A Comparison of MPS Algorithms

Steve Lyster

Multiple-point statistics are high-order spatial moments that contain more information than traditional variograms or covariances. A number of algorithms have been published in recent years that use multiple-point statistics. This paper compares several of them based on a variety of criteria and also compares the results to a variogram-based simulation method.

Introduction

In many cases where geostatistical models are used, the facies or rock types have significant control over the continuous variables of interest such as porosity, permeability, or grade. The spatial structure shown by facies often cannot be captured fully by variograms or covariances. Multiple-point statistics (MPS) may be used in these cases to reproduce structure such as curvilinear connectivity or complex facies relations. MPS is a term that is used to describe both spatial moments of order greater than two and methods that use these high-order moments. The decision of which MPS are important is a modeling choice; most MPS methods use training images (TIs) to derive MPS as data are rarely sampled densely enough to allow for direct inference.

The decision of which MPS algorithm to use is also an important one in the modeling process. Resources for choosing between one algorithm and another are scarce. This paper will discuss the differences between several MPS algorithms and compare the results for two examples. There are a number of published MPS algorithms, and three will be considered here: MPS-GS, SNESIM, and FILTERSIM. These three algorithms all use TIs to infer the complex spatial moments used in simulation.

MPS Algorithms

The MPS-GS algorithm (Lyster and Deutsch, 2008) is an iterative method that uses a Gibbs sampler framework to converge an initial image to the desired target statistics using conditional distributions. The conditional distributions in a Gibbs sampler may be calculated in any way deemed acceptable, including using MPS; in the MPS-GS algorithm the indicators of multiple-point events (MPEs) are used in a kriging-like framework:

$$P^{*}(k)=\sum_{i,\alpha}\lambda_{i,\alpha}^{k}\left[I\left(E_{i}^{\alpha}\right)-P\left(E_{i}^{\alpha}\right)\right]+P(k) \tag{1}$$

where E_i^a is discrete MPH class *a* for MPE *i*. The MPEs are each discrete parts of a template of points, and are easier to infer from a TI than the full template. This is illustrated in Figure 1. The central 24-point template would be difficult to infer from a TI as there are 2^{24} possible combinations (or classes); each individual four-point event has only 2^4=16 possible classes.

The SNESIM algorithm (Strebelle, 2002, Liu, 2006) is the oldest and most well-developed MPS algorithm. SNESIM is a sequential simulation method that uses Bayes' Law to directly infer the conditional probabilities of facies from a TI, using Equation 2. Figure 2 shows an example of Bayes' Law as it is used in the SNESIM algorithm. The probabilities are stored in a search tree so the TI only needs to be scanned once.

$$P^{*}(k)=\frac{P(k\cap D)}{P(D)} \tag{2}$$

The third MPS algorithm considered for the comparisons here is FILTERSIM (Zhang et al, 2006, Wu et al, 2008). The FILTERSIM algorithm uses filters to group similar patterns of facies to reduce the dimension of the statistics. Each pattern is assigned a score for a number of filters and those patterns with similar scores for all filters are considered to be sufficiently similar for grouping. These similar patterns are represented by prototypes; as simulation proceeds, the prototype that best matches the conditioning data is selected and then one pattern represented by this prototype is chosen and patched into place. An example of a number of patterns and their prototype is shown in Figure 3.

MPS-GS has been developed at the CCG and the code is available to industrial sponsors. The SNESIM and FILTERSIM algorithms are freely available in the SGeMS software package. In addition to the MPS algorithms, SISIM (Deutsch and Journel, 1998) will be used for comparison to a more traditional geostatistical method. SISIM uses only the variogram of each facies as the model of spatial structure.

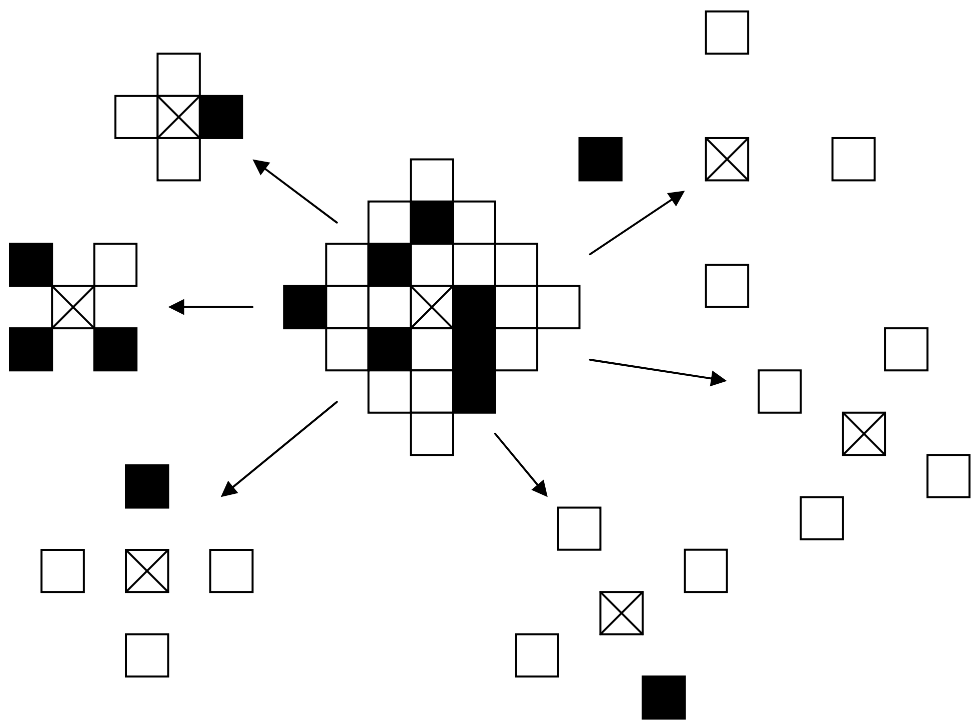
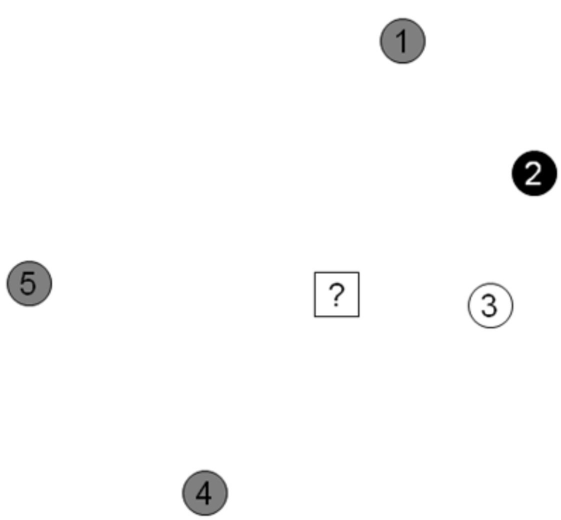


