Response to Referee #3 comments

We thank Referee #3 for his/her valuable comments and suggestions. The response to the individual comments is given in the following. The original review is quoted in *italics*, whereas the response of the authors is given in **bold** font.

Multiple-point simulation (MPS) is a geostatistical simulation technique first developed at Stanford University in the early 2000's. An MPS algorithm is used to reproduce spatial patterns, such as connectivity, that are depicted in a training image (TI), which contains the possible spatial configurations for any given geological object and relationships between objects. A TI contains only spatial patterns and their respective likelihoods. A frequent pattern appears more often in the TI than a rare one. The actual position of a pattern in the TI is largely irrelevant, the MPS algorithm sees a set of patterns and tries to set them together through a randomization process. In order for an MPS simulation to produce a reasonable representation of any given geological system, it must have to honor some conditioning data and have a method of accounting for spatial trends in the probabilities of selecting patterns from the TI.

The title of this paper suggests that, in this case, these MPS simulations are to guide subsequent hydrogeological applications. The title also suggests the main thrust of the paper is the importance of developing realistic 3-D TI's and strategies for conditioning MPS models. However, this topic is only discussed rather briefly (in 10 lines) in section 5.4 on page 7!

In the manuscript, the TI characteristics are described in "5.4 Training Image", while the details of the strategy we have used during the TI development are provided at the beginning of the section "6. Results", and discussed thorough Fig.s 9 and 10. Along the paper (e.g., at page 7, lines 15-22; page 9, lines 19-32; and page 10, lines 1-8), we emphasise the importance of unconditional realizations to assess the quality of the selected TI (in combination with the used simulation algorithm).
Now, we feel that it would be probably better to move some of the material from page 7, lines 15-22, to the methodological section "5.4 Training Image". We will do this in the revised version of the manuscript. Moreover, we will add further details to the description of the TI development.

Regarding the conditioning strategies, they are discussed throughout the entire manuscript.
In the first place, they are introduced in the methodological section "5 Defining MPS input information". For example, sections "5.1 Seismic data" and "5.2 Existing 3D model" are devoted to the reasons for using the seismic data and the boundaries of the pre-existing geological model as hard conditioning, while section "5.3 Borehole data" describes how to translate boreholes into soft conditioning data by means of a moving window approach and discusses the importance of kriging the "localized" borehole information to enforce the necessary spatial trend.

The results of the application of the different conditioning strategies - introduced in the section "5 Defining MPS input information" - are then presented and compared in section "6 Results".
Section "7 Discussion" examines the assumptions and choices we made and the future possible developments of the presented conditioning strategies and TI development workflow.

Thus, we do not agree with the reviewer when s/he writes that we discussed "*realistic 3-D TI's and strategies for conditioning MPS models***" "***rather briefly (in 10 lines) in section 5.4 on page 7!***". We definitely spent a large portion of the manuscript to go through the optimal approaches for conditioning and TI construction.**

I thus found the paper rather confusing and yet I recognize that the authors and their organizations have considerable experience in 3-D geological modeling for hydrogeological applications, and that there is a considerable body of observational data in southern Denmark that can support the development of multiple-point simulations of the subsurface environment.

In short, while individual sentences and paragraphs use consistently good English, I became uncertain about many important details of their research and objectives. The paper is quite lengthy in its current form, yet it leaves many questions unanswered. If fact, I believe the paper raises more questions than it answers.

I therefore propose that, for publication, the authors undertake to reorganize the current text into something similar to the following:

Section 1: Introduction: This should clearly define the background, objectives, a scope of this project. These topics, I believe, include: âĂć A desire to evaluate the Miocene sediments over a 2810 sq.km. area of southern Denmark where they provide the source of most drinking water. âĂć For about 22% of the area, there is a detailed 3-D stratigraphic model (lithostratigaphic and/or hydrostratigraphic?) developed by deterministic methods (the Tonder model) âĂć Southern Denmark has some high-resolution seismic surveys that can be used as conditioning data for MPS simulations (However, the authors need to provide more information about the spatial adequacy of these surveys, not just that they total 170 km and are shown (rather poorly) on Figure1). âĂć While existing borehole records are available, they are of relatively low quality (WHY?) and most borehole are relatively shallow, so these can provide only limited-value conditioning data. âĂć The project was undertaken to determine if MPS could produce 3-D subsurface information over the entire southern Denmark area more efficiently than deterministic modeling, yet still produce information of value to further hydrogeological models.

