Solid Earth Discuss.,
https://doi.org/10.5194/se-2019-4-AC2, 2019
© Author(s) 2019. This work is distributed under
the Creative Commons Attribution 4.0 License.

Interactive comment on "Bayesian geological and geophysical data fusion for the construction and uncertainty quantification of 3D geological models" *by* Hugo K. H. Olierook et al.

Hugo K. H. Olierook et al.

hugo.olierook@curtin.edu.au

Received and published: 3 May 2019

Reviewer #2:

I thank the editor for inviting me to review this paper. Unfortunately, I find this paper quite underwhelming. The paper describes very little that has not already been shown in previous works, nor do the results convincingly reveal new understanding from the region. It is thus difficult to understand what contribution this study makes either to probabilistic methods in geophysics, or geological understanding of the Gascoyne Province. This is compounded by the authors' inadequate review of existing work and failure to place theirs in context with the discipline. The authors make statements

about the "importance" or their results with no justification, and how that their method "is the only technique that provides a range of solutions" which is false. To reiterate, the authors needs to spend more time reviewing the existing literature. One positive is that the manuscript is well-written and structured. I suspect the authors will be able to remedy many of the deficiencies listed here and produce an adequate revision.

We thank the reviewer for their rapid and detailed review of our manuscript. Please find attached a zip file containing comments to all reviewers, a manuscript with track changes and a clean manuscript with all changes incorporated.

I list the major issues directly below, followed by relatively minor comments. Major comments. Valid criticism is made of existing work in the Introduction (P2, L9-14), and refers to how "these approaches still require a significant degree of human decision making into how to fuse disparate geoscientific datasets." And "these approaches still largely elide the question of how the joint distribution of such parameters is meant to be derived." This infers the manuscript will then address these important issues, which it barely does. These statements are then followed by another which claims the presented method "will fuse all available constraints in a probabilistically rigorous fashion." The method doesn't fuse all available constraints (see discussion, where this is admitted), in fact it only uses a small subset of available data. One major omission is structural and drillhole data, which is used or can be used in all the methods described in the papers cited in this paragraph. These claims are at best poorly made, and at worse false. Pakyuz-Charrier and Giraud both address the issues of how joint distribution of parameters (geophysics, drill holes, petrophysics) are made. A far better justification of these statements needs to be made in order to emphasise the contribution of this paper to advancing this important area of research.

AGREE. Also in light of comments made by reviewer 1, we have modified the second and last introduction paragraph significantly to incorporate the above comments. The structural data would be an excellent addition but, unfortunately, Obsidian is not capable of integrating this data. Future work will involve modifying the Obsidian code but we felt that the joint inversion of gravity, magnetics, petrophysical and lithostratigraphic data was sufficient for this contribution. Drill hole data is scant in our study area, with only a few $\sim$10 m deep holes located in the SW corner.

Comparing models results with maps. If the maps were made using interpretations from geophysics, then the model, which is based on geophysics, matching the map is not surprising, and expected, and thus not an adequate validation exercise. Please better justify the validation method.

PARTLY AGREE. The geological maps were primarily derived from geological mapping. There is some significant regolith cover in the southern portion of the region but the majority of the region is able to be mapped without the need for geophysical input.

How is the geological model built? Figure 2 implies five units are modelled, and then it's revealed deep into the discussion that only two were modelled. Differentiating between two geological units is not that exciting, not useful, especially at the scale of the study, so the authors need to show better justification as to how this method is novel, and worthy of publication.

One point of novelty we believe we have underemphasized in the previous version of our manuscript is that while the parametrization of our 3-D geological model is not very sophisticated, it incorporates very little prior information about the region in question, in comparison to more detailed parametrizations that start from a detailed geological survey. This may make our approach useful in a greenfields exploration context, perhaps to produce initial 3-D models that can be refined spatially in separate runs as more data become available. The limitations on what kinds of geological features can be represented by our choice of parametrization is something we hope to address in future work.