Figure 1: An example of a 24-point MPS template broken into six discrete four-point MPEs.



$$P(white) = \frac{P\begin{pmatrix} 1 = grey \cap 2 = black \cap 3 = white \\ \cap 4 = grey \cap 5 = grey \cap ? = white \end{pmatrix}}{P\begin{pmatrix} 1 = grey \cap 2 = black \cap 3 = white \\ \cap 4 = grey \cap 5 = grey \end{pmatrix}}$$

$$P(grey) = \frac{P\begin{pmatrix} 1 = grey \cap 2 = black \cap 3 = white \\ \cap 4 = grey \cap 5 = grey \cap ? = white \end{pmatrix}}{P\begin{pmatrix} 1 = grey \cap 2 = black \cap 3 = white \\ \cap 4 = grey \cap 5 = grey \end{pmatrix}}$$

$$P(black) = \frac{P\begin{pmatrix} 1 = grey \cap 2 = black \cap 3 = white \\ \cap 4 = grey \cap 5 = grey \cap ? = white \end{pmatrix}}{P\begin{pmatrix} 1 = grey \cap 2 = black \cap 3 = white \\ \cap 4 = grey \cap 5 = grey \end{pmatrix}}$$

Figure 2: An example of Bayes' Law as used in SNESIM.

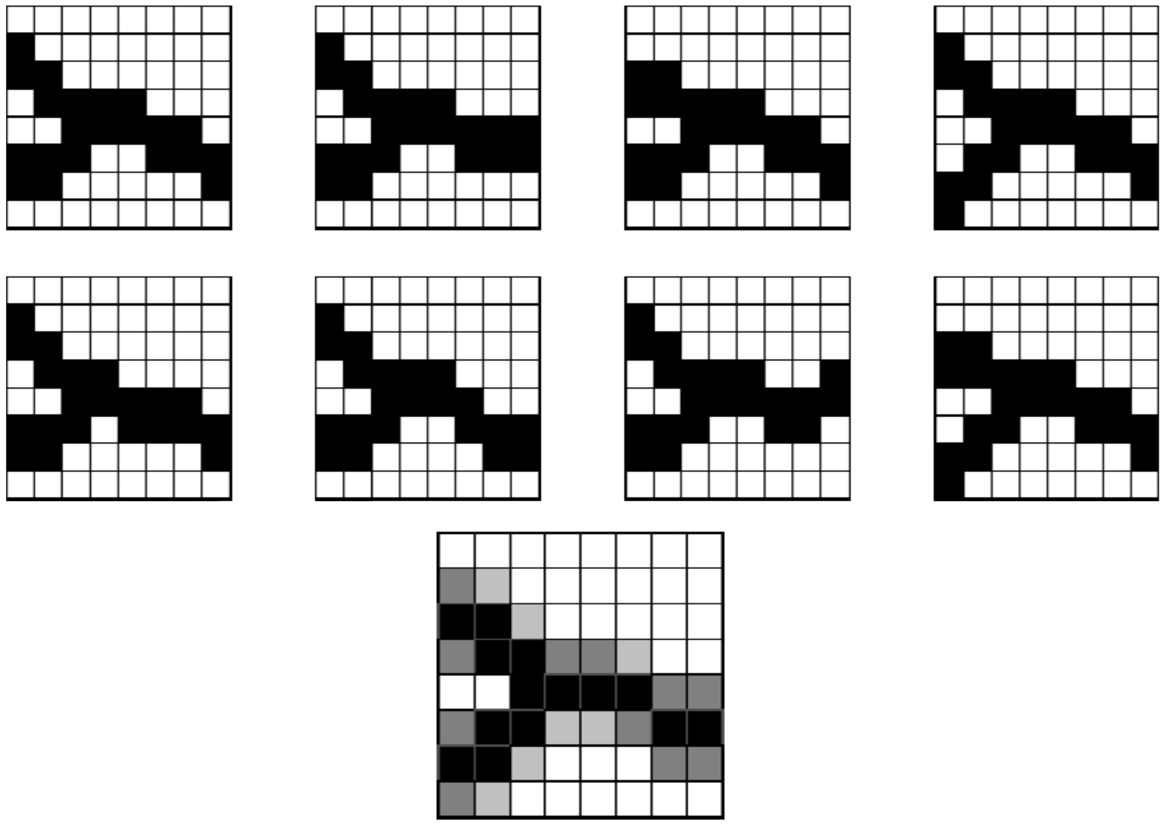


Figure 3: Eight patterns that could be grouped together in FILTERSIM and the prototype representing these patterns.

Comparison Criteria

The criteria that are used when comparing realizations created by using different methods will have an impact on the results of the comparison. For this reason, a variety of criteria must be used. In this paper a wide variety of criteria will be used:

- Visual inspection: this is a good qualitative first-look test for the feasibility of realizations. In MPS it is important that the results "look" like real geology.
- Time required for simulation: an algorithm must be capable of producing results in a reasonable amount of time on a typical desktop computer.
- Variogram reproduction: the variogram is the simplest multivariate spatial statistic and is the easiest to calculate and compare. The MPS algorithms should reproduce the variogram of the TI used.
- Trend reproduction: if a trend exists, this secondary information should be incorporated into the algorithm and properly represented.
- MPH reproduction: the multiple-point histogram is a relatively simple MPS. The algorithms being tested should match the MPH of the TI.
- Runs reproduction: runs, or straight-line connectivity, should be reproduced by the algorithms. This statistic does not capture curvilinear connectivity but is easy to calculate and visualize.

3D Fluvial Channels

Figure 4 shows a TI that will be used for the first comparison. The TI is representative of a simple fluvial channel system with only channel/non-channel facies. The TI is 100x100x50 cells for a total of 500,000.

