

## **Response to the comments from anonymous Reviewer #1.**

### *Not well supported conclusion* -----

The reviewer is right that the conclusion ‘Comparison of groundwater heads indicated that a training image with higher accuracy helped improving the model simulation’ is based on un-calibrated models. Even though both training images (TI1 and TI2) are developed from the same geophysical data, it is correct that the values for the hydraulic conductivities used in the analysis are obtained from calibrating a model, which resembles the geological geometry of TI1 more than the one of TI2. Thus the above conclusion may not hold if all models are calibrated. We agree in this and in consequence we have redone the analysis and calibrated all models.

### *Methodological issues with the MPS approach* -----

The reviewer casts doubts about the first TI (TI1) with argument that the geophysical measurements smooth out the high frequency heterogeneities, and therefore the high frequency structure is lost in the TI. However, as the reviewer states a TI is normally a conceptual model of the geology, it could even be hand drawn by a geologist (Strebelle, 2002). In a highly heterogeneous system like ours, TI1 does not necessarily hold less information than a conceptual model of the geology. Although some details obviously have been smoothen out in TI1, we believe that it is still representative, especially for an attribute like proportion. As mentioned by Strebelle (2002) “The first recommendation is to generate training images with global proportions reasonably similar to the target proportions one wish to impose to the final model”. We realize that we should have provided a more thorough discussion of the approach and will do that in the revised version of the manuscript.

TI2 was developed by using the TiGenerator and the reviewer has raised concerns about the ellipsoid shape that was chosen for clay lens and states that other shapes could also have been chosen. We realize we should have included a more thorough geological interpretation to justify the shapes. Also the reviewer has some doubts about the procedure used to define the dimensions of the ellipses. We agree with the reviewer and will justify and perform validation on TI2 in the revision.

The reviewer states that with the geological conceptualization of our study area, a simple SIS method could give similar quality simulation as with MPS. We basically agree but the reason we applied the MPS method is that we would like to test the applicability of the high resolution geophysical data (SkyTEM) with the framework of MPS and also to test the sensitivity of the simulations to the definition of TI. The reviewer also suggests providing the parameter values used in MPS. We certainly agree and will add this information in the revision.

*Detailed comments* —————

(1) *Abstract, lines 19, and 21: what are the criteria used to define if a TI image is acceptable or lead to better simulations. The text would be better if more explicit.*

We will elaborate more on this issue in the revised version.

(2) *Page 2, line 27: Other authors did 3D MPS simulations from 2D TI images. The reviewer gives the following references:*

- (1) *Okabe, H. and M. J. Blunt (2005). "Pore space reconstruction using multiple-point statistics." Journal of Petroleum Science and Engineering 46(1-2): 121-137.*
- (2) *Okabe, H. and M. J. Blunt (2007). "Pore space reconstruction of vuggy carbonates using microtomography and multiple-point statistics." Water Resources Research 43(12).*
- (3) *Comunian, A., et al. (2012). "3D multiple-point statistics simulation using 2D training images." Computers & Geosciences 40(1): 49-65.*

*These methods are not always as complicated as the authors suggest.*

Okabe and Blunt (2005) state in their paper: "In order to generate 3D structure from 2D information, measured multiple point statistics on one plane are rotated 90 degree around each principal axis. In other words, measured statistics on the XY plane are transformed to the XZ and the YZ planes with an assumption of isotropy in orthogonal directions". This method is simple but maybe not suitable for all geological environments, since the statistics from XY plane are normally quite different from the ones on XZ and YZ planes. In fact, we started out with using this method but the simulations were not at all representative of the geology.

Okabe and Blunt (2007) state: "We extend this work by generating 3-D realizations from 2-D patterns by assuming isotropy in orthogonal directions".

Further Comunian et al. (2012) first stated the importance of having 3D-TI for 3D simulation. They compared three methods, and they stated that all three are useful. However, the methods require simulation to generate the geology in between the 2D images and this involves certain statistical assumptions.

We agree with the reviewer that methods have been developed from which 3D MPS simulations can be generated from 2D TI images and some of these methods are perhaps not always that complicated. However, assumptions need to be invoked about the statistics in the 3D space. Thus we believe that more reliable simulations are obtained by using 3D TIs.

(3) *Page 3, line 3, Is equifinality really analyzed in this paper?*

'Equipfinality' is an improper word in this context, we will change the wording.

- (4) Page 5, line 4: "The critical step of sequential simulation is the conditional probability" – needs to be rephrased, "the conditional probability" is not a step.

Agree, will be modified.

- (5) Page 5, lines 5 to 14. The explanation of snesim needs to be clarified. The terminology is not described in proper order in the text.

Agree, will be modified.

- (6) Page 6, lines 8 to 10. This part is important. I suggest to extend the explanation. One key aspect is that Krishnan (2008) has shown that the value of tau depends on the values of the conditioning data as well as the patterns. In general it is not a constant. In practice, setting the value of tau is complex, and this technique is not always ensuring that the ensemble of simulations will respect the probability maps that are given as input. This is discussed in detail for example in Allard et al (2012) or Schaeben (2012). One thing that can be done is to adjust by trial and error the value of tau and check the quality of the results. Here, the value seems to have been set a priori. The point is slightly discussed later on but as the authors claim that this is one of the first application of MPS constrained by geophysical data, it would be good to provide more details. It would also be good to show probability maps to check that they are close to the target.

Allard, D., et al. (2012). "Probability Aggregation Methods in Geoscience." *Mathematical Geosciences* 44(5): 545-581. Krishnan, S. (2008). "The Tau model for data redundancy and information combination in earth sciences: theory and application." *Mathematical Geosciences* 40(6): 705-727. Schaeben, H. (2012), *Comparison of Mathematical Methods of Potential Modeling. Mathematical Geosciences* 44 1 : 101–129. doi:10.1007/s11004-011-9373-2.

The reviewer is right that the Tau value is set a priori. We did not use trial and error to set the value, instead we used the most extreme value 1.0 directly in S2 and S4 implying that the highest possible weight is given to the soft data. In another scenario we performed simulations for a value of 0.5 to illustrate the redundancy effect.

The reviewer also suggests adding probability maps of the simulated models. We will generate such maps and evaluate their significance and on this basis decide whether or not they should be included in the revised version. However, cross section figures, whether lateral or vertical, can only be shown at certain places while Fig. 7 provides an overall view.

- (7) Page 7, line 5, I suggest to rephrase the sentence. It is not clear what is a "geostatistical descriptive map".

Will be modified.

(8) *Page 8, lines 1-5. I think that the method is not correct. See General comments above.*

As stated above, we will improve the method.

(9) *Page 8, lines 7-19. The flow and transport model is not described accurately enough nor a citation is given to provide the details. This is not sufficient.*

Will be expanded.

(10) *Page 8, line 24. What does it mean “head tends to be stable”. Please provide a more accurate statement.*

Will be more thoroughly explained.

(11) *Page 9, lines 23 and 24. The notation S1, S2, etc. should be introduced in the text and not only in Table 1.*

Will be added in the text.

(12) *Page 10, lines 20-29. A figure showing the probability maps and allowing to compare them with the target probability derived from the geophysical data would help understanding the results.*

As stated above, we will add such figures.

(13) *Page 12, lines 16-17. “Comparison of groundwater heads indicated that a training image with higher accuracy helped improving the model simulation”. What does “training image accuracy” means here?*

We agree that the statement is too general and will be elaborated in the revised version.

## *Reference*

Strebelle, S., 2002. Conditional Simulation of Complex Geological Structures Using Multiple-Point Statistics. *Mathematical Geology* 34, 1–21.