Two important papers that are not referred to and are very relevant to this work are: Wellmann, J. F., M. de la Varga, R. E. Murdie, K. Gessner and M. Jessell (2018). "Uncertainty estimation for a geological model of the Sandstone greenstone belt, Western

Australia – insights from integrated geological and geophysical inversion in a Bayesian inference framework." Geological Society, London, Special Publications 453(1): 41-56. Guillen, A., P. Calcagno, G. Courrioux, A. Joly and P. Ledru (2008). "Geological modelling from field data and geological knowledge: Part II. Modelling validation using gravity and magnetic data inversion." Physics of the Earth and Planetary Interiors 171(1-4): 158-169.

Both these works describe methods similar to that being described here and deserve a good review in this paper. In particular Wellman et al. 2017 presents a Bayesian framework for geophysics that authors would benefit from during their review.

AGREE. Reviewer 1 also noted that these papers were omitted. Both papers have now been cited in the text, with particular reference to the structural measurement integration covered in Wellmann, et al., 2018.

No figure shows any 3D model, either the initial, or geophysically constrained geological version, nor the inverted geophysical volume. This is a critical thing to show to the readers of Solid Earth, most of whom are geoscientists. How can we appreciate your endeavours without seeing the results, especially when "3D geological models" is in the title?

DISAGREE. Figure 11 shows two 3D models.

Downsampling of "geological" (really lithostratigraphic) observations. You have detailed "petrographic, geochemical and geochronological knowledge obtained on a subset of WAROX data" which would surely give far higher lithological resolution that the five bulk units that make your model (legend of fig. 5). How did you downsample these observations into the five major groups? As you hint, there is significant uncertainty, not just in correctly identifying the correct rock unit (though with the data you have this source of error should be reduced, but not irreducible). This alertoric uncertainty is inadequately addressed in P9, L19. How did you determine the error in these observations? How was it translated into a Beta distribution? There is also the loss of information from the process of downsampling – i.e. epistemic uncertainty (which is reducible). You refer to section 3.4 in this matter, but section 3.4 barely describes this in a geological context. Other issues with section 3.4 exist. . . next paragraph.

We are not sure what the reviewer means by this.

Section 3.4 needs significant work. It is quite disjointed from the previous section. For example, how are the survey forward models determined? Why alpha = 1 and Beta = 2? There is no effort to make the text link with the previous sections, explain the importance of a Beta distribution to a Bayesian framework, nor appropriate translation of known uncertainties within a geological context or even related to widely understood sources of uncertainty in geoscientific data. In its current form this section is incomprehensible.

AGREE. We have completely rewritten this section in the hopes of improving clarity, explaining that the alpha and beta parameters (for the inverse-gamma and beta priors on sensor noise) are elicited from experts, and that they are fairly vague priors.

The results are not presented well. Section 4.2 Residuals from forward models: Statistics are presented without any indication as to whether they are acceptable (e.g. "Aeromagnetic residuals display an approximately Gaussian distribution of 0 +358 –31725 nT (2σ, 21% of the total magnetic range" – so what?) or even higher or lower than expected.

PARTLY AGREE. Results should be as transparent as possible with as little as possible interpretation. Stating that something is acceptable or unacceptable is subjective and is appropriate in the discussion, not the results. The implications of the distribution and ranges are already discussed in section 5.1. Nevertheless, we agree that the results could be better streamlined and have made comparisons between the total range of each of the sensors to the petrophysical data to help the reader understand this link.

Section 4.3 Probability density of layer locations. Text associated with figure 9 states

that rock observations near the contact between the Halfway Gneiss and Durlacher Supersuite are misclassified. None of this is very surprising given it's a contact which any geologist knows are hard to define. But the relevance of this finding is difficult to discern given the method for building the model isn't described anywhere, the cell sizes of the model are not given (see comments below), nor how the contact was defined in the first place. All it infers (given no other information) is that Obsidian doesn't manage to determine the geometry of this contact well. This may not be true, but none of the other results presented show that Obsidian has done a good job in this regard. This is not helped with the lack of description for geological model construction.

AGREE. We have clarified these aspects in our revised text. The primary aim of our work is to provide a quantitative framework for the uncertainty in the location of the contact in our model, which makes the statement that contacts are "hard to define" more precise. Unlike more detailed parametric models, we arrive at our posterior probability distribution for the contact geometry starting only from a minimum feasible resolution – set by the discretization and the correlation scale for the interface geometry – and the knowledge or strong prior belief that a contact exists in the modeled volume. We also set out additional details of the geological model in revisions to sections 2.3 and 3.1, describing the grid of control points, the covariance kernel for the interpolating surface defining the contact, and the discretization resolution.