All of the realizations that are generated will be this size. Twenty realizations were generated using MPS-GS, SNESIM, FILTERSIM, and SISIM. The times required for simulating 20 realizations for each algorithm are shown in Table 1. SISIM is the fastest algorithm, which is to be expected as only lower-order statistics are used for these realizations. Of the MPS algorithms MPS-GS is the fastest but this advantage changes with the number of realizations; MPS-GS calculates and stores the MPS in a file for later use. This calculation is slow but only needs to be done one. In this case with 20 realizations there is a significant speed advantage but for about 5 realizations MPS-GS and SNESIM would take the same amount of time. FILTERSIM is very slow compared to the other algorithms.

Table 1: Time required to simulate 20 realizations of the fluvial channel example.

Algorithm	MPS Calculation	Unconditional
SISIM	-	2:04
MPS-GS	12:04	6:56
SNESIM	-	52:50
FILTERSIM	-	230:54

Figures 5 through 8 show the results of unconditional simulation for the four algorithms. Each algorithm has two figures: one that shows an isometric view of one realization and 2D slices from the same realization, and one that shows the variogram and runs reproduction of the realizations compared to the reference TI values. Visually the SISIM realization is by far the worst, with no curvilinear structure or distinct flat tops on the channels. The MPS realizations all show these features, although the long-range connectivity is not always apparent when looking at 2D slices as the continuity is present in 3D.

The MPS algorithms all reproduce the indicator variograms reasonably well considering that the variograms were not directly used in simulation. SISIM performed surprisingly poor by this measure. For the distributions of runs, the MPS realizations all encompass the TI reference within the 90 percent range of uncertainty for the realizations while the SISIM realizations show far too much short-range connectivity and not enough long-range.

Overall, the MPS results for this example are too close to determine any significant difference between the algorithms.

Stanford V Data Set

The Stanford V data set (Mao and Journel, 1999) is a synthetic reservoir model that is useful for research purposes. This data set includes a TI and hard conditioning data; these two parts will be used to test the algorithms in this paper. The conditioning data also display a clear vertical trend that will be used to examine this aspect of the algorithms. Figure 9 shows an isometric view and 2D slices of the Stanford V TI. This TI has three facies and is 100x130x30 cells for a total of 390,000. The addition of crevasse splays as a third facies significantly increases the complexity of MPS as the order of MPS is K^N is where K is the number of facies and N is the number of points.

Table 2 shows the simulation time for three algorithms in four cases: unconditional, using hard conditioning data, using the vertical trend model, and using both hard data and the trend. The number of realizations was reduced to 10 to reduce the amount of time required. FILTERSIM was not used for this comparison as the program is too slow to be practical. The difference between MPS-GS and SNESIM is again quite pronounced for 10 realizations but the total time (including MPS calculation) would be about the same for 2-3 realizations. SISIM is once again several times faster than the MPS algorithms. The use of hard data has no effect on the time for simulation; using a trend as soft data increases simulation time slightly for the MPS algorithms.

Table 2: Time required to simulate 10 realizations of the Stanford V example.

Algorithm	MPS Calculation	Unconditional	Hard Data	Trend	Hard & Trend
SISIM	-	1:02	1:01	1:03	1:03
MPS-GS	10:45	4:07	4:08	4:52	4:49
SNESIM	-	66:25	66:24	67:01	67:16

Figures 10 through 12 show the results of unconditional simulation. As before, the first figure for each algorithm shows an isometric view and 2D slices from one realization and the second shows variogram and runs reproduction. The results here do differentiate the algorithms. MPS-GS shows too-high covariance (too-low variograms), and the short-range runs distributions are too low while the long-range runs connectivity is too high. SNESIM matches the TI variograms quite well but has too-high short-range runs distributions and too-low long-range runs. SISIM again does surprisingly poorly at variogram reproduction, particularly in the vertical direction. The runs show similar results as before, with the SISIM realizations having far too much short-range connectivity and not enough short-range.

Discussion

Visually the MPS algorithms produce realizations that are significantly better than the variogram-based SISIM program. The time required for simulation favours MPS-GS over SNESIM for many realizations, but this advantage would be lessened with fewer realizations. FILTERSIM performs well from a statistical perspective but the algorithm takes too long to be practical in most cases.

To assess the range of uncertainty in MPS simulation, three sources of uncertainty need to be considered: uncertainty in the TI, uncertainty between realizations, and parameter uncertainty. The parameter uncertainty was not considered in this case and general parameters were used for each algorithm. No TI uncertainty was considered; if this is deemed important than the speed advantage of MPS-GS would be very small or could even become a disadvantage if a large number of TIs were to be used with only 2-3 realizations per TI. If there is significant certainty in the geology (and therefore the TI) then MPS-GS is significantly faster than the other MPS algorithms.

All of the algorithms explicitly reproduce hard conditioning data. MPS-GS has some noticeable discontinuities near conditioning data when the data disagree with the TI; SNESIM shows discontinuities on average as data are more likely to form the edge of geo-bodies than the middle. SISIM has no such problems but no realistic-looking geology is evident in the results.

Integration of soft data is vital for simulation methods to be considered robust. Seismic data or trends are available for many projects and should be incorporated into the models. All of the algorithms tested are capable of reproducing the trend on average; it is this average reproduction that is the goal of “soft” data integration.

The SNESIM algorithm has seen significantly more development and has more features incorporated than other MPS methods. MPS-GS is a more recent development and could be improved in a number of ways and have additional functionality implemented. Overall the MPS methods offer improvements over more-traditional SISIM and, when combined with geologic knowledge in the form of TIs, geologic layering or zone mapping, and fault modeling, can result in realizations that both honour the input statistics and look like real geology.

References

Deutsch, C.V. and Journel, A.G. (1998) *GSLIB: Geostatistical Software Library and User’s Guide, 2nd Ed.* Oxford University Press, New York, 369 p.

Liu, Y. (2006) Using the Snesim Program for Multiple-Point Statistical Simulation. *Computers & Geosciences*, Vol. 32, No. 10, Dec. 2006, pp 1544-1563.

Lyster, S. and Deutsch, C.V. (2008) MPS Simulation in a Gibbs Sampler Algorithm. *8th International Geostatistics Congress*, 10 p.

Mao, S. and Journel, A. (1999) Generation of a reference petrophysical/seismic data set: the Stanford V reservoir. *Stanford Centre for Reservoir Forecasting*, No. 12.

