Comment on gmd-2021-273

Anonymous Referee #1

Referee comment on "Calibrating the soil organic carbon model Yasso20 with multiple datasets" by Toni Viskari et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-273-RC1, 2021

General comments:

- I am not sure about the significance of this work. This study proposed two hypotheses and performed some numerical experiments to test the hypotheses. However, as to me these two hypotheses are more like facts. It is well-known that advanced parameter optimization methods (e.g., global parameter optimization methods, Markov chain Monte Carlo methods that can sample multi-modal distributions) perform better than the simpler methods (e.g., local parameter optimization approaches and the MCMC methods that cannot sample multi-modal distributions such as the AM method used in this work) in finding likelihood maxima. Additionally, it is also known that the models calibrated against multiple datasets perform better than those calibrated with a single dataset as long as the uncertain parameters are sensitive to the calibration data and the calibration data provide information to constrain the uncertain parameters.
- I am not sure about the novelty of this work. I do not see new methodology development or new insights.
- I do not think the model calibration performed in this work was completely reasonable. First, a global sensitivity analysis is necessary before the parameter estimation, especially when the multiple datasets are used for calibration. Secondly, when the multiple datasets have dramatically different size, a weighted likelihood function should be constructed to consider the imbalanced data contribution. I am not sure whether the authors consider this point.

Specific comments:

- Line 85-86, the authors mentioned that they present a new model formulation and a calibration protocol. Please elaborate what is the "new" here, and please explain what the authors meant "calibration protocol".
- Line 88 two "not only". Line 92, maximas-->maxima.

- Line 91-95, I do not think these two are hypothesis, they are more like facts to me. It is well-known that advanced parameter optimization methods (e.g., global parameter optimization methods, Markov chain Monte Carlo methods that can sample multi-modal distributions) perform better than the simpler methods in finding likelihood maxima. Additionally, it is also known that the models calibrated against multiple datasets perform better than those calibrated with a single dataset as long as the uncertain parameters are sensitive to the calibration data and the calibration data provide information to constrain the uncertain parameters.
- Line 433-435, I assume the authors gave all the calibration data equal weights; the imbalanced data size may dilute the influence of Hobbie3 on the parameter estimation. How about giving different datasets different weights in model calibration?
- Subsection of impact of prior parameter information: in Bayesian calibration, the prior parameter distribution matters, not only the distribution type but also the prior parameter range. I think a global sensitivity analysis is necessary before the model calibration. At different portion of the parameter space, the parameters may have different sensitivity to the calibration data.