

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

G. Caumon (Referee)

guillaume.caumon@ensg.univ-lorraine.fr

Received and published: 14 March 2017

General comments

This paper is a nice contribution on the topic of geological uncertainty assessment. It addresses more specifically the problem of how structural uncertainty evolves as subsurface data is integrated into a set of possible 3D models. This is an important area of research which has implications both in fundamental and applied geosciences.

As in Wellmann and Regenauer-Lieb (2010), this paper makes a significant use of information entropy to visualize and analyse geological uncertainty. The main novel aspects covered are:

- A true case study in SW Germany is considered, which allows the authors to

Printer-friendly version

Discussion paper



consider four base models built from more and more detailed data acquired to understand ground heave related to geothermal exploitation.

- The uncertainty model considers both spatial data perturbation as Wellmann and Regenauer-Lieb (2010) and Lindsay et al. (2014), but also includes geometric model perturbation away from the data, which should better capture uncertainties.
- Model comparison is considered in the analysis

In addition, this paper is easy to read and well-organized, so I consider it deserves publication after minor to moderate changes. Below are some remarks which I hope will help improve the paper.

Specific comments

1. The title of the paper is very general and should probably be more specific to better reflect the content. The term “assimilation” is a bit inappropriate, as it is generally used to incorporate data through time in dynamic systems.
2. Overall, the paper clearly shows that adding data helps globally reducing uncertainty, and confirms that information entropy can be used to quantify this reduction. However, it also clearly shows that uncertainty can locally increase when new data call for making the model more complex, in particular when faults are added. So, if the area of interest were smaller (localized around the inserted faults), or if the averaging were weighted (e.g., with regard to where the ground heave is largest), the total entropy could decrease as new data are considered, which would be counter-intuitive. The fundamental reason, which is mentioned in the paper, is that the initial model, being parsimonious, under-estimates the uncertainty. The message that parsimonious models can locally under-estimate

[Printer-friendly version](#)[Discussion paper](#)

- uncertainty, is in my opinion, an important conclusion of this work, which should probably be better stressed.
3. It would be useful to indicate all the uncertainty model parameters used in the study: amplitude of geometric perturbation for the various faults and horizons, range of the variograms. Did you change these parameters from model 1 to model 4?
 4. There are several issues in the mathematical notations: symbols should be the same in the text and in the equations; please also check that all symbols are defined and consistent (math style, upper or lower case) in the text (e.g., geological units are currently denoted by U , u or U ; a model is denoted by u_i in p. 9, l. 24 and by i on line 9 of p. 10, then u_i on line 11 denotes a sub-region. I am lost... Indicator functions by \mathbb{I} or $\mathbf{1}$; q_i, q_j on page 9 are not defined (is it the same as Q_U in Eq (7)?).
 5. In terms of geology, I am surprised that a reverse fault had to be inserted in Model 3. What is the strike of that fault, and is it supported by the tectonic history of the area? Could it be a normal fault whose dip has been mis-estimated?

Technical corrections

- p. 1, l.17: “high data consistency and superior data visualization”: Unclear, please reformulate.
- p. 2, l.33: “bur”: but ?
- p. 4, l. 15 “initiated, because”: remove coma.
- p. 5, l. 3: “information that has”: that have

- p. 6, l. 6-10: Why did you create explicit models and not directly implicit ones?
- p. 6, l. 11: The implicit method by Lajaunie et al and Calcagno et al follow the same idea as the one in SKU, but their formulation is different (dual kriging Vs. discrete optimization. A Reference for the method in SKUA is Frank et al. (Computes & Geosciences, 2007).
- p.6, l1: Please give the variogram range and value of maximum perturbation applied on the surfaces for the various models. Some discussion about how these values were chosen would also be useful. NB: the maximum displacement and the Gaussian distribution suggest that a truncated Gaussian is used? Please check. References for this method are Tertois et al. (GSL Spec. Publication, 2007); Caumon et al. (EAGE Petroleum Geostatistics, 2007 and Math. Geosci. 2010); Mallet (EAGE book, 2014).
- p.8, l.11: “U” should have the same font as in Equation (U). Same for all math symbols.
- Eq.(1) “if otherwise”: no need for if.
- Section 3.4: please check mathematical formalism, as the explanations are difficult to follow. Indicating how the distances are computed and their values would also be didactic.
- p.11, l.4: “NO-SW” should read NE-SW; Displacement of about 10m for the fault seems small (the cross-section in Fig. 6 suggests about 50 m).
- p.12, l.4: “direct problem specific data”: please remove direct (source of ambiguity, as the term “direct problem” is often used in a completely different sense in geophysical inversion.

[Printer-friendly version](#)[Discussion paper](#)

- p.12, l.10: This sentence seems to suggest that seismic data are a source of uncertainty. Actually, they aren't, as they do provide information. The uncertainty revealed by seismic sections was present previously, but ignored by interpreters / modelers who tend to strive for the simplest explanation. Seismic data forces to add complexity down to a certain scale. BTW, it could be useful to explain and discuss a bit more what was done in terms of interpretation and modeling by showing one or two seismic lines.
- p.12, l.11: "their connexion to fault networks": Unclear to me; please rephrase.
- p. 12, l. 25: "displacement": perturbation would be more precise (as displacement may also refer to fault slip in the context of 3D structural modeling)
- p. 13, l. 1: Model should read Models
- p.16, l. 8-9: Not sure I understand this sentence.
- p. 18, l.19-20 "adequate parameters for dip and azimuth": Not sure I understand this point. Could you explain further is this is a software interface problem or a more fundamental issue?
- p.18, l.24: It would also be appropriate to cite the seminal work of Holden et al (Math Geol., 2003) on stochastic structural modeling.
- In the conclusion, some more general discussion on the sources of uncertainties could be useful. For example, are the base models 1 to 4 reliable? This study rightfully suggests that Model1 is not reliable in the light of data acquired recently. The corollary is also that Model4 may still be wrong on low data density. Even in high data density areas, isn't there a risk that wrong interpretations have been made? The author's thoughts about these issues would be interesting to read.

[Printer-friendly version](#)[Discussion paper](#)