Strebelle, S. (2002) Conditional Simulation of Complex Geological Structures Using Multiple-Point Statistics. *Mathematical Geology*, Vol. 34, No. 1, Jan. 2002, pp 1-21.

Wu, J., Boucher, A., and Zhang, T. (2008) A SGeMS code for pattern simulation of continuous and categorical variables: FILTERSIM. *Computers & Geosciences*, doi: 10.1016/j.cageo.2007.08.008

Zhang, T., Switzer, P., and Journel, A. (2006) Filter-Based Classification of Training Image Patterns for Spatial Simulation. *Mathematical Geology*, Vol. 28, No. 1, Jan. 2006, pp 63-80.

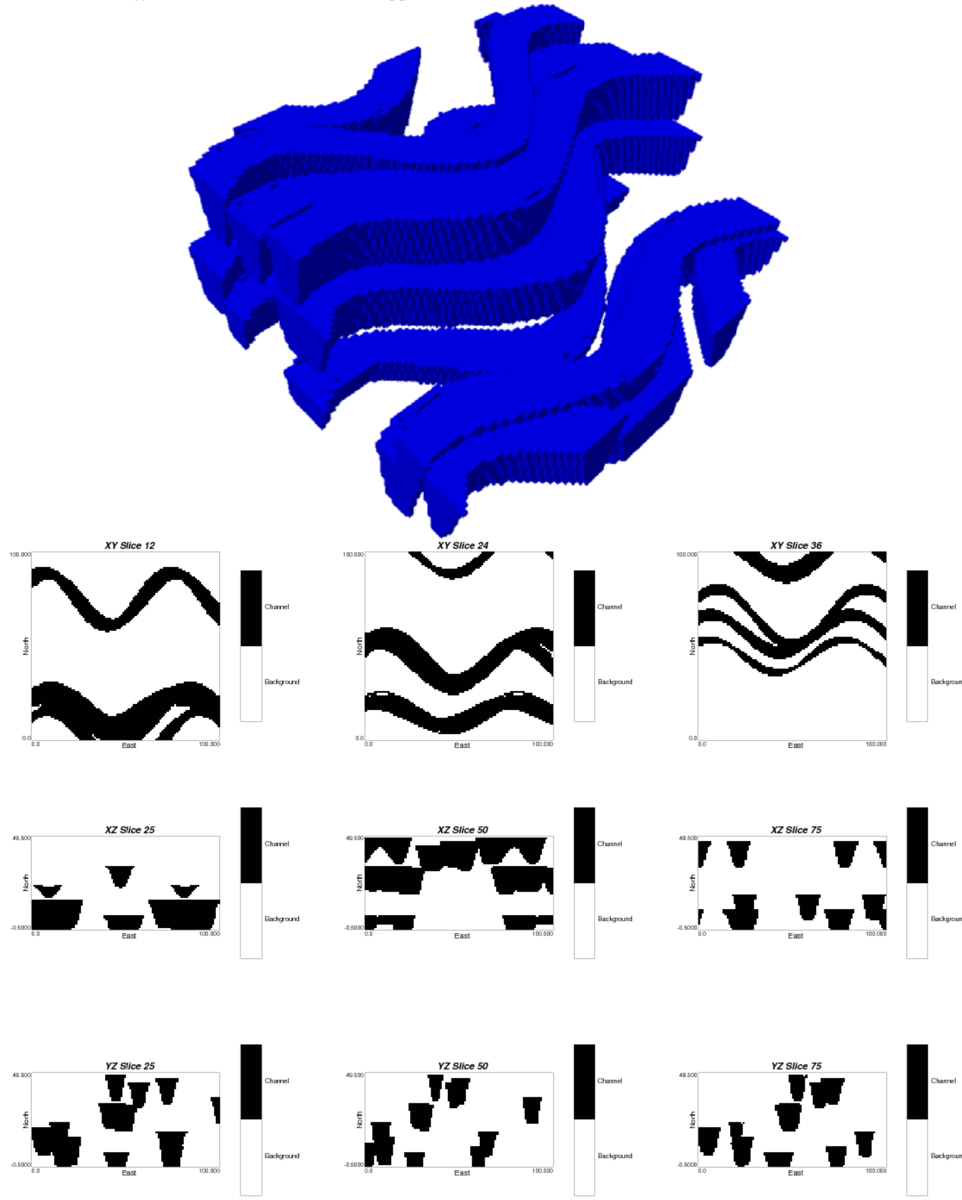


Figure 4: Fluvial channel TI. Top: isometric view; bottom: 2D slices.

Figure 5: One realization created using MPS-GS. Top: isometric view; bottom: 2D slices.

