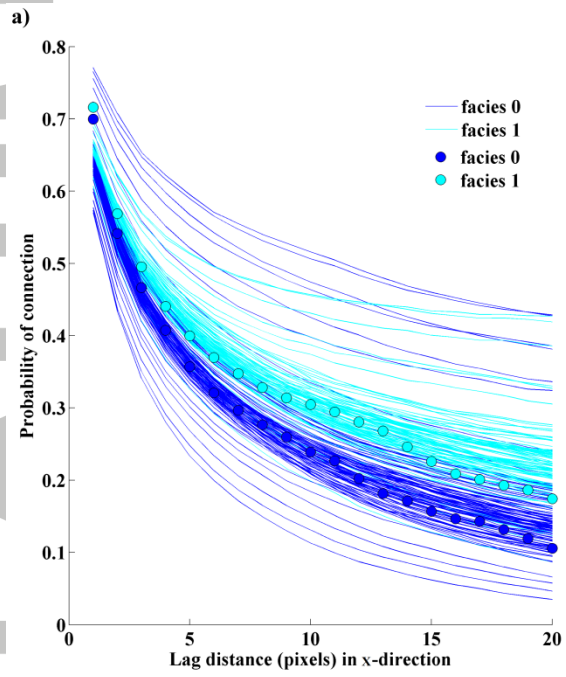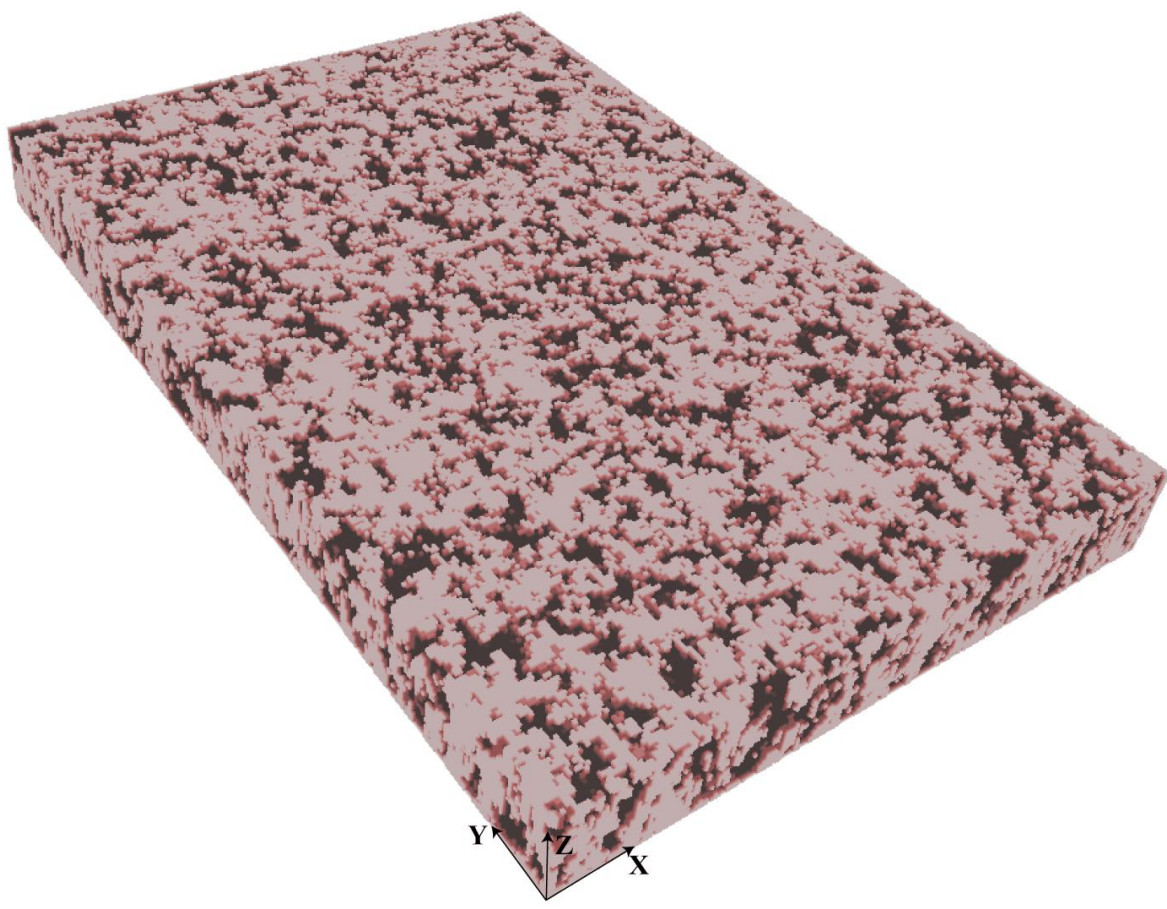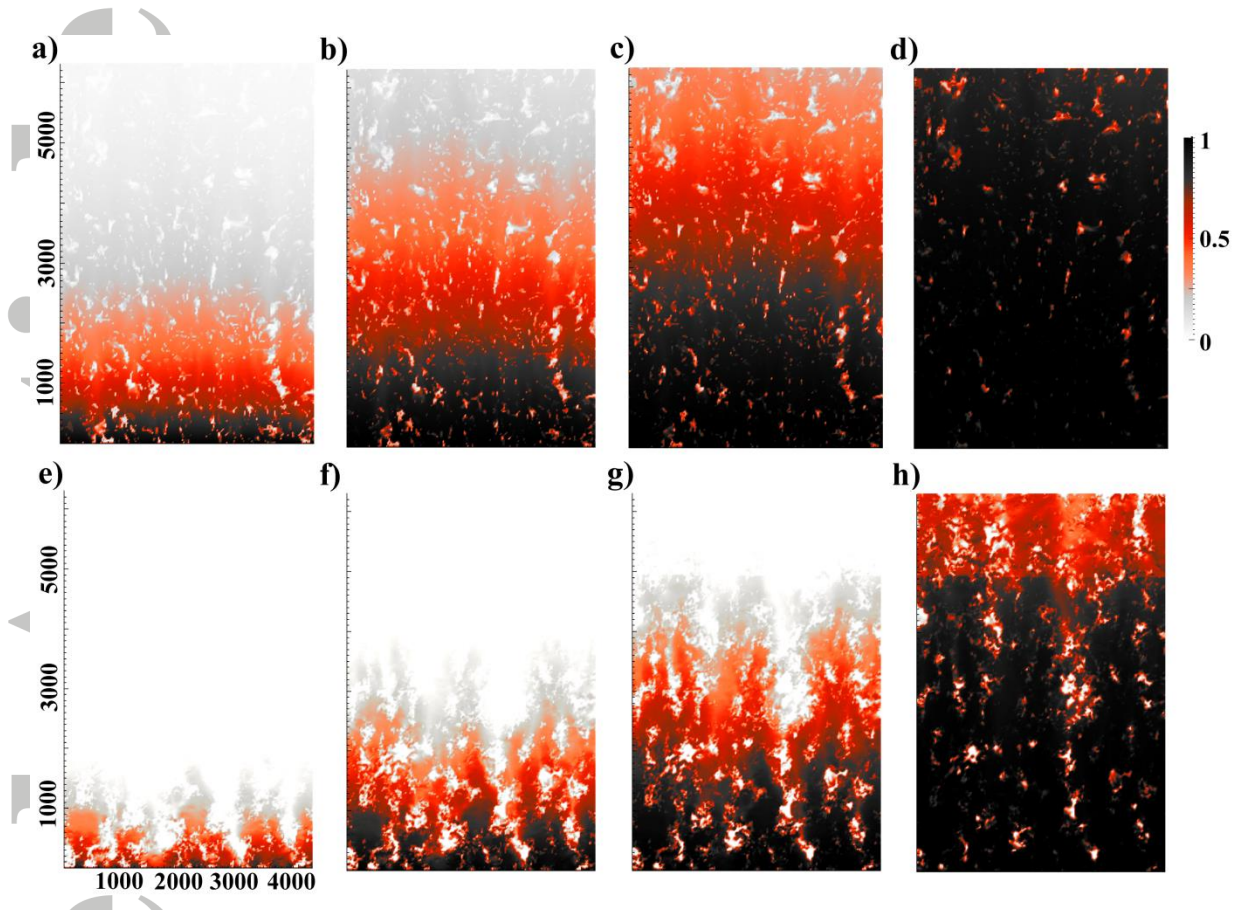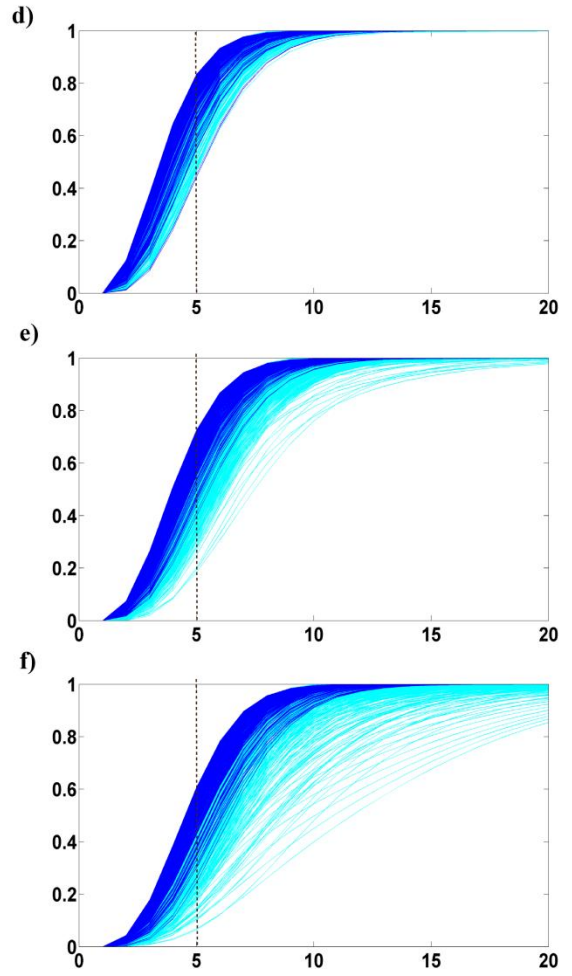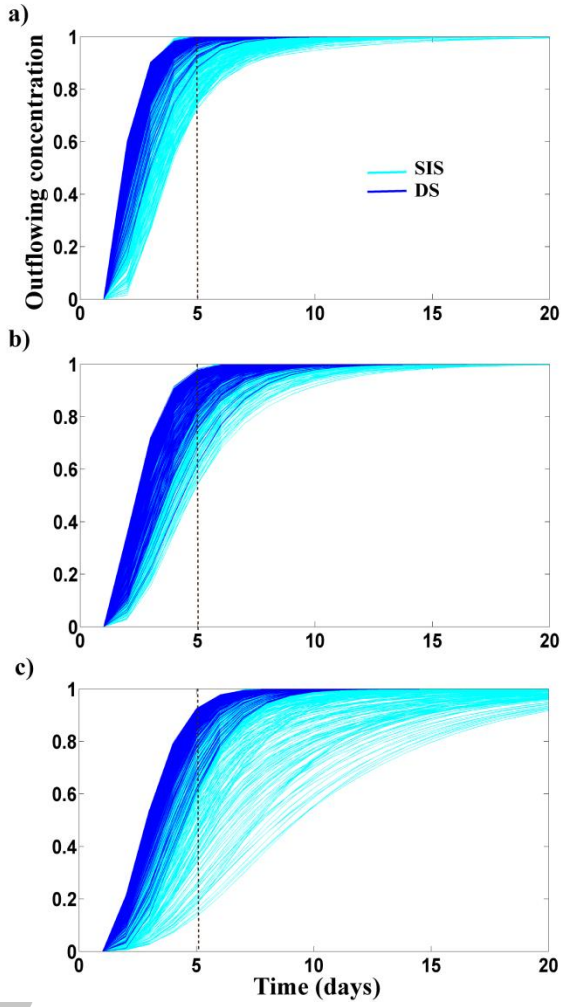
a) N_x
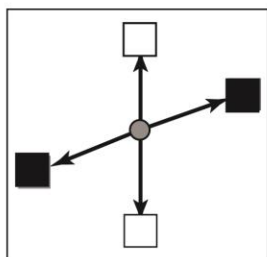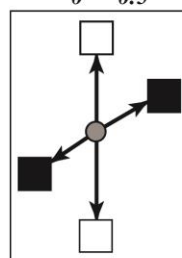
b) $R_{45}A_1(N_x)$
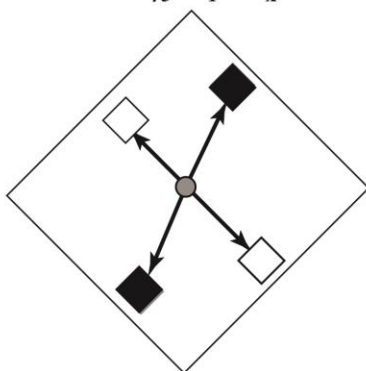
c) $R_{80}A_1(N_x)$

