Defend the conclusions despite the large posterior uncertainty.

Thomas Wutzler (Referee)

Referee comment on "Parameter uncertainty dominates C-cycle forecast errors over most of Brazil for the 21st century" by Thomas Luke Smallman et al., Earth Syst. Dynam. Discuss., https://doi.org/10.5194/esd-2021-17-RC1, 2021

The study of Smallmann et al. presents a model-data integration study where a suite of terrestrial ecosystem models of increasing complexity is inverted and evaluated using spatially resolved data across Brazil. The finding that already with quite simple models, parameter uncertainty is more important than model structural uncertainty and uncertainty in forcing data has a large impact of the earth system science and is worth to be published.

I enjoyed reading the beginning of the paper and appreciated the well designed study using multi-model, multi-biome, site-specific spatially resolved setup, and varying input data for future climate scenarios, within a fully Bayesian inversion setting.

However, I got disappointed when I more closely inspected Table 2. Even with the simplest model, the confidence intervals of the predictions are so large, that only vague and general statements or conclusions can be drawn from the results. All the elaborations and conclusions on model complexity and structural error would be a really good presentations, if the results were more constrained. However, with the large uncertainties I would recommend to only summarize them and omit the detailed presentation, because they are base on vague ground. Instead of studying increasing model complexity, the results show that the data is already not enough to constrain the simplest model variant. Statements like LL259 "simulated NEE was consistent with CTE ensemble at the 90% CI" does not tell me much about the goodness of the model, if the CI range is 400% of the median estimate. In order to defend the insights despite the large uncertainty, the posterior density of the parameters in comparison to the priors should be provided as a supplement or appendix.

The main conclusion about the dominance of parameter uncertainty is strengthened by this large uncertainty and should be published, with a much shortened presentation of the (to my opinion vague) comparison across model structures.

An alternative route, which requires a larger reanalysis effort, is based on the claim of the authors that the model can be constrained by repeated EO observations of biomass. In addition to the current model inversion, I suggest generating an artificial observation of this biomass data stream using the most complex model variant add noise and some slowly-changing bias and repeat the inversion including this artificial data. If the uncertainties decrease as much, the presentation about model structure could be kept, but based on this new (artificially) more constrained inversion results.

Specific comments

To gain an conception about the computational effort: At how many pixels was the model inverted?

Line 220ff: It did understand how "future climate is imposed by determining the anomaly from the end of the analysis until 2100". Please, extend this explanation.

Line 223: I assume the model structural uncertainty was estimated for each climate scenario separately (and the climate uncertainty for each model variant separately), right? Or does the "between model range" span across all climate scenarios?

Fig 4: In my opinion, the stippling (indicating consistency within confidence range) does not tell much when considering the large uncertainties.

Fig 5: putting all the labels the center panel confused me first, I suggest putting the observation legends to the panels. Almost all the streams are encoded by color, which for me were difficult to read.

L286, 294: Why is the NEE not improved with model complexity, if fire is improved and makes up 3 to 30% of NEE?

L315: How did you assign priors to the MRT parameter? The sentence suggests a Normal distribution that includes also negative values. A lognormal prior would be more appropriate and you could report the multiplicative moments of the posterior and avoid negative residence times.

Sec 3.3. reads lengthy. Are all the details necessary in the main text. I have, though, no specific suggestion how to shorten.

L331: Hints to model error. Thanks for the discussion at L462ff, that could be referenced at this point. For me it did not become clear, how biomass removal was accounted for in the DALEC simulations, and the future scenarios.

L 347: Can you quantify "most likely"? Can you infer $p(deltaBiomass(t) > 0)$ from the posterior?

L 450: may replace "a function of three factors" by "There are three possible interacting explanations"