Responses to comments provided by the referee C.A. Sierra

This manuscript presents an evaluation of the Yasso07 model of soil organic matter dynamics. The original Yasso07 model was parameterized with litter decomposition data from North America and Europe with an important underrepresentation of tropical regions. This manuscript is important because it evaluates the performance of Yasso07 predicting data from a litter decomposition experiment from a tropical forest. The authors do a good job reparameterizing the model to include the new data; however new questions arise regarding the identifiability of the parameter set of Yasso07 using only litter decomposition data (see below).

Authors' response: Noted

Main comments

My main concern with this analysis is the large differences between the different sets of parameters among Y07, Y07A and Y07B. Table 2 shows the parameter values for these sets. It is particularly intriguing the large differences among the _i parameters. For example, the parameter _A is 5 times larger in Y07B than the original Y07, but in Y07A this parameter is 0.5 times lower than the original. Similar large differences are presented for other parameters. Why are the differences so large and why are not the values of Y07A somewhere in between Y07 and Y07B? To me, this is an indication of poor identifiability of the parameters given the available information.
I think this issue should be better discussed in the manuscript. One option would be to show the posterior frequency histograms alongside the scatter plots of the parameters of the model runs. Similarly, the correlation coefficients among the parameters may give some insight on possible identifiability issues. Depending on the software you use, you may be able to perform a full identifiability analysis.
In R, this can be easily done with package FME. Unfortunately, I am not aware of similar procedures implemented in Python or other languages.
In any case, I think the authors should address better this potential issue of the identifiability of the parameter set of Yasso07 using litter decomposition data. A good reference for this topic is Brun et al. (2001, Water Resour. Res. 37: 1015-1030).

Authors' response:
We agree that these differences appear striking. However, it is not the parameter values as such that matter, but the decomposition rate factors ki, for A, W, E, and N, defined in Eq. (2), that are crucial in controlling the decomposition rate and thus model predictions.
Calculating these factors based on the estimates in Table 2 shows that these rates are consistent for the Y07A and the original Y07 given the uncertainties of the parameter estimates. It thus remains to be explained why the parameter estimates of Y07B are significantly different. The simple explanation for the different values of the Y07B estimates is that they are based on data from Benin only, which means that the very narrow range of climate conditions effectively prevented us from determining the dependence of mass loss on climate. Moreover, the effects of leaching and mass-loss based on decomposition have become mixed, which can be seen in the Table 2 when looking at the parameter \omega_B. Given all the data, this leaching appeared to explain most of the mass-loss but given the Beninese data alone, the interpretation was instead that the mass-loss is caused dominantly by decomposition. With a narrow scale of climate conditions (Table 1), namely precipitation that affects both decomposition rate and the leaching, it is difficult to differentiate between these two effects.

We also note that the uncertainties of the parameter estimates of the Y07B are much greater (typically a factor of 10) than those of Y07A and the original Y07 because much less data was available for parameter estimation with Y07B. The differences are thus not necessarily an indication of poor identifiability and thus poor modelling.

Furthermore, studying "identifiability", i.e. whether the model is consistently accurate in such modelling problems, is beyond the current study and deserves a paper of its own.

Changes in the text: in line 23, p. 3013, text changed to:

"The analysis indicated large differences between the different sets of parameter values among Y07, Y07A and Y07B (Table 2). The parameter values αi were larger in Y07B than the original Y07, but in Y07A the differences were smaller compared with Y07. Similar differences were observed for other parameter values. As a result, the decomposition rate factors of chemical fractions (A, W, E, and N) were substantially higher when determined from Benin data only, as defined in Eq.(2). …"

In line 19, p.3015, text changed to:

"It is not the differences in parameter values (Table 2) as such that matter, but the decomposition rate factors ki, for A, W, E, and N, defined in Eq. (2), that are crucial in controlling the decomposition rate and thus model predictions. Calculating these factors based on the estimates in Table 2 shows that these rates are consistent for Y07A and the original Y07 given the uncertainties of the parameter values. It thus remains to be explained why the parameter values of Y07B are significantly different. The simple explanation for the different values of the Y07B estimates is that they are based on data from Benin only, which means that the very narrow range of climate conditions effectively prevented us from determining the dependence of mass loss on climate. Also, the uncertainties of the parameter estimates of Y07B are much greater than those of Y07A and the original Y07 because much less data was available for parameter estimation with Y07B. Although the level of decomposition predicted with Y07B was higher in Benin, the major pathways of the decomposition process (mass flow between chemical compound groups) were rather similar for Y07A and Y07B."

Comment: One important question that emerges after reading this manuscript is whether the new parameterization proposed with Y07A significantly changes the predictions obtained with the original Y07 parameterization? Do this new parameterization drastically change previous predictions with Y07? I think a discussion on this subject would be pertinent for this manuscript.

Authors' response: the issue of comparison of predictions obtained with Y07 and Y07A are addressed in the result section, page 3012 (from line 17) to page 3013 (line 15). The following text will be added in the discussion section: page 3015, line 19:

"This is also confirmed by the large differences observed between the predictions of Y07 and that of Y07A."

Minor and technical comments

The description of the methods for inverse parameter estimation using MCMC needs to mention the criterion to assess convergence of the chains. Also in the results, what proportion of chains were accepted by the algorithm? Can you show some convergence diagnostics?

Authors' response: text to add to section 2.3 data analysis (page 3012, line 11)
"We assessed the convergence of the Markov chains by checking the Gelman-Rubin diagnostics (Gelman & rubin, 1992) that compares the variance of the parameters within chains to that between chains. Although we only calculated three chains for each case, as the computational requirements of such samplings were rather high, we could easily see that the chains had consistently the same stationary distributions enabling us to state that the chains were indeed sufficiently close to

convergence. This was also visually clear as the chains did not jump between different states but stayed around the MAP estimate at all times, which is indicative of good mixing properties of the chains."
Our samplings had typically acceptance rates between 0.2-0.3, which, we believe, is what the ref. #1 means by the question "what proportion of chains were accepted". This is a typical acceptance rate for the adaptive Metropolis algorithm in a case where the posterior is unimodal and resembles multimodal Gaussian density.

Gelman, A. and Rubin, D. B. 1992. Inference from iterative simulation using multiple sequences (with discussion). Statistical Science, 7:457-511.

Attached the posterior densities of the model parameters for the global Y07, although we do not consider them very useful (especially as they are a bit crap plots) when the basic information has already been presented in Table 2. Unfortunately we don't seem to have correlation plots available, although there would be more than 400 of them for the 22 parameters in Table 2, so we do not fully understand why they are very necessary. We think even showing the 22 distributions in the paper would be simply waste of space, not to mention showing them for all three solutions.

Pg. 3005, ln. 17-18. Please reword sentence. It's not clear what are you trying to say here.

Authors' response: text changed to:
"Overall changes in the SOC stock may be quantified with measurements, but repeated measurements of soil carbon stocks are laborious, and time and resources consuming making difficult the quantification based only on measurements."

Table 2. Please add Y07B, Y07A, and Y07 to the column names after Benin data, All data, and Original calibration, respectively.

Authors' response: accepted.
Change in Table 2:
"Benin data (Y07B), All data (Y07A), Original calibration (Y07)."

The use of the English language can be improved in many sentences. Please review the next version carefully.

Authors' response: the comment will be implemented.