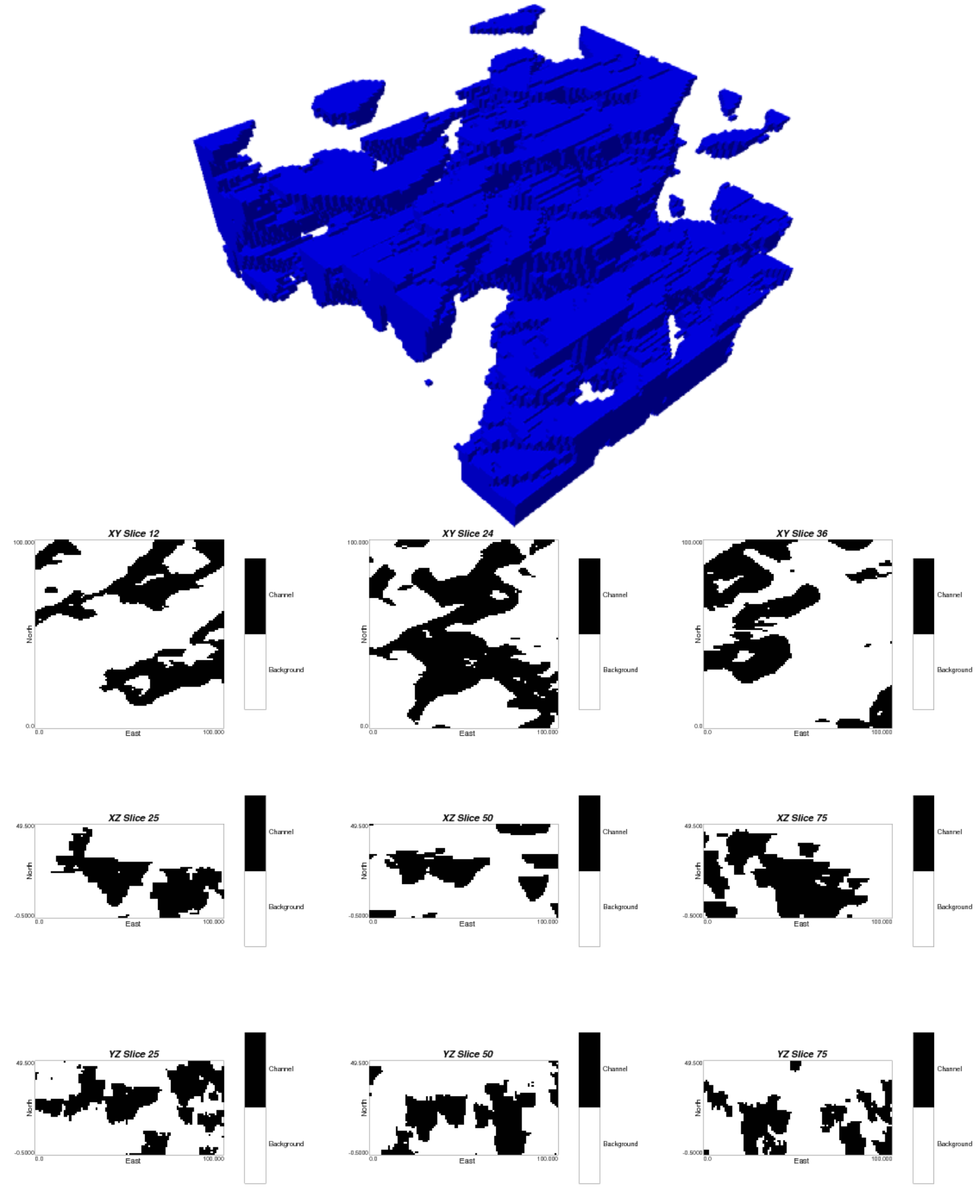


Figure 6: One realization created using SNESIM. Top: isometric view; bottom: 2D slices.

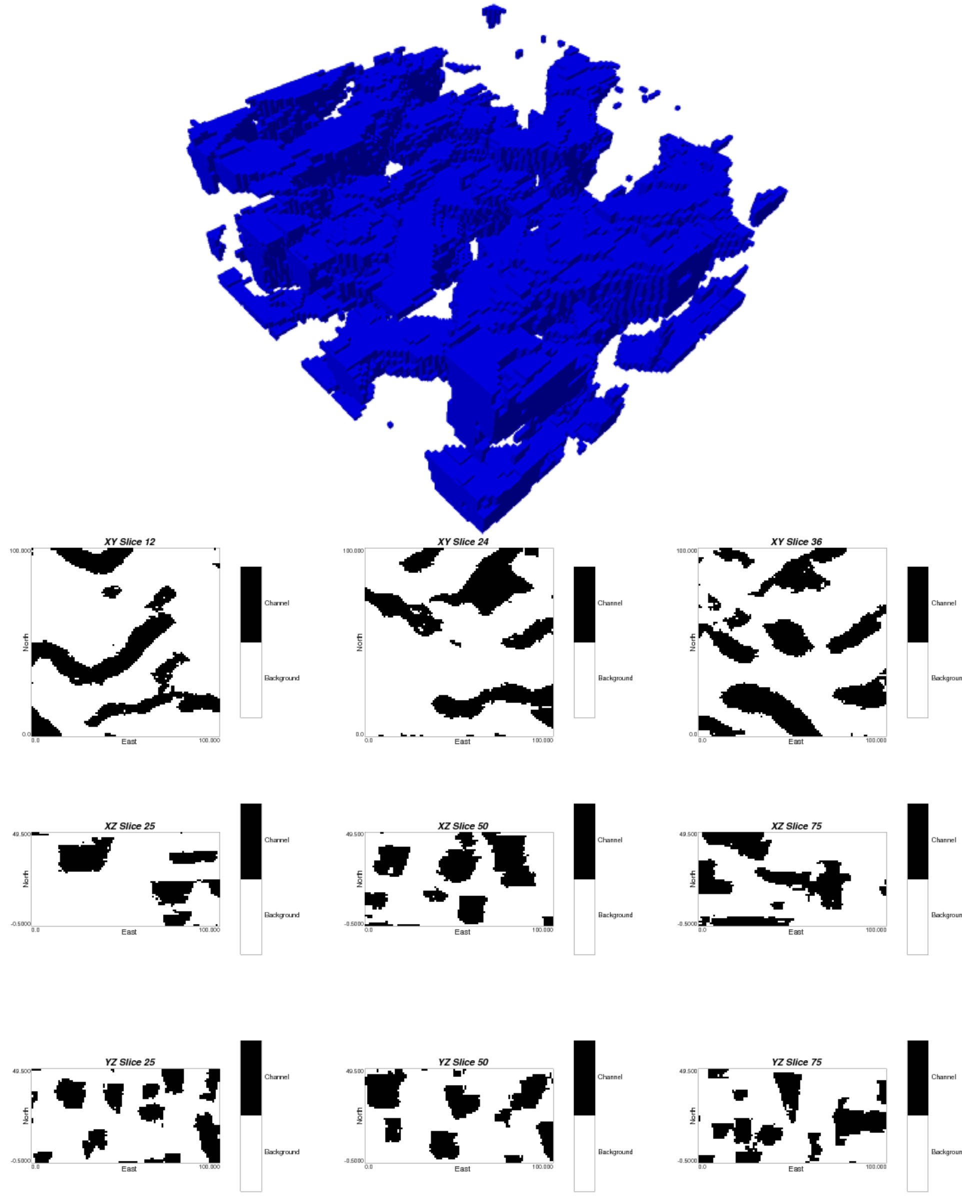


Figure 7: One realization created using FILTERSIM. Top: isometric view; bottom: 2D slices.