You state that results show the Durlacher Supersuite to be in two "domains". This isn't surprising given it is a Supersuite, and by definition made up of multiple suites, which can be defined as domains. This interpretation is also not supported by any geological data, nor is the importance of this made apparent during the introduction, discussion or conclusion. The results are described as being far more successful than they really are.

DISAGREE. The petrophysical data, which is inherently tied to the geophysical data, shows there is no spatially-controlled differences between the different domains of the Durlahcer Supersuite.

Page 14, line 13: "Highly similar" – not really. There are a quite few differences, plus you have only shown the probabilities of two units, when the 3d model was built using 5 units (maybe? Again, describe how the model was built). What about the other three units? The assessment of this method is thus inadequate. As such much of the discussion unconvincing, especially when two select slices of the probabilistic model are shown.

AGREE. We agree that parts of the methods section were ambiguous and we have now made these more transparent. Only the 2 most voluminous units were modelled as the other 3 units are volumetrically and areally minor. We have also changed some of the figures of Figure 11 (3D model) to make it clearer where regions of uncertainty are.

You admit that structural data is not used in the discussion (P14, L29). This needs to be stated clearly in the method (where a description of model construction is required – see previous comments) and makes earlier comments criticizing previous work disingenuous (see major comments). It is self-evident that structural data is very useful for geological models. The use of structural data is shown in other methods that have been around for almost a decade (see uncertainty work by Wellmann, de la Varga, Bond, Lark, Lindsay, Jessell), or general modelling (see Calcagno paper). Why can you not do this? The same can be said for drillhole data, which other methods also use. The main problem is that both structural and drill hole data from the area is publically available, but not used. So it appears that Obsidian, or the described method cannot use these data presently, otherwise they would have done so. Other methods (as cited earlier) can use both structural, seismic and drillhole. How do the authors then justify this method as novel, or one people should adopt given is has severe limitations to inputs? Simply being Bayesian is not enough, especially as Bayesian methods are well suited to integration of different data types. Other Bayesian techniques have been proposed (de la Varga, Wellmann). This aspect needs a much fuller justification and discussion.

PARTLY AGREE. This has now been made upfront in the final section of the introduction instead of in the methods: "There are a few datasets available in the region that are not utilized in our model. (1) There are only a few $\sim$10 m-deep drill holes in the southwestern corner, so drill hole data is omitted as it does not add further detail than surface observations provide. With a lack of drill hole data, our contribution is able to address the impact of solely surficial geological data on the model accuracy. Applications such as greenfields mineral exploration or tectonic analysis of hard rock terranes without drillholes would benefit from this understanding. (2) Our study excludes the use of structural geological data because other workers have recently focussed on this problem (e.g., Pakyuz-Charrier et al., 2018b; Wellmann et al., 2018) and because Obsidian cannot yet incorporate structural data. (3) The single 2D active seismic line immediately to the west of our model (Fig. 1) is not utilized in a Bayesian framework because the vast majority of hard rock terranes do not have seismic data coverage."

Minor Comments Page 3, line 7: define "data-rich". Rich in diversity or coverage, both? Quantify this richness.

AGREE. This has been changed to: "We demonstrate the validity of our techniques by building models of a 13.5 $\times$ 13.5 km subsection of the Gascoyne Province, Western Australia (Fig. 1), that is rich in data diversity and coverage"

Page 4, line 11: technically measurements of gravitational acceleration and magnetic field strength

AGREE. This has been changed.

Page 4, line 16 – Be clear about the shortcomings of other work. Giraud et al. in review does acknowledge alternative geophysical, petrophysical and geological sce- narios. Are you referring to alternative forms of parameterisation for regularization?

AGREE. Also in line with reviewer 1's comments, we have changed this to:

"Ways to introduce such constraints include 3D geometry inversion (Fullagar et al., 2008; Guillen et al., 2008), level-set inversions (Bijani et al., 2017; Zheglova et al., 2018) and (cross-)gradient regularization (Giraud et al., 2019; Scholl et al., 2016). However, these techniques are deterministic, yielding a single geological-geophysical inverse model that represents only one scenario."