d) $R_0A_2(N_x)$

e) $R_0A_{0.5}(N_x)$

Table 1. Values of parameters used in unconditional DS simulations.

DS parameters	Values
Number of grid nodes	60×100×60 in X, Y, and Z-directions
Grid size	20m × 20m ×1m in X, Y, and Z-directions
Maximum number of neighbouring pixels	30×30×30 in X, Y, and Z-directions
Distance threshold	0.1
Maximum fraction of the TI to scan	0.5
Affinity	0.1 to 1 in X, and Y-directions; 1 in Z-direction, no tolerance in any direction.
Rotation	0° in X, Y, and Z-directions. 120^0 tolerance in Z-direction only.

Table 2. Values of parameters used in conditional DS simulations.

DS parameters	Values
Number of grid nodes	212×316×94 in X, Y, and Z-directions
Grid size	20m ×20m ×1m in X, Y, and Z-directions
Maximum number of neighbouring pixels	30×30×30 in X, Y, and Z-directions
Distance threshold	0.1
Maximum fraction of the TI to scan	0.5
Affinity	0.3 in X, and 0.4 in Y-direction; 1 in Z-direction, no tolerance in any direction.
Rotation	0^0 in X, Y, and Z-directions. 120^0 tolerance in Z-direction only.
Conditioning data	Orientation map and borehole logs

Table 3. Variogram parameters used in SIS.

	X-direction	Y-direction	Z-direction
Nugget	0	0	0
Sill	0.25	0.25	0.25
Range	104m	100m	10m

Table 4. Values of flow and transport parameters used in the Groundwater code.

Parameters	Values
Hydraulic conductivity	$K_{xx} = K_{yy} = K_{zz} = 10^{-7}$ m/s for clay, and 10^{-3} m/s for sand
Drainage porosity	0.01 for clay and 0.25 for sand
Dispersivity	2m in X, 20m in Y, and 1m in Z direction.