Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2018-306-RC2, 2019

Geoscientific Model Development Discussions

Interactive comment on "Efficiency and robustness in Monte Carlo sampling of 3-D geophysical inversions with Obsidian v0.1.2: Setting up for success" *by* Richard Scalzo et al.

Anonymous Referee #2

Received and published: 18 April 2019

GENERAL COMMENTS:

This study explores the influence of various practitioner decisions on MCMC posterior sampler efficiency for a geophysical joint inversion with a layered paramitrization; specifically, the influence several of proposal, prior, and likelihood function options. The tests are well designed and succeed in addressing the questions asked. I personally did not find much of the results and conclusions surprising, most of it could be deduced from purely theoretical grounds. However, the topic is important and this paper gives a good empirical basis from which future geophysical posterior sampling work can draw. On these grounds, I think it deserves to be published.

Most of the paper is well written and clear, with the introduction being the exception. It seems rushed and the odd use of Bayesian/statistical/probabilistic terminology (in the introduction only) suggests a lack of familiarity.

I personally don't see the need for additional detail on the software use and implementation since that is not what the paper is about and if anything the scope should be more contained not expanded.

The review of the MCMC literature review is extensive and was interesting to read. Complexity of the shape of the posterior is discussed several times but seldom in the context of previous work. For example, the non-linearity and complex correlations of physical parameters for non-unique magnetotelluric inversion is well known, but no overview is given here on that. I believe the discussion section could be improved by relating more to the known properties of the different geophysical forward problems.

I think the discussion section is needlessly bloated. Here the authors go into detail on many topics which the experiments shown here had no bearing on. Various things that could be done or might work are listed here which are in no way related to what the study presented actually did. I strongly suggest rewriting this section to be more on topic.

SPECIFIC COMMENTS:

P1L12) What does "improve inversion results" refer to? Most readers would assume that it means a more accurate inversion. Since accuracy of results,compared to reality, is never quantified in this work, I don't see how this claim is backed up. One might argue that if true sensor noise levels are known, uninformative priors on them would only increase chances of their miss-estimation. Counterarguments based on model inadequacy could be raised of course, but these are not things that this study shed light on so please remove this claim.

P1L13-15) I don not see why this claim about using gradient information is in the ab-

stract. The statement is probably true, but this study did not show anything new to support it.

P2L25) It's not clear what is meant by posterior ensembles being a 'gold-standard'.

P2L29) Online updating is not necessary or sufficient for optimal for decision-making; these are separate things. The only relevance I can see here is that it could speed up decision making.

P3L3) What about overestimating uncertainties?

P3L4) The "no 'one-size-fits-all' solution exists" comment is very important. Perhaps give the reader some direction by citing something (e.g. Wolpert et al., 1997, No free lunch theorems for optimization: IEEE transactions on evolutionary computation, 1, 67-82.)

P4L21-29) This paragraph should probably lead with the last sentence (lines 27-29). The parts about deterministic inversion reads like an odd tangent and I didn't see the relevance and purpose of it until a second read through.

P7) I did not pick up on the fact that all your tests use PTMCMC until the second read-through; this section should probably make that more explicit.

P12L1-5) Where the seismic lines used to inform the layer interface Gaussian process variogram?

P14L8) What is 'global posterior shape'? is it always defined?

P14) Could you have used the layer Gaussian process covariances directly to create a proposal function, it seems like that is what the CNp effectively does?

P16L4) "Figure 2 shows that iGRW and aGRW have more trouble travelling between different posterior modes than pCN" Tell us how you deduce this from that figure.

P17) Please mention why MT noise levels were fixed.

P21) Increasing the amount of data points by interpolation seems like a terrible idea. Why would anyone even attempt it? The observed effects should be obvious. If there are actual examples of Bayesian posterior analysis papers which do this, please cite one; otherwise, this seems like an odd and unnecessary test to include.

P21L16-19) I don't know about gravity and and magnetic, but Gaussian process likelihood functions have been used for MT and seismic MCMC. Relevant work should be cited here, E.g.:

Agostinetti, N. P., and A. Malinverno, 2010, Receiver function inversion by trans-dimensional Monte Carlo sampling: Geophysical Journal International, 181, 858–872.

Bodin, T., M. Sambridge, H. Tkalčić, P. Arroucau, K. Gallagher, and N. Rawlinson, 2012, Transdimensional inversion of receiver functions and surface wave dispersion: Journal of Geophysical Research: Solid Earth, 117.

Xiang, E., R. Guo, S. E. Dosso, J. Liu, H. Dong, and Z. Ren, 2018, Efficient hierarchical trans-dimensional Bayesian inversion of magnetotelluric data: Geophysical Journal International, 213, 1751–1767.

Also, there are ways to learn the correlation during sampling:

Steininger, G., J. Dettmer, S. E. Dosso, and C. W. Holland, 2013, Trans-dimensional joint inversion of seabed scattering and reflection data: The Journal of the Acoustical Society of America, 133, 1347–1357.

P22L19) "The clearest lesson we can draw ..." I'm not sure why this is the lesson you lead with, in the introduction it was stated as known; almost every MCMC application to geophysics show this and it was not among the questions that your tests were set up to answer.

P22L20) "Our results were sensitive to ..." Each point raised in this sentence will be true for for any difficult posterior sampling problem. This is not a new result and this sentence adds nothing to the manuscript.

P22L34) I don't agree with the claim that either of these outcomes are counter-intuitive. Tighter constraints lead to narrower local optima, hence more sampling is needed. Cauchy likelihood functions are more likely to give multi-modal posteriors than Gaussian likelihood functions, even for the most trivial problems (e.g. with just one parameter).

P23L6-10) None of these three dot-point listed statements were informed by the experiments presented in this manuscript. The claims are also obvious and well known.

P24L10) "However, proposals using gradients from auto-differentiation are probably needed to make further progress in this area." This claim, while probably true, is not really backed up by what is in the manuscript. Why is it listed as a conclusion?

P24L14) This is a trivial claim by itself. How can it help design future work. Will the better fit derived from uninformative priors lead to more accurate results in terms of uncertainty estimation. This ties in with my comment for P1L12.

P2417) Without some guiding principle for how to do the sub-sampling, this is not useful.

TECHNICAL CORRECTIONS:

P2L20) "..., but about uncertainties." Awkward use of terminology, a Bayesian probability is an uncertainty and an assumption. Assumptions are specified as uncertainties quantified by probability distributions.

P2L26) "The posterior distribution is a representation of all possible outcomes and hence provides an internal estimate of uncertainty." The world parameterization is the representation of all 'possible' outcomes. What does the 'internal estimate' mean? This sentence is incoherent.

P6L12) Spell out what SGR stands for here.

P6L26-27) Grammar mistake.

P13L1) Table 1, what is N? First it was iteration count, then number of layers, then what? Readers shouldn't have to fish through the past 12 pages to find out.

P18L4) Grammar mistake.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-306, 2019.