This document contains referee comments to bg-2021-327 manuscript: ‘Implementation and initial calibration of carbon-13 soil organic matter decomposition in Yasso model’ and details on the modifications we have made to the manuscript.

Referee 1 comments

This manuscript describes new stable carbon isotope capabilities added to the Yasso model. The new model capabilities are described clearly. The model updates were parameterized and evaluated using measured datasets in a way that was well described and justified. Overall, I though the manuscript was a clear and concise description of a valuable new model capability. 13C measurements are a common metric for understanding soil organic matter decomposition processes and adding this capability to a SOM model is a valuable advance.

I did think that in some areas the introduction and conclusions went beyond the scope of the actual results. Specifically, the model developments and testing were entirely focused on 13C fractionation and did not include changes to or evaluation of overall soil C decomposition rates. Therefore, the hypothesis in the introduction about “significant improvements in SOM decomposition predictions” seems broader than is justified. The study does yield improvements in predictions of 13C dynamics, but this was not used to improve overall SOM predictions.

This is a valid critique. We have reformulated this part of the text in the introduction as “allows model development for future improvements in SOM decomposition predictions.” And we have added more discussion on the different avenues of how the modified model could be used in the future. We view this work more as a proof-of-concept – examination of actual SOM decomposition predictions would require a different approach, and this goes beyond the scope we have set for this paper. The changes to actual SOM content values, induced by the model modifications, are too small to be of direct use. However, 13C (or rather delta13) can be used as a natural tracer for determining how, e.g. different environmental conditions or management would affect the stability of soil carbon storages and carbon accumulation.

The first two paragraphs of the introduction (lines 10-20) provides a good justification for improving SOM models. However, the focus in these paragraphs on agricultural soils and carbon monitoring is not well related to the actual model structure and evaluation which only includes litter decomposition and peat systems. Carbon sequestration in mineral soils is sensitive to mineral-organic interactions and mineral-associated organic matter accounts for a large fraction of SOM (e.g., Lugato et al., 2021). However, Yasso does not include mineral interactions and treats humus as a passive pool and was only evaluated using litter and peat decomposition. Therefore, it does not seem justified to introduce the model in the context of agriculture soils. Since the model seems intended to simulate peat systems, I think it would be more reasonable to introduce it in the context of better understanding and predicting carbon dynamics in peatland or organic soils.

Reference: Lugato, E., Lavallee, J. M., Haddix, M. L., Panagos, P., & Cotrufo, M. F. (2021). Different climate sensitivity of particulate and mineral-associated soil organic matter. Nature Geoscience, 14(5), 295–300. https://doi.org/10.1038/s41561-021-00744-x

We have modified the introduction and now initially introduce 13C in the context of ESMs and the sensitivity of soil carbon w.r.t. changing climate. We have not entirely removed the sections related to agricultural soils as this is one of the future development aspects of Yasso. We also note here (as was mentioned in our previous response) that Yasso was originally built and used primarily for C decomposition in mineral soils and C measurements from mineral soils have been an important part of the extensive calibration data (Viskari et al. 2022). Additionally, humus pool decomposition has been part of Yasso calibration on bare fallows (Viskari et al. 2020). We have expanded the Yasso model description by mentions that Yasso was originally developed for forested areas, previous version of the model has been implemented into JSBACH (land-surface component of MPI-ESM) and that the recent recalibration used multiple global datasets.

Other comments:

Section 2.1: The peat depth profile measurements that were used to validate the model should also be described in this section.

This information has been added to the manuscript.

Figure 1: It would be helpful if the figure axes used the L notation that is used in the text so it is clearer what is being plotted. Is marginal likelihood in these plots the same as L?

We have added y-axis labels and now show a wider range for each parameter – previously the image was focused on the area of highest likelihoods. We have also removed the “heatmaps” to simplify the image (these can be added if needed).

Figure 2: Consider using different symbols for the branch and needle data to accommodate red-green colorblindness (which is common) or in the case of printing the paper in grayscale.

Different symbols are now used as well as a different color-scheme.

Line 138: It was not immediately clear to me how relative 13C content can change over time in the default model without any fractionation included. I think this occurs because the initial pools have different isotope ratios and are mixing over time which causes the isotope ratios to change. But a more specific explanation of this would be helpful. It might also be helpful to show a diagram (perhaps in the appendix) of transfers among the different pools so it is more clear what kind of mixing over time can occur.

That is correct, also the flow of matter between the pools differs so both the initial ratios and the different decomposition rates drive this behavior. This information has been added to the manuscript.

Line 145: The actual depths should be included. And I suggest including a more detailed explanation of why the depth sampling was consistent with the 10 year age assumption. Was there evidence from that site that the age difference was actually close to 10 years across depths?

We have added the peat depth profiles and other related information to the “measurements” section of the manuscript.

Figure 3: I suggest splitting this figure into separate panels as in Figure 2. The large number of lines and colors makes the figure difficult to interpret. Also, can bulk 13C in the model be calculated to compare with the bulk 13C measurement from peat?

We have split the image to separate panels and added bulk 13C as well as trendlines for the measurements.

Line 162: The negative parameter values are consistent with the theoretical expectation of slower 13C decomposition rate (as described in the introduction) which is a good result for the model and would be valuable to point out more explicitly.

We have amended the text and now point this out more explicitly.

Line 167: “This situation is not ideal” – why not? Is it inconsistent with measurements or theoretical expectations? It doesn’t seem particularly unreasonable to me.

This statement has been revised.

Line 179-180: It’s not clear to me how the results demonstrate improvement to SOM model accuracy and predictability since they were not used to inform any changes to the overall C decomposition rate or structure. Improvements were limited to 13C dynamics.