Page 4, line 27: it is unclear what you mean by "geophysical processes". Are you talking about how well the models represent the geology?

AGREE. It was unclear to us exactly what we meant with "geophysical processes" as well. Apologies about that. We have now changed the text to:

Page 6, line 18: "PTMCMC" define your acronyms before using them.

AGREE. This has now been defined in the last paragraph of section 2.2.

Subheading 3.1. "World" is an expansive term that infers all parameters, data, models, inferences, assumptions are under consideration in the following paragraph, which isn't true. "3D geological model parameterization" is more specific and less confusing. The same applies to all references of "world" models. This is important as you use and describe more than one model through the manuscript, including statistical, geophysical and conceptual are present as well.

AGREE. The title of subsection 3.1 has been changed to "3D geological model parameterization" and any "world" text replaced by "3D geological model" in the rest of the manuscript.

Page 6 line 30: should be "magnetic susceptibility and density data"

AGREE. This has been changed.

Section 3.2: You need to show where these petrophysical data were acquired on a map (Fig. 2?). Presumably the petrophysical data locations will be different to the "surface observations" shown in Fig. 2. Given the caption describes them as geological surface observations

AGREE. These have now been shown in Fig. 1.

Page 7, line 21. Please describe what Bayesian "fusion" is, or just say the data were input to the Obsidian framework.

AGREE. This was confusing and has now been updated to: "These types of geophysical surveys were already available for incorporation in the Obsidian framework".

Page 7, line 31: explain the source of these "correlations", what they are correlated with, and why this could be a problem. You do this later with the gravity data (P8,L10), so move that explanation here. But you still need to better explain the source of the biases and how they produce incorrect results in context of the Bayesian methods you describe earlier.

AGREE. The section following the gravity data has now been moved up.

Page 9, Line 23 PT-MCMC or PTMCMC (as P6,L18)

AGREE. All is now referred to as PT-MCMC.

Page 11, Line 18 Discussion of large Gweke scores

AGREE. We have now discussed this in the discussion (not in the results)

Page 11, Line 21 Figure 8 needs to show the measured interpolated image with the forward models for easy comparison, rather than forcing the reader to switch between figures on different pages.

Figures 8a and d already do this as stated in the caption: "Modelled mean contours of (a) Bouguer anomaly and (b) magnetic intensity compared to interpolated mean colored data."

The reader is also referred to figure 2 when describing magnetic lineaments, but figure 2 is a geological map. Are the authors assuming the NW strike of the geology will also produce NW striking magnetic lineaments? This is a reasonable interpretation, but the authors need to first make that interpretation for the statement in line Page 11, Line 22.

AGREE. This was poorly worded. It now reads: "Aeromagnetic models effectively identify the NW–SE strike of magnetic lineaments in the northern half of the modelled volume (Fig. 8) that would be predicted from geological maps (Fig. 2)."

Page 11, Line 25 – okay, but so what?

AGREE. This paragraph has been overhauled to make it clearer for the reader. We also discuss the implications of these ranges more fully in section 5.1:

"Aeromagnetic residuals display an approximately Gaussian distribution with a mean of 0 nT (i.e., equivalent to measured aeromagnetic data) and a range of "+358" Âę"–317" nT (2σ), which covers approximately 21% of the total magnetic range (Fig. 8). The $\sim$21% residual standard deviation is comparable to the standard deviation range of magnetic susceptibility values (Fig. 4). Only one region in the northwestern portion of the map has significantly higher magnetic field strength than modelled (Fig. 8)."

Page 11, Figure 8 caption. Labelling of figure parts appears to be incorrect. b) shows units in mGal, so not magnetic intensity, c) shows a histogram, not the modelled contours

AGREE. We thank the reviewer for identifying this mistake. The lettering in the legend has been corrected.

Page 12, Line 14. Distances have no meaning without telling us what the model cell sizes are first. How many cells does 300-1000m represent?

AGREE. The model cell sizes are 500 m, as now stated in section 3.1.

Page 12, Line 14. "The ellipsoidal Durlacher Supersuite inlier is heterogeneously constrained." What does this mean and why is it relevant?