Section 2: The Study Area and Available Data Sources: This should summarize its geological character and assess the various data sources. This can be accomplished by revising as necessary existing sections 2 and 3 and Figures 1-5. A discussion about trends should be enhanced.

Section 3: The Experimental Process. This needs considerable expansion from existing section 4. Several questions arise from reading the existing paper. Chief among them: Was the Tonder model the source used to develop the TI? Currently this is unclear. Later the use of two TI's is noted. How were they developed/selected? What are the spatial characteristic patterns desired in the TI?

Section 4: Analysis of Results. This will combine information from existing sections 5 and 6 and some of 7. It also should address several of the limitations defined below.

Section 5: Discussion and Conclusions. This should be relatively short, but include some of the ideas in existing sections 7 and 8. It also should address the need to determine what level of subsurface detail is required to produce an acceptable groundwater management tool for regional and more site-specific applications in southern Denmark (see my final comments).

We arranged the paper as follows:

1) **"1 Introduction" - the first section after the abstract - is devoted to a very general and brief discussion of the existing literature on MPS.**
2) **The second section "2 Study area" deals with the overall geological framework of the area. So, this section delineates the general geological setting where the unit investigated in the paper (the Miocene) is embedded. This, at the same time, contextualizes the presented research and better defines its limits.**
3) **Section "3 Data" describes in detail the amount and characteristics of the different kinds of data available (borehole data and seismic measurements). In particular, a large part of the paragraph "3.2 High-resolution seismic data" discusses - to a reasonable level of detail for the purpose of the present study - the different specifications of the seismic data available.**
4) **Section "4 Establishing framework-model constraints" goes into details about the specific geological unit targeted in the research.**
5) **"5 Defining MPS input information" is the most methodological section of the paper. Here, the approaches used in the comparison throughout the manuscript are described. For example, it is discussed how (and why):**
   - **(i) the seismic data and the existing manual adjacent model have been incorporated as hard conditioning (paragraphs "5.1 Seismic data" and "5.2 Existing 3D model");**
   - **(ii) to translate the boreholes into soft probability to address their uncertainties (paragraph "5.3 Borehole data");**
   - **(iii) to build an effective TI based on the outcomes of the unconstrained simulations ("5.4 Training Image").**

6) **Section “6 Results” is about the detailed comparisons of the outputs resulting from the application of: (i) the iterative approach for the construction of the TI and (ii) the different conditioning strategies detailed in the previous section.**
7) **The second last section, “7 Discussion” discusses the assumptions/choices made across the paper and their possible limitations and possible, future, developments.**
8) **“8 Conclusion” is a very concise section where we simply summarize our results.**

We trust this is a reasonable way to present our research.

Clearly, other scientists might think differently. To some extent, it is simply matter of taste, as long as the rationale behind the choice is evident. And we believe that already in the present form, the overall logical organization of the paper is quite clear and effective.

Regarding the specific questions posed by the reviewer:

1) **Many factors affect the quality of boreholes: (i) the drilling methods (e.g.: hydraulic-rotary drilling and air core drillings); (ii) the drilling purpose (sometimes, if the goal is simply to reach a specific target, the lithological description is not a priority); (iii) the age of the boreholes (older boreholes are generally less reliable); (iv) the presence of simultaneous wireline logging data. During the geological modelling phase, all the borehole records are (or should be) checked for inconsistencies. In the study area, as stated in the manuscript, the few available deep borehole are characterized by high level of uncertainty.**
2) **We acknowledge that, even if the procedure for the construction of the TI is discussed in several parts along the paper (not only in the paragraph “5.4 Training Image”, but also at: page 7, lines 15-22; page 9, lines 19-32; and page 10, lines 1-8), expanding the associated explanation could make our point clearer. And we will do that in the new version.**
3) **MPS has not been considered as a way to “***produce 3-D subsurface information over the entire southern Denmark area more efficiently than deterministic modeling***“. With the available input data (in terms of types, quality, and amount), MPS is the only way to produce a meaningful estimation of the geological variability within the Miocene at the scale that is reasonable for groundwater investigations.**

As I reviewed the current draft, I assembled a list of what I consider to be its current limitations. These include:

1) Apparently, the research so far has focused on examining the role of various hard and soft conditioning data, but only a single realization is used for each setting. This is clearly insufficient. Repeated applications of MPS will produce a sequence of slightly different realizations even with the same conditioning setup, and it should be possible to quantitatively evaluate their similarities and compare this to the differences introduced by the changed conditioning strategies.