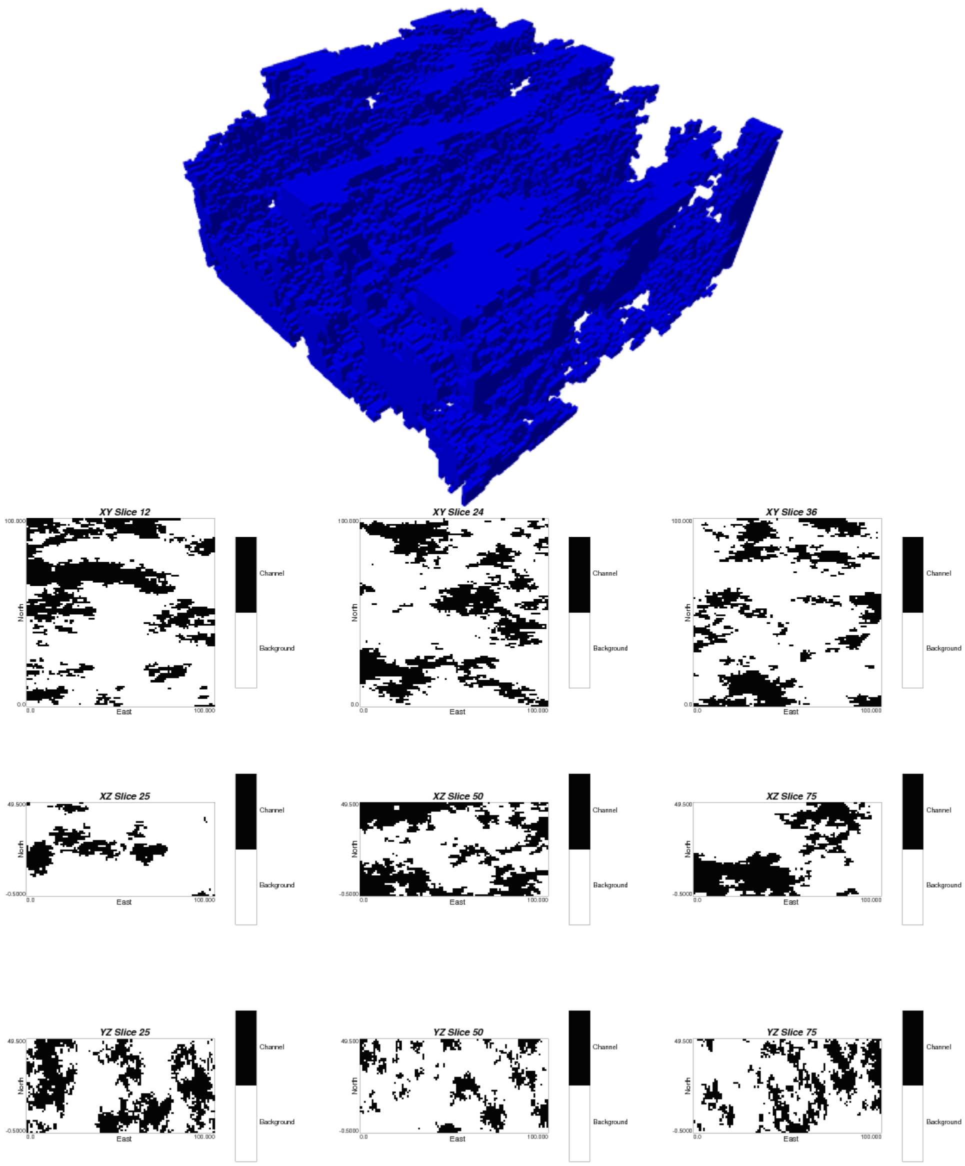


Figure 8: One realization created using SISIM. Top: isometric view; bottom: 2D slices.

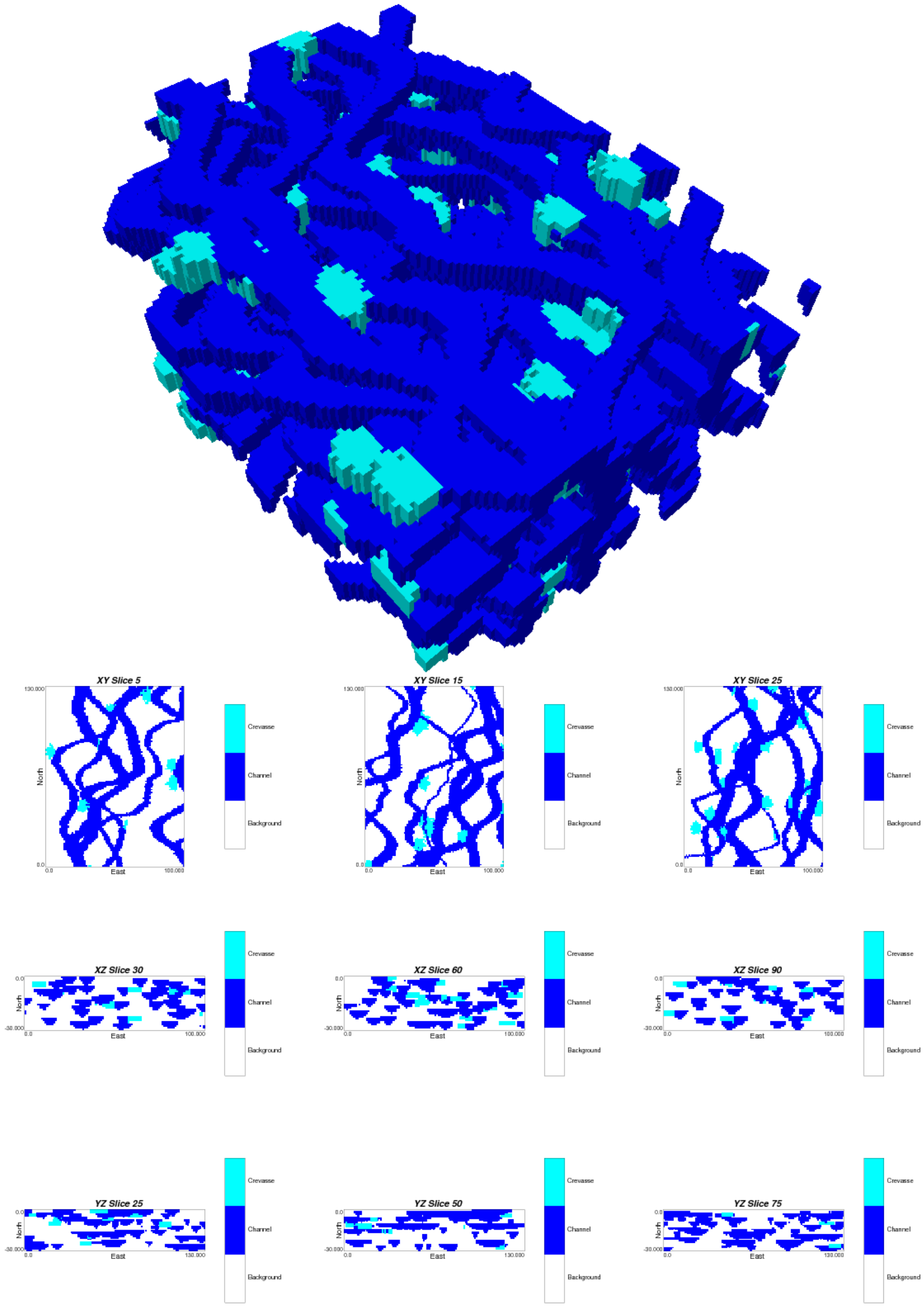


Figure 9: Stanford V TI. Top: isometric view; bottom: 2D slices.

