Solid Earth Discuss.,
https://doi.org/10.5194/se-2019-78-RC2, 2019
© Author(s) 2019. This work is distributed under
the Creative Commons Attribution 4.0 License.

Solid Earth Discussions

Interactive comment on "Topological Analysis in Monte Carlo Simulation for Uncertainty Estimation" *by* Evren Pakyuz-Charrier et al.

Gautier Laurent (Referee)

g.laurent.research@gmail.com

Received and published: 9 July 2019

The manuscript entitled "Topological Analysis in Monte Carlo Simulation for Uncertainty Estimation" proposes to combine mathematical description of topology as presented by Thiele et al. (2016) to Monte Carlo approaches to geological uncertainty estimation as presented by Wellman (2013) and Lindsay et al. (2013). In particular, the authors demonstrate the use of DBSCAN clustering algorithm for highlighting the main topological families.

The proposed approach provides very interesting complements for this overall uncertainty estimation approach. It's interest lies in the fact that topological changes in produced models often have a first order impact on resulting physical simulations (eg., for

Printer-friendly version

Discussion paper

fluid flow simulations in reservoir modelling). It is only natural to use topology as a metric for clustering the sampling outputs.

However, there are several limitations to the proposed approach, which should not prevent its publication but rather call for further discussion:

- First of all, this manuscript inherits from the limitations of the approach to uncertainty sampling that consist in varying the input data independently. This is making the assumption that uncertainty is mainly due to measurement errors, while the matter of data representativeness and lack of knowledge account for a probably larger part in the geological uncertainty. (cf., discussion below)
- A second limitation comes from the use of adjacency matrices as a unique measure of topology. This is quite efficient indeed, as demonstrated in this manuscript, but this metrics is unable to completely account for topology. As a matter of fact, two models that are not homeomorphic could have the same signature (cf. attached drawings), which limits the justification for calling it a signature. This should be better discussed. As a consequence, considering your remark at p8L37 "These results indicate that topological signatures may help differentiate favourable scenarios in ore reservoir or oil and gas modelling applications." most of these traps are actually invisible in the topological signature whereas they represent peculiar topological features. This is a serious limitation in my opinion and should be discussed.

Besides these drawbacks, the manuscript is relatively well written and illustrated, even if the figure captions should be more informative. Some recurrent grammatical errors are still making the reading difficult in places. Avoiding the inadequate use of possessive apostrophes should greatly improve the reading. Acronyms that are not commonly used in the community are used throughout the document, which makes it difficult to follow the discussion in place.

CC BY

Based on these appreciations and the detailed comments below, I would suggest that major revisions should be made before accepting this manuscript for publication.

1 Detailed comments

1.1 Lack of explanation of some basic concepts

- p2L.9: produce probabilistic geological models (PGM) and uncertainty index models (UIM). should be defined
- p2L. 15 "Plausible models' incompatibility is damaging to the relevance of MCUE because the PGMs and UIMs implicitly assume plausible model homogeneity." -> I don't understand this statement. What is this assumption for homogeneity?

1.2 The source of sampled uncertainty

- p2L.39: these are only a (probably small) part of the uncertainty. Problem of representativity is much bigger for data points, and what about other type of uncertainties
- p3L7: the problem is that simplifying the source of uncertainty as a measurement error make the following statement possible "Sampling is usually made independently as errors do not show any spatial dependency, because measurements are physically independent (Pakyuz- Charrier et al., 2017b).", which is hiding the fact that different measurements may be sampling a same structure thus making the measurements physically dependent. This situation seems closer to a rule than an exception for geological structures, on the contrary to what is suggested by the following paragraph (limiting it to cyclicity or observed heteroscedasticity).

CC BY

Although disturbance distribution is not the main topic of this paper, this limitation to the method should be further discussed and stated more clearly.

- p4L31: "Moreover, these methods work best if errors are spatially dependent. This is normally not the case for sparse geological structural measurements taken individually. Actually, there is no logical reason to consider that the measurement errors related to, for example, two foliations measured with a compass in different areas are dependent on one another." Is it equivalent to say measurement errors are independent and the uncertainty about the measures are independent? If two measures are informing about a same structure, then one is also providing information about the other measurement and they are not independent. Consider an uncertain measurement A dipping 15 degree to the north with a possible error of + or - 5 degrees, if you add another measure B dipping 5 degree to the north + or - 5, you are probably more likely to have a somewhat flat layer dipping 10 degrees to the north than to have a large fold structure. Therefore the errors should not be looked at independently. It does not mean you can not make this assumption to simplify the mathematical treatment of a dataset, but you can not justify it by stating that uncertainties are independent.

1.3 Unimodality

- p4L9: I don't see why computing stratigraphic variability or information entropy would have any assumption about unimodality?

- p4L10-14: entropy and stratigraphic variability are scalars that inform about the variability in a set of models. They are not meant to reconstitute a whole distribution. It is not sufficient information for a bimodal distribution but it would not be for a general unimodal distribution as well. I don't see your point and your argument should be clarified.