AGREE. This was difficult to follow. We have changed it to: "The uncertainty on the boundary between the ellipsoidal Durlacher Supersuite inlier and the Halfway Gneiss

is constrained differently in different parts of the model." This follows better to the following sentences.

Page 13, Line 3 "a function of a long-wavelength (i.e. deep) gravity response" Careful here. A long wavelength is not always deep. It can be laterally extensive but shallow.

AGREE. To avoid ambiguity, this sentence has been changed to: "This is primarily a function of a long-wavelength gravity response that is probably attained from the deep subsurface (Johnson et al., 2013)."

Page 13, Line 6. I would think more petrophysical data from "other geological units" (see previous sentence) would be more useful to define the lithostratigraphic diversity than marginally tightening the standard deviation of the Halfway Gneiss and Durlacher SS units.

PARTLY AGREE. More petrophysical data from all lithostratigraphic units in the Gascoyne Province would be beneficial (see Fig. 1). There are only a few on the periphery of our model currently. In any case, more petrophysical data from the two volumetrically-major units in our modelled region would have improved the local-scale model, which is an important point for mineral explorers.

Page 13, Line 11-20 Annotate the figures with the various features being described here (Chalba SZ, Durlacher SS 'spur' and 'sliver' etc.)

AGREE. This has been added to Fig. 10a. We also considered adding it to Fig. 2a, but Fig. 2a is already relatively cluttered.

Page 13, line 19-20. Tells me the initial 3D model is wrong.

While we acknowledge that our parametrization is limited, we should also point out that there is no "initial" 3-D model in the way the reviewer would understand it. Instead we use some rather restrictive assumptions about the form a contact would take (the square exponential covariance for a Gaussian process defining the interface, and the control point geometry) but an essentially random initial guess for its geometry. While most methods thus start close to a known or believed model (or combination of parameters), our sampling process explores all models of the specified parametric form that are compatible with the data. The chosen parametric form has substantial limitations discussed in our text, but appears to be good enough to capture the overall contact geometry on scales larger than the discretization resolution. Page 13, line 31. “Importantly, our method is the only technique that provides a range of solutions and quantitatively accounts for all the input assumptions” This grandiose statement needs far more justification. I actually think this should be removed entirely, given the technique is poorly described in the first place.

AGREE. We have removed this statement.

Page 14, line 1-9. Figure 10d? Over-interpreted results. “Definitively separated” Plus given the sections only extend to 4km, how can you be sure the Durlacher remains separated beyond that? Figure 10 d looks like the Halfway Gneiss only extends as far as 4km depth (which is also probably a function of the model volume parameters? Also needs discussing). You then state correctly in later lines (lines 4-5) that “it was difficult to know whether this spur of Halfway Gneiss between the two Durlacher Supersuite domains continued at depth or was truncated in the near subsurface”. Hardly definitive!

AGREE. The earlier statement is misleading and has been changed to: “. . .is separated into two domains at the surface and shallow subsurface”

“This important contribution shows that small geological volumes on the scale of a few km can be resolved accurately and will be important when this modelling output is up-scaled to larger regions.” Small volumes can be detected given appropriate geophysical data resolution and corresponding model parameters. Small volumes have also been detected by many other methods which I suggest you spend some time reviewing (Li and Oldenburg papers, Peter Fullagar, Guillen, etc etc) so you realise this is not a world first. Upscaling models to larger regions is also commonplace. If you are to make this kind of statement, please explain how upscaling should be done.

AGREE. We have removed reference to upscaling here. We talk about upscaling in more detail in section 5.2, paragraph 3.

Page 14, line 29. How are drill holes going to help Bayesian methods >4km when they rarely extend beyond 400m?

AGREE. We have clarified this to “petroleum wells”, which do go down to as deep as $\sim$5 km.

Page 15, Line 22. Interesting concept, and I agree should be done, but expand on how this would be achieved?

PARTLY AGREE. At this stage, we’re not sure how it should be done yet.

Please also note the supplement to this comment:
https://www.solid-earth-discuss.net/se-2019-4/se-2019-4-AC2-supplement.zip

Interactive comment on Solid Earth Discuss., https://doi.org/10.5194/se-2019-4, 2019.