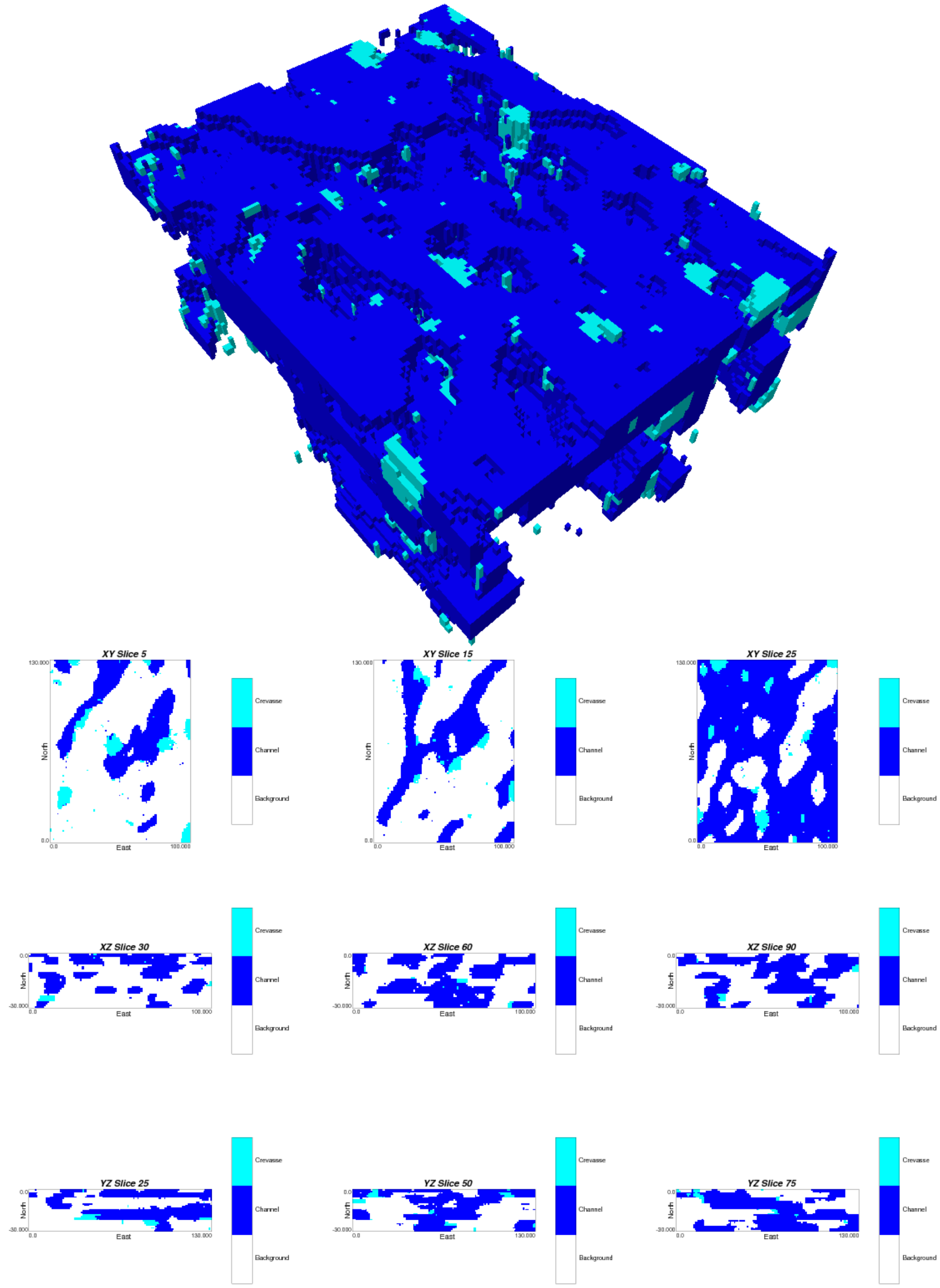


Figure 10: One realization created using MPS-GS. Top: isometric view; bottom: 2D slices.