Clearly, the differences we highlighted for each conditioning setup are not realization-dependent. This means that they will appear consistently for every realization obtained with the same conditioning settings. Just to mention an example, all the realizations generated with the "borehole as hard data" (Fig. 12b) will be "perfectly" matching all the borehole samples (page 8, lines 11-18) by construction. The same is true for the "borehole as soft data" (Fig. 12c); in this case, we confirmed empirically that the SNESIM ignores (Hansen et al., submitted) "localized" soft data (e.g.: page 8, lines 4-7).

Of course, considering more realizations would allow uncertainty assessment. And this would be, in principle, very interesting. However, as the reviewer has pointed out, the paper is already quite long and a discussion about the uncertainty would fall out of the original scope of the manuscript. This article is, in fact, meant to be simply about the best conditioning strategy and the optimal preparation of the TI.

2) Regarding the Training Image aspect of MPS, it seems that two TI's were used. It is unclear how they were developed and what underlying geological concepts or knowledge were used to develop them. Figure 9 does not clearly show the sand/clay layers – the colors are not ideal for this, and the text (line 15 Page 7) merely states one realization has more layers than the other. Does either seem more likely with the geological knowledge available? Are these layers defined as channels or sheets (continuous layers)? How does either TI relate to conditions within the Tonder model?

One of the main points this manuscript would like to convey is the importance of unconditional realizations to evaluate the effectiveness of the selected TI. In particular, we discuss this in "5.4 Training Image" (where the general TI characteristics are detailed) and in "6 Results" (concerning the strategy for the TI development).

However, for sake of clarity, it seems necessary to add, in the new manuscript, further details. So, in the revised version, we will underline that: (i) the first TI test was based on the existing 3D geological model (the Tønder model) covering an area adjacent to the simulated one; (ii) then, during several iterations, this first TI was manually adjusted based on the unconditional outputs. In the paper, for simplicity, only the initial and final TIs are shown in Fig. 9 together with the associated unconstrained realizations (Fig. 10); (iii) this iterative process was stopped as soon as the corresponding realization was able to mimic the geological features expected in the Miocene.

Those expectations about the geology are based: (i) on our prior geological understanding of the area, (ii) the available seismic lines, and (iii) the few existing deep boreholes.

3) The assessment of the results is mostly qualitative. Quantitative tools to do exist and should be used.

We do not agree with this remark. For example, with respect to a quantitative analysis of the proportions variability, we think that our point concerning the effectiveness of using the kriged sand probability to enforce the proper spatial trend is clearly demonstrated via the comparison between Fig 14b and Fig. 8. The two figures allow a voxel-by-voxel

comparison between the soft conditioning distribution and the final corresponding realization. So, it is not clear to us what kind of more quantitative argument we should use instead.

Maybe, regarding the quantitative assessment of the patterns produced by using the two different TIs in Fig. 9, it could be useful to report the values of the mean volume of the sand bodies in the realizations in Fig. 10 generated respectively with TI1 and TI2. So, even if the difference between the patterns in Fig.s 10a and 10b is evident, if requested, we will specify these values in the revised version of manuscript.

4) The cited literature appears to miss several important more recent studies. Attached are a few representative paper citations.

We see the reviewer's point. In the new version, we will include additional, relevant references.

FINAL COMMENTS

MPS is an interesting and potentially powerful method for developing very useful subsurface geological models. I am aware that at least some of the authors have experimented with other simulation approaches, such as TPROGS to apparently successfully simulate facies heterogeneity in buried valleys. It would be interesting for them to include at least a short comparison between MPS and other simulation approaches. The current paper assumes the reader to be proficient in MPS concepts. This may not be true in many cases, so a short comparison in the introduction might broaden the readership and understanding of the importance of this line of investigation.

