Interactive comment on “Data assimilation and uncertainty assessment in 3D geological modeling” by Daniel Schweizer et al.

F. Wellmann (Referee)
wellmann@aices.rwth-aachen.de

Received and published: 9 February 2017

General comments:

In their paper “Data assimilation and uncertainty assessment in 3D geological modeling”, the authors present an interesting, novel and well-structured analysis of uncertainties in the context of 3D geological modeling. This topic is in itself of high relevance, and the authors show this here specifically with an application to a case study where a 3D geological model is used as a basis for a subsequent analysis of ground heave damage in the city of Staufen, Germany. Specifically, the authors compare uncertainties in subsequent stages of detail in model construction - an aspect that has, to my knowledge, not been investigated systematically in an applied case study for 3D geological modeling before.

Specific comments:

One important point that should be adapted in my opinion is the use of the term “data assimilation”. This term is typically used in a very different context (as an update of parameter and state space in the course of a dynamic simulation, but for the same model)! In order to avoid confusion (or wrong expectations by readers), I would suggest to remove any reference to “data assimilation” from the paper and replace it with a better fitting term. For example for the title: “Uncertainty assessment in 3D geological models of increasing complexity”.

Specific comments to location in text (identified by page/ line number):

- 1/17: I would claim that 3-D models are mostly preferable because our object of study is intrinsically 3-D...
- 6/5: Unclear what exactly “model complexity” refers to: the number of parameters or structural features?
- 6/5: What is meant with “data acquisition”? Do you mean “data integration”? Please clarify;
- 6/15: The listing of the steps already contains details about the specific model that is used later. A clearer separation from the general approach (here) to the specific application (in Sec. 4.2) would be better;
- 9/23: I would generally suggest to use “average entropy” instead of “total entropy” (even though I am probably to blame for the second term, but it may lead to confusion);
- 9/16-20: Other interesting aspects at this point could be the “geodiversity measures” of Lindsay et al., or the topological analyses of Thiele et al., 2016;
• 9/23: Note that this definition is highly sensitive to outcomes with small probability! Could be more robust when using a threshold value of probability.

• 10/10: I am not sure that the term “city block distance” is correct here. Equation (10) seems to be simply the L1 norm over the cells for two combined sub-regions (as N is the number of cells). In the definition of Paul and Maji this is the number of features (“m” in their equation 1)! Interpreted in this context, each cell would be one “feature”. Is this what you intend to express here? Please check and/or clarify;

• 12/19ff: Important points considering the specific implementation:
  provide all parameters and the assigned probability distributions in table form; also, please describe the reason for the choice of these parameters (even if based on educated guess); two other choices are made: generating 30 realizations, and using a cell size of 5 m³. What is the reason for these choices? Especially concerning the number of realizations: is this based on an estimate of convergence (note that, for example, average entropy could be used here).

• 13/12: In my opinion, this is not a limitation of this specific approach, but the general statement of epistemetic uncertainty and related to missing knowledge;

• 13/13: Model 4 not only may, but surely will, underrepresent true structural complexity by definition - because it is a model. In my opinion, the question is only if it represents complexity sufficiently for the specific purpose of the model. Please adjust or discuss this point (also in conclusion);

• 13/19: Please note that measures of information theory are not limited to point estimates (see e.g. Wellmann 2013 for exactly this context, please excuse self-citation);

• 15/19: What exactly does “data specificity” refer to at this point? Please clarify;

Figures: All figures are excellent throughout the manuscript and already of publication quality.

Technical corrections (identified by page/line number):

• The names in the author list seem to be in the wrong (first name/last name) order;
• 2/17: “we hypothesize”?
• 12/4: “in addition to”?
• 12/10: “ambiguous” instead of “equivocal”?
• 12/20: “includes minimal constraints”?

Interactive comment on Solid Earth Discuss., doi:10.5194/se-2016-174, 2017.