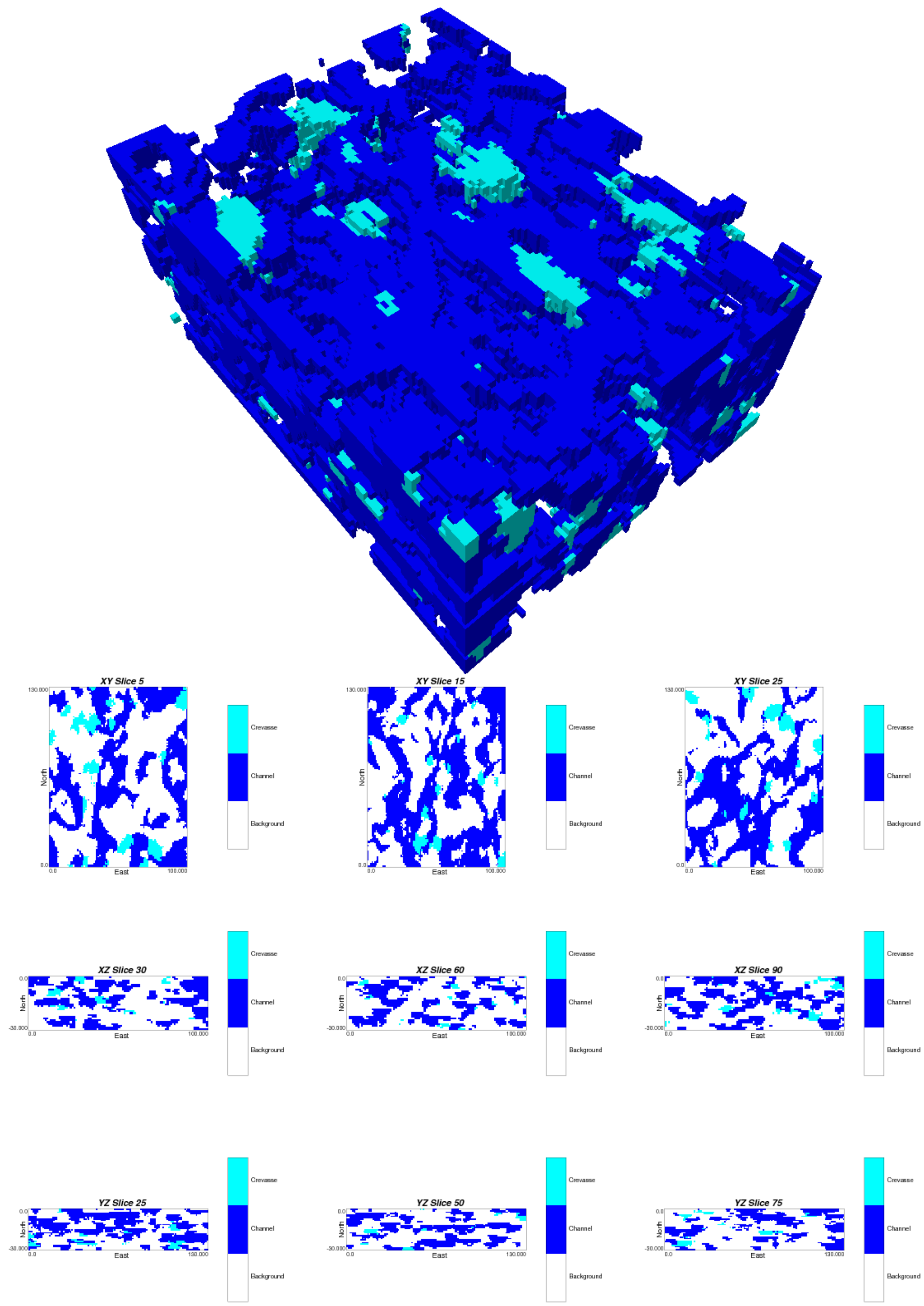


Figure 11: One realization created using SNESIM. Top: isometric view; bottom: 2D slices.

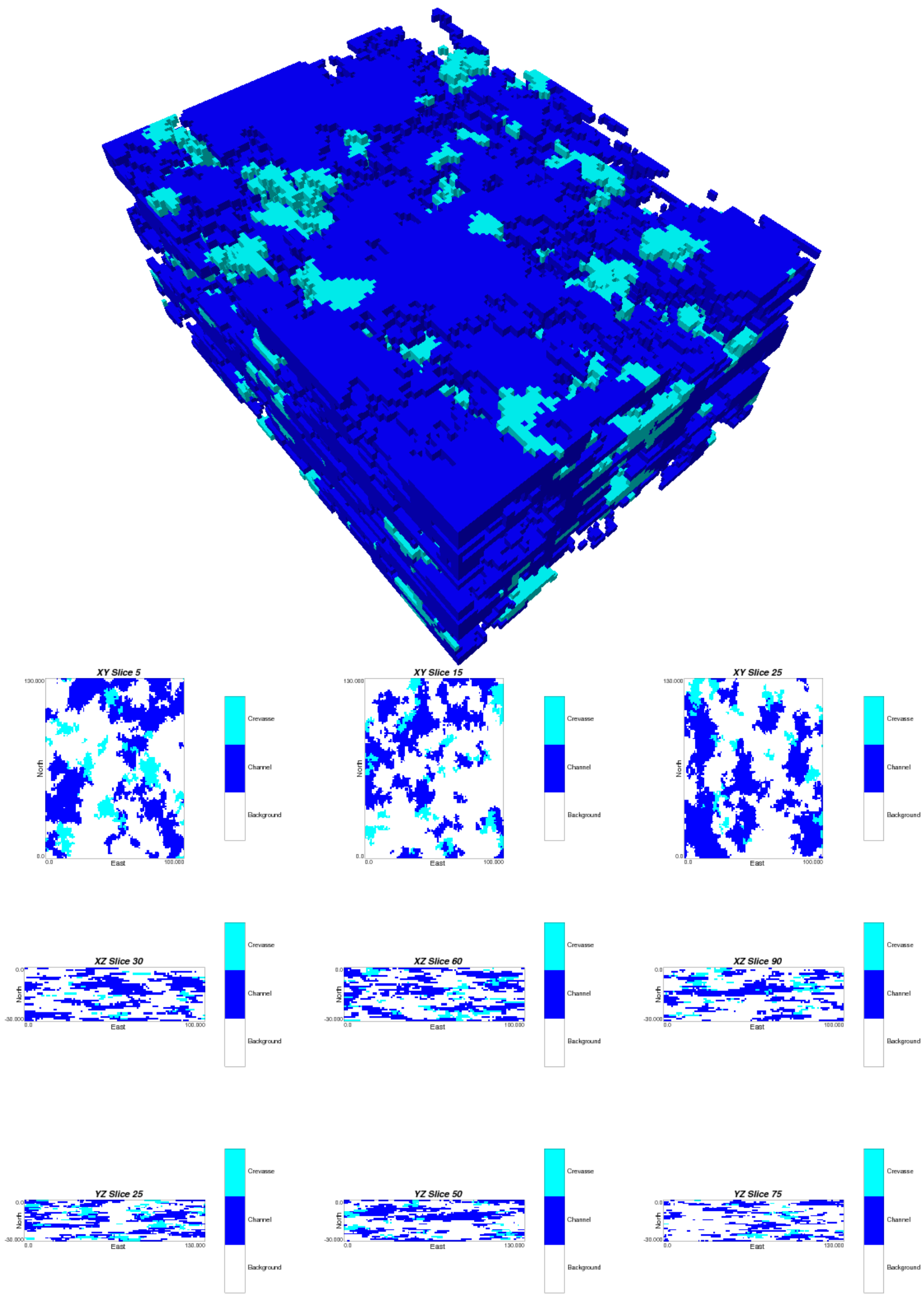


Figure 12: One realization created using SISIM. Top: isometric view; bottom: 2D slices.