Solid Earth Discuss.,
https://doi.org/10.5194/se-2017-115-RC3, 2018

Interactive comment on "Monte Carlo Simulations for Uncertainty Estimation in 3D Geological Modeling, A Guide for Disturbance Distribution Selection and Parameterization" *by* Evren Pakyuz-Charrier et al.

G. Caumon (Referee)

guillaume.caumon@ensg.univ-lorraine.fr

Received and published: 25 January 2018

On the use of spherical distributions for data-related structural uncertainty quantification

General comments

This paper shows on a synthetic and real examples that using spherical orientation distributions to describe structural data uncertainty is important in 3D structural uncertainty quantification. This is important, as most work considering structural uncertainty (including some papers I co-authored) have neglected spherical distributions and used simpler and independent statistical models for plane strike and dip. The results show that a more careful consideration of spherical distributions can have an impact on Uncertainty Quantification results, and some interesting statistical insights are provided. The authors also provide their models as supplemental material, which I consider very useful, and give practical guidelines to use Fisher distributions in the Appendix, which is also useful for practioners. So, I think this paper deserves publication. However, I have a few problems and I am still unclear about some parts of the paper. Therefore, I am making several comments and recommendations below, which I hope will help the authors to make the paper easier to read and more precise.

Specific comments

- I had some difficulties to understand the paper. There are locally some purely formal aspects to it, which a careful reading could easliy fix. More importantlty, part of the reason is that several of the statements appear as general truths in the paper, whereas they only hold in some cases or under some assumptions that are not explicitly described. I think these inappropriate generalizations should be addressed before publication. Another reason is that some of the ideas and principles are not always clearly expressed (in particular in Section 3.2). Overall, I find the beginning of the paper not very easy to read and to understand. I have higlighted several of these issues in the annotated pdf manuscript, I hope this will help the authors make the paper easier to follow.
- I think the term MCUE does not precisely describe what the authors have in mind (by the way, "Monte Carlo simulation **for** uncertainty estimation" (as in the title) seems clearer to me than "Monte Carlo simulation uncertainty estimation" as in the abstract and main text). Indeed, the fact that this paper focuses on ** data ** perturbation is not clear from this wording. Indeed, MC perturbation of model parameters is another (and widely used) approach to sample uncertainty

Printer-friendly version

Discussion paper

in structural modeling (see seminal work by Abrahamsen (Geostats 1993) for horizons; Lecour et al (Petrol. Geosci 2001) for faults and many other papers since then). MC simulation is also used to change how data can be connected in structural modeling (see work by my co-authors N. Cherpeau and C. Julio). I understand these model perturbation approaches are not the focus of this paper, but it should be clear to the readers from the outset that there is more to geological uncertainty than orientation data perturbation. Therefore, I would recommend to replace MCUE by a more specific term (inclusing in the paper's title). My two-dime suggestion would be "data-related structural uncertainty quantification", but the authors may find a better term. This distinction is essential and should be clarified.

- I disagree with the statement (line 26, page 3) that data perturbation and kriging are "equivalent to running geostatistical simulation": has this been mathematically shown or experimentally proven? I have 99

- Geological modeling is bound to be implemented with software; however, a paper can gain much by clearly separating the mathematical and methodological principles from the software platform used for demonstrating these principles. So, in clear, I consider that this work is compatible with other implicit structural modeling methods, and that this could be argued before the introduction. Also, a key feature of implicit modeling schemes and a motivation for using them is this work is that they use orientation data. This could be stressed in Section 2.

- Maybe merging Sections 1 and 2 could help more effectively set the scene and introduce the contributions of this paper?

- I would recommend to comment on the links between this paper and Carmichael and Aillères, (JSG 2016). They also use spherical distributions in a 3D structural modeling context.

- I am not sure I agree with the statement that “Uncertainty will then be best represented by disturbance distributions that are consistent with the Central Limit Theorem”. The CLT holds when N independent random variables are added; in the case of orientation data independence may be assumed but is not granted, and N is likely small. I would, therefore, rather state this as a convenient hypothesis for the argument than as an ideal objective. More generally, the main point of Section 3.1 seems to be that Normal and Von Mises-Fisher distributions are appropriate for structural uncertainty management. I am ready to accept that they are very convenient and useful, but I don’t think I agree with the term “appropriate”. For example, Thore et al (2002) show that propagation of seismic velocity uncertainties through reflection data imaging yields non-symetric distributions about horizon positions. The same comments holds for the first sentence of the discussion (Section 5).

- Section 3.2 is not really clear to me. When I first read the paper, I understood that Eq. (4) suggested that all orientation data samples have the same mean / dispersion (and was not clear about why this should be the for samples taken at different locations). I now think that (4) concerns a single orientation at a given location. But then, it would be good to state that $X_1, \ldots, X_n$ on line 26 correspond to repeated measurements (or whatever it stands for), and to specify the meaning of n. Overall, expressing the idea and assumptions behind the equations would help the non-mathematically-inclined readers better understand Section 3.2 and how it can be applied. For example, it would be worth mentioning that x_i are assumed to be independent samples before Eq. (9); That Eq. (10) is nothing else than the total probability formula. I don’t get what N stands for in (11). Overall, after reading through, I get the general idea but I am still unclear about the details and whether n stands for the number of repeated measurements (which is not available in most data sets) or for some number of simulated points.

Printer-friendly version

Discussion paper

I don't understand the principles behind Eqs. (12-17). More explanations would be much welcome. In the end, I am not fully sure at the end of the section about how the development can be used in practice, and what a "propoer parameterization" (page 7, line 26) exactly means. If some orientation measurement device came up with a good evaluation of the orientation probability distribution based on a sound error propagation, would we still need Section 3.2?

- Overall, maybe I missed a point, but it seems that section 3 summarizes some facts from the statistics literature. This is probably useful, but I feel that this is a bit long and I don't clearly see connections to the experiments made in Section 4, which essentially address the use of spherical distributions for sampling data uncertainty. So I am not really clear about the point on posterior predictive distributions (PPD) in section 3.2. Is it really needed in this paper? What does it bring in? Similarly, the discussion in Section 3.3 seems to mainly serve the point that two angle distributions instead of a spherical distribution entail heteroscedastic effects. So maybe it would help to get more rapidly to that point directly in Section 4.

- on p. 10, I am a bit puzzled by the term "dip vector". It is clear that a plane should be represented by the spherical distribution of its pole; but it seems to me that the experiment mainly describes the "dip vector" by the dip and dip angles. If so, the term dip vector seems inappropriate (as it could also be described by a spherical distribution).

- some more details about how the input uncertainty used in the Mansfield case study would be useful (see comments on page 11). Does this connect to Section 3.2 ?

- In the discussion, the authors may want to add a point on spatial correlation or orientation data. Does it make sense to sample orientation data independently? Wouldn't something like Gibbs sampling be good in that case?

- Form: please check that all math symbols in the text have the same font as in the equations. (e.g., line 10, p5)

Technical corrections Please see annoated pdf file.

Please also note the supplement to this comment:
https://www.solid-earth-discuss.net/se-2017-115/se-2017-115-RC3-supplement.pdf

Interactive comment on Solid Earth Discuss., https://doi.org/10.5194/se-2017-115, 2017.