CC BY

- p4L16: how do you define a "homogenous population of plausible models" or heterogeneous (P4L19)

1.4 The discussion about linearity seems confused

- p4L21 : the concept of non-linearity that you are using here is not clearly explained. On one side you are using linearity in terms of the mathematical combination of data values (p4L22) but the non-linearity that is then discussed is based on the geometrical and more particularly the topological continuity of the discussed changes. Of course, modifying the geological ruleset introduces more extreme changes in the model and would certainly impact its topology, but simple (linear) perturbation of a given dataset may also result in topological changes even with the same geological rules applied. This is one of the disadvantages of implicit approaches vs. explicit ones. You could refer to Collon Caumon 2017 for these aspects: DOI: 10.3997/2214-4609.201701144 I would suggest to clarify what you describe as a linear change and to rely on topological terms such as the concept of homeomorphism here.
- p4L27: (ii) plausible dataset variography is not a reliable indicator of plausible model homogeneity. Why would that be? What is your point?

1.5 Topological signature

- Fig. 5: why discretizing, boundary model (eg. surface model) should contain enough information?
- how are faults described? It looks like they are not whereas this is a major aspect in model topology.

CC BY

- the diagonal of 1 is useless and should be removed for compression. Actually, later in the manuscript, with the second example we realise that this is actually encoding the presence of the layer. This should be explained more clearly.
- P6L9-10 : This is not clear: "Note that clustering the topological signatures of the plausible model suite implies that quantitative topological stationarity is not required." To me it seems that clustering doesn't required stationarity, but this is not what this sentence is saying.

1.6 Entropy

- eq. 3: why is there a k factor inside the log parenthesis?

1.7 Post-process

- p7L7: Central statistics: this term should be explain instead of calling it straightforward This process is very interesting but its implementation needs to be better explained In addition, this process is not demonstrated in the examples. Please either detail or remove.
- p8L25: the threshold of 60 is not justified. Could you discuss this choice?

1.8 References

- inconsistency of reference ordering, should generally be alphabetically or chronologically, e.g. P2L1-2: "Schweizer et al., 2017;Wang et al.,2016;Nearing et al., 2016;Aguilar et al., 2018;Mery et al., 2017;Dang et al., 2017;Lark et al., 2013"
- p2L13. and p4L18 (Thiele et al., 2016a;Thiele et al., 2016b) should be gathered into (Thiele et al., 2016a,b)

CC BY

- P4L15: Pakyuz-Charrier et al., 2018;Pakyuz Charrier et al., 2018 : he same reference is repeated twice
- It seems that the three appendices are not referenced in the main text.

1.9 Grammar remarks and other miscellaneous remarks

- p3L22-23: the construction of the sentence and the enumeration are making it difficult to read.
- p3L35: the concept of using triple lines for comparing models require a reference or further explanation
- p4L2: groundtruthing doesn't look like a formal english word and should be explained anyway.
- p4L9-11: These two sentences are repeating the same information, I suggest to remove the first one which is less detailed. "The underlying assumption is that the plausible models constitute a unimodal population and may be analyzed as such. The UIM used in MCUE are scalar proxies for categorical uncertainty and one of the critical conditions for a single scalar to be representative of the uncertainty of a variable is that it has to be distributed unimodally."
- p4L29: remove "are"
- in the whole manuscript: authors are making an extensive use of possessive apostrophes in forms than are grammatically incorrect after eg. https://www.ef.com/wwen/english-resources/english-grammar/forming-possessive/
- p6L2: there is a double dot at the end of the line

CC BY

- p6L4: there is no ’s’ at the end of signature in "Cumulative observed topological signatures graphs"
- p6L17: each points -> each point
- P6L24: implementation wise -> implementation-wise
- p6L39: (3 -> Eq. (3)
- p7L4: incorrect grammar construction
- p7L11: central, -> central statistics,
- p7L33: "with a," remove the coma
- p8L39: Error! Reference source not found.

1.10 Image quality

- The quality of fig. 3 is poor (to the point it is difficult to identify colors)
- The quality of fig 4 should be improved as well.
- The figures are not sufficiently described and rely to much on the text for explanations. For example, the caption should state why annotation such as ellipses are made for.

1.11 Fig 3

- is lacking description and explanation

CC BY

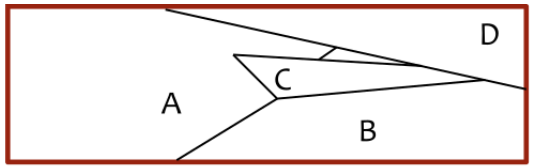
- there are misleading geological terms: for the models labeled grabens, only the one in the top right corner is actually a graben, similarly for the horsts, only the top left one is actually a horst the different tilted blocs are in fact not tilted and a layer is missing for the two model in the lower left corner of the tilted models.
- why are the arrows labelled plausible models?

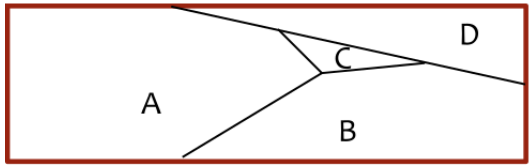
Interactive comment on Solid Earth Discuss., https://doi.org/10.5194/se-2019-78, 2019.

Printer-friendly version

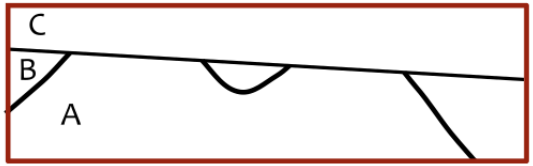
Discussion paper

CC BY





	A	B	C	D
A	1			
B	1	1		
C	1	1	1	
D	1	1	1	1



	A	B	C
A	1		
B	1	1	
C	1	1	1

Fig. 1.

CC BY