Solid Earth Discuss.,
https://doi.org/10.5194/se-2020-21-RC1, 2020


# *Interactive comment on* "Uncertainty assessment for 3D geologic modeling of fault zones based on geologic inputs and prior knowledge" *by* Ashton Krajnovich et al.

**Alexander Schaaf (Referee)**

a.schaaf@abdn.ac.uk

Received and published: 14 April 2020

## 0.1 General comments

The paper presents a step forward in the uncertainty-aware modeling of subsurface faults in structural geomodels. The authors make use of Monte Carlo simulations to simulate uncertainty of fault zones based on a specific fault zone parameterization (surface traces, vertical termination surfaces, structural orientation and fault zone thickness) using a proprietary software suite. The authors elaborate the use of anisotropic spherical distributions for parameterizing orientation data for uncertainty simulation,



which is a valuable contribution.

The manuscript is overall well structured, except for a few re-arrangements necessary to increase readability (detailed in the specific comments). The authors give proper credit to related work and clearly indicate their own contribution. The title clearly reflects the contents of the paper. The figures presented will require some work to improve legibility and to avoid confusion of the reader.

But the authors appear to be confusing their simulation approach: They introduce MCUP (i.e. Monte Carlo simulation) in the methodology and properly parameterize their stochastic geomodel using probability distributions. But they then erroneously describe that they use Markov Chain Monte Carlo (MCMC) sampling. MCMC sampling is used for exploring the *posterior* space, which does not exist in a Monte Carlo simulation (i.e. MCUP). As the probability space is known in a Monte Carlo simulation, it needs no exploration. In a Monte Carlo simulation we only don't know how the combination of samples effect the output of the simulator function (the geomodeling software), thus we randomly sample (Monte Carlo sampling) from the parameter distributions to create a geomodel ensemble that shows us how the uncertainty in the input parameters effects the geomodel output. Luckily, to my knowledge, the used probabilistic programming framework *pymc3* defaults to Monte Carlo sampling when no likelihood function is given (and thus no Bayesian inference can be conducted). Thus the authors appear to have accidentally conducted the simulations they wanted to do (MCUP/Monte Carlo sampling). The use of trace plots (as in Figure 6 and 8) for Monte Carlo simulation results is meaningless though (and potentially misleading), as no sampler is being used that requires determination of convergence. As, luckily, the presented simulation results appear to be valid MCUP results, the authors only need to change their writing accordingly, without the need for re-running simulations.

In its current state, mixing up the terminology of Bayesian inference and MCUP, I can not recommend the paper to be accepted. But if the authors fix their method descriptions and discussions of the results to fit the MCUP simulations they actually conducted, I believe this could become a valid scientific contribution that is worth publishing in Solid Earth.

## 0.2 Specific comments

### 0.2.1 1 - Introduction

The Introduction of the paper needs to clearly state the scope of the study / hypothesis to be tested or explored.

### 0.2.2 2 - Model implementation

L52 - Both paragraph (lines 52-68) need to incorporated into the introduction as they define the scope and motivation of the study.

### 0.2.3 3 - Probability distributions for MCUP

L138 - How do you evaluate the likelihood of the proposal step in a Markov Chain during MCUP? This is only possible when doing a Bayesian inference, not a Monte Carlo uncertainty propagation, as you don't have any likelihood function.

Overall the description of simulation/sampling should be moved into Section 3.2.

L140 - The paper de la Varga & Wellmann (2016) uses *pymc* to conduct a Bayesian inference - thus not MCUP.

L141 - *pymc3* has not been implemented into *GemPy*, but rather *GemPy* in implemented in *theano*, which is also used in *pymc3*. Thus *GemPy* integrates seamlessly with *pymc3*, providing the gradients necessary for advanced gradient-based sampling



techniques such as Hamiltonian Monte Carlo.

### 0.2.4 3.2 - Simulation

A more adequate name for the section would be “Sampling”.

### 0.2.5 3.3 - Rotation

This section is part of sampling and should be merged into Section 3.2

### 0.2.6 4.2 - Surface trace

L291 - It is unclear to me what the “approximate geographical error of known landmarks” is.

### 0.2.7 4.3 - Vertical termination depth

L312 - What is a deterministic distribution? Do you mean a derived distribution? Or an empirically parametrized distribution? A distribution should be by definitiv non-deterministic.

### 0.2.8 4.5 - Simulation quality assessment

L334 - Without a likelihood function you can't use a MCMC sampler, as you are unable to evaluate the step proposals.

L336 - In an MCUP simulation, you have no posterior uncertainty space, as you are not using any likelihood function. You are mixing up terminology of MCUP and Bayesian inference. Again, MCMC sampling is only possible with a likelihood function (thus not in MCUP).

#### 0.2.9 6.2 - Model parametrization

L388 - The use of "posterior distribution" is false, as you are doing MCUP, not a Bayesian inference.

#### 0.2.10 6.3 - Parameter relationships

L411 - The meaning of the entire paragraph is unclear to me and needs to be revised.

L419 - Gibbs sampling is not applicable to MCUP, as no likelihood is used.

#### 0.2.11 Figures

#### 0.2.12 Figure 2

The dotted volume texture makes annotations for uncertainty extremely hard to read. The same goes for there fault zone signature/texture. I'd highly recommend removing as much texture as possible from the plot to improve legibility.

"The Visual Display of Quantitative Information" by Edward Tufte provides ample of additional reasons for reducing distracting "ink" from scientific visualizations and is well worth a read :-)



#### 0.2.13 Figure 4

Highlighting of fault traces is really difficult to see. I highly recommend making this figure more legible to the reader by removing visual complexity: e.g. remove coloring of geological map in the background.

#### 0.2.14 Figure 5

- legend is barely legible - please increase text size
- entropy plot of the fault zone thickness barely shows any uncertainty. If your discretization is not fine enough to resolve the simulated uncertainties, then is it worth incorporating into you model?

#### 0.2.15 Figure 6

The use of trace plots is only useful if evaluating convergence of (e.g.) Markov chains. MCUP uses Monte Carlo simulation, thus the use of trace plots serves no purpose and is confusing. Also the rug plot on the left size shows the same information as the histogram of the vertical termination depth in the lower right. I'd recommend just using the histogram to demonstrate that you've sampled enough samples.

#### 0.2.16 Figure 8

Same as for Figure 6.

---

Interactive comment on Solid Earth Discuss., https://doi.org/10.5194/se-2020-21, 2020.