Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-567-RC2, 2016

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Multiple-point statistical simulation for hydrogeological models: 3D training image development and conditioning strategies" *by* Anne-Sophie Høyer et al.

Anonymous Referee #2

Received and published: 21 December 2016

This paper presents a case study of multiple-point statistical simulation of sand/clay occurrence. The paper focuses on two aspects of multiple-point statistics: (1) 3D training image development and (2) different conditioning strategies to incorporate borehole data and geophysical data. This is a very relevant topic. Especially the construction of 3D training images is indeed still difficult. There are definitely some very interesting ideas in this paper, such as the different ways of using borehole data as hard or soft conditioning data. I would like to see, however, some more discussion on the following points:

In the title and aims of the paper, the authors stress the importance of realistic 3D training images. They write that they present a workflow to build a training image. The part

about how they build the training image (section 5.4) is however very short. From this short description, it is not clear to me how exactly the training image was constructed. On what data is the TI based? On seismics or on the existing Miocene model? How were the shapes/geometry/position of the clay/sand featureds determined? What was the input in the Geoscene3D model? How exactly does this model work? What are the assumptions? Is this a manual or an automatic method? What are the "interpretation points" and where do they come from? They refer to a methodology in an earlier paper, but this paper uses completely different input data, i.e., airborne EM data?

On page 7, the authors write that "the results are evaluated and compared against the structures expected from the Miocene model"? How exactly? Was this model not already used as a basis for the TI? Is it then fair to use it again to evaluate the resulting structures?

The simulated structures are relatively simple and uniform. Is it really necessary to use MPS? If you would model sand/clay occurrence with a more simple method such as indicator kriging, you would probably get similar results? It would strengthen the paper if you can prove somehow that this relatively complex approach has significant advantages over simpler approaches.

More in general, the conclusions of the paper are only based on visual inspection of the simulated clay/sand patterns. There is not objective or quantitative way of comparing the different results. For example in Figure 11: could you not use cross-validation or something similar to come to a more objective comparison of the different realizations? I also wonder how relevant the differences between the different realizations are, e.g. when you state that "the realizations showed a significant sensitivity to the TI". If you put these different realizations in a groundwater flow model, it is quite plausible that they all give similar results. It would be really interesting to using your geological model for some flow runs to see whether the different realizations based on different conditioning strategies really result in different groundwater flow patterns.

The authors claim that in the study area many of the borehole records are of low quality. Why is that? In what sense are they of low quality? How can users in other study areas determine the quality of boreholes and decide whether they can treat the borehole data as hard or soft data?

The results of the paper are based on visual comparisons of individual realizations using different conditioning strategies. It would be really interesting to see multiple realizations for each conditioning strategy to see the variability and uncertainty of different realizations.

On page 9, the authors write that realization rarely have the same spatial variability as the training image. I find this a really strange statement. The idea of MPS is that you produce realizations with similar spatial patterns as the TI. If the realizations do not show similar patterns, this usually means that the parameters have not been chosen optimally or that the TI was for example not large enough. They also claim that therefore a TI should be chosen together with a specific MPS algorithm? If find this really strange. In my experience, all MPS algorithms can produce realizations with similar patterns to the TI given that they are used in the right way. If the authors really want to claim that different MPS algorithms have different capabilities in reproducing patterns, they should show a comparison of the different algorithms.

On page 12, the authors write that "probabilistic models need to be developed and refined in order to utilize the multiple realizations and the uncertainty the represent". There are however many methods available and applications of MPS using multiple realizations to assess the uncertainty. The methods to do this are available but the authors have chosen to work with single realizations.

The paper is clear and well written. The figures are of good quality.

P4, line 24-25: "along extended profile": error in grammar of sentence?

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-567, 2016.