We believe that the paper (already "*quite lengthy in its current form***") would not benefit from a comparison with other simulation approaches as the manuscript is not about checking the performances of MPS as a tool to incorporate complex statistical information against other "simpler" approaches. This article deals solely with the optimal strategies (within the MPS framework) for the soft/hard conditioning and for the construction of the most (geologically) effective TI.**

In addition, the approach of the submitted article is based on TIs for the description of the information about spatial structures, while TPROGS does not make use of TIs and, instead, requires the user to define transition probabilities from which the realizations are then simulated. Therefore, given the topic of the paper, it is not natural to consider the use of TPROGS.

I believe the ultimate goal of this research is not to model the Miocene of southern Denmark as a purely academic exercise, but to use this information to guide groundwater management schemes. I wonder what groundwater model sensitivity *analysis would yield in terms of necessary subsurface detail for producing acceptable groundwater resource management at regional or site-specific scales?*

Obviously, somewhat less precise spatial definitions are likely to be required for regional assessments. On the other hand, site-specific studies may be unreliable if based only on MPS inputs without careful additional local conditioning. So, the question arises, "Where does MPS fit within this overall objective?" This question also reflects on the limitation noted on page 9 (lines 13-15) the present inability of MPA to handle graben structures and faults.

The present paper is about: (i) how to include, in the most effective way, the available data as simulation conditioning, and (ii) how to build a TI that is capable to reproduce the geological features we want to see in the realizations. In particular, regarding point (ii), the ability to generate meaningful geological realizations must be seen in a general perspective and, so, relevant per se.

Moreover, MPS is a way to "see" beyond the (strictly speaking) available data. In fact, via the prior geological knowledge formalized by the TI, it is possible to "reconstruct" features compatible, at the same time, with the data (e.g.: the geophysical measurements), but also with the desired prior geostatistical information. This is a very important aspect in the paper: a significant amount of data - but not as much as we desired (as always happens) - was available together with the knowledge of the geological features we wanted to see in the realizations. Thus, we investigate the best workflow to utilize the full potential and flexibility of MPS methodology to jointly exploit those two pieces of information and obtain a realization that is, simultaneously, matching the data (i.e., boreholes, geophysical measurements, and pre-existing geomodels), and having the required spatial characteristics.

The scale of the geofeatures necessary for flow modelling purposes is hard to decide before entering in the more hydrological aspects of the problem. In the paper, we show how to get those features in an effective way and we test the proposed approach at the scale we think is the more appropriate one. Further analysis would be definitely necessary, but we think they are not falling in the scope of the present study.

Despite my several criticisms, I think this is a potential important paper and hope the authors will consider reorganizing it and adding in a few details on some important methodology issues, while at the same time focusing on the TI and conditioning strategies.

SOME SUGGESTED REFERENCES:

Boucher, A. (2011) Strategies for Modeling with Multiple-point Simulation Algorithms.IN: "Closing the Gap", 2011 Gussow Geoscience Conference, Banff, Alberta. 9p.

Kessler, T. (2012) Hydrogeological Characterization of Low-permeability Clayey Tills – the Role of Sand Lenses. PhD Thesis, Department of Environmental Engineering, Technical University of Denmark, Lyngby, Denmark. 80p.

Klenner, R., Braunberger, J.R., Dotzenrod, N.W., Bosshart, N.W., Peck, W.D. & C.D. Gorecki (2014) Training Image Characterization and Multipoint Statistical Modeling of Clastic and Carbonate Formations. PowerPoint Presentation, 2014 Rocky Mountain Section AAPG Annual Meeting, Denver, Colorado. 27 slides.

Meerschman, E., Pirot, G., Mariethoz, G., Straubhaar, J., Van Meirvenne, M. & P. Renard (2013) A Practical Guide to Performing Multiple-Point Statistical Simulations with the Direct Sampling Algorithm. Computers & Geosciences. 52. p. 307-324.

Straubhaar, J., Walgenwitz, A. & P. Renard (2013) Parallel Multiple-Point Statistics Algorithm Based on List and Tree Structures. Mathematical Geosciences. 45. p. 131-147.

Straubhaar, J., Renard, P. & G. Mariethoz (2016) Conditioning multiple-point statistics simulations to block data. Spatial Statistics. 16. p. 53-71