This was meant more as a general statement, and we have now included examples in the beginning of discussion on how 13C can be used to examine and analyse C cycles in more detail. However, the referee is correct that improvements to overall C decomposition were not demonstrated in the paper. Therefore, we have modified this sentence to: “We have demonstrated how 13C can be implemented into a soil carbon model, so that carbon isotope signals could then be used to analyse carbon cycles in more detail and to improve model capabilities, accuracy and predictability.”

Line 189: Similarly, it’s not clear that the study made improvements to SOM decomposition in general outside the direct comparisons to 13C content of organic matter pools.

This sentence has been modified to: “The capability of a model to simulate soil 13C and isotope-specific SOM decomposition improves the applicability of Yasso-C13 model to scale process from ecosystem level to regional and global using δ^{13}C as a tracer.”

Referee 2 comments

The preprint manuscript “Implementation and initial calibration of carbon-13 soil organic matter decomposition in Yasso model” describes calibration of the Yasso model to 13C data collected from a litterbag decomposition experiment. The model was calibrated using 13C values measured on sequential extracts of pine litter and branch samples from a 4-year litterbag experiment. The decomposition parameter matrix of the Yasso model was modified to account for 13C using simple scalars. After optimization, three out of 4 scalars were negative, which was consistent with the hypothesis that 13C is preferentially retained in decomposing organic matter. The optimized model was applied to data from a peat core and produced more realistic predictions than the default model.

This manuscript is clear and concise. However, I think this manuscript should be framed differently to better showcase the results. The manuscript is framed narrowly in terms of soil carbon sequestration as a climate mitigation tool. However, the analyses and results are not directly relevant to soil carbon sequestration efforts. Specifically:

- The study system is unmanaged and focused on C cycling in litter and organic soils, and has no obvious connection to the agricultural soil carbon management strategies listed in the introduction.
- The 13C calibrated model performs no better at predicting changes in bulk C, hence its relevance to soil carbon measurement and verification efforts are unclear or at the very least indirect.

This critique is on point and on similar lines as Referee 1. We have now reframed the work more in the context of ESMs, although not entirely removed the sections related to agricultural soils as this is one of the future development aspects of Yasso. The inclusion of 13C is not expected to change bulk C estimates radically (as we demonstrate in the manuscript) and the benefits come from the model capability of using 13C as a tracer. We have added explanations and examples (with references) to the discussion.

Later in the manuscript the significance of the 13C calibrated Yasso model is described differently, in terms of integration with 13C enabled ESMs. This seems like a much clearer justification for the calibration effort. Taken at face value, the results presented here are nearly trivial: calibrated the Yasso model to 13C data results in a better fit to 13C data. As a technical result, this is to be expected. What is the concrete significance of this incremental

> advance for our understanding of soil carbon cycling? What can the calibrated model eventually tell us about the cycling of the bulk C pool or the broader functioning of soil beyond fractionation of 13C?

We do agree with the face value analysis. The concrete significances come from isotopic signatures in soils, litter and respiration and the several ways these can be used to examine carbon cycling in more detail; to improve models; and to analyse the effect of different environmental drivers (and different management practices in the future).

> If the 13C modifiers are generalizable to other systems (which may or may not be the case), I can see how they might enable the Yasso model so that it could be calibrated based on tracer experiments or in cases where the d13C of vegetation has shifted, or how it might be useful for interpreting time series of 13CO2 data to attribute fluxes to different soil C pools. These sorts of application are alluded to, but perhaps the manuscript would stand on its own more clearly if it was framed more clearly as an intermediate step towards these larger scientific goals.

Thank you for the comment and hopefully the changes in the manuscript now better reflect this view. It is now explicitly stated that benefits for modelling etc. Come from the perspective of using 13C as a tracer. We have also added references to papers demonstrating this use.

> **Detailed comments:**

> Abstract: Details of the calibration dataset are not given in the abstract – consider including them.

We have modified the abstract, which now includes the statements: “The model modifications were calibrated using fractionated C, 13C and delta13 measurements from litterbags that were left to decompose in natural environment for four years. The modifications considerably improve the model behaviour in a 100-year long simulation that is compared against fractionated peat column carbon content.”

> Line 1; Line 10: I agree that strategies for increasing soil carbon as a climate mitigation strategy have received increasing attention over the years. However, I think this initial framing is an innapropriate place to start this manuscript (see broader comments above). Carbon cycling in soil is a fundamental aspect of terrestrial ecosystem function. Soil carbon influences the climate system and a whole range of global biogeochemical cycles regardless of how we try to manage it.

We now begin the introduction by introducing C cycling in the context of ESMs.

> Line 7: I suggest deleting “despite of their simplicity”, as it implies that we expect that simple modifications will not generate improvements.

Deleted.

> Lines 21-32: This paragraph begins by addressing the challenge of deciding which processes to include in models, but the application for 13C seems to mostly relate to parametrization. Is 13C useful for both determining model structure and fitting parameters? Are these distinct challenges?

This paragraph was also modified, but the underlying message is still the same. Our approach to include 13C was very straightforward, but alternatives could be more complicated. In our view, model structure and parameter fitting are not distinct challenges and the inclusion of 13C processes adds another approach to test how well different model formulations behave. 13C could be useful for, e.g., determining interaction between soil layers or to test if adding clay content as a driver would be beneficial, providing we would have relevant isotope measurements (something that has been talked about).

> Line 28: Writing edit -- delete “By” before “estimating”.

Deleted.

> Lines 114-115: In other words, the precipitation and temperature dependence was the same for both isotopes? These factors are included in the original “alpha” term?

We have modified the text and now give more details on what exactly has been changed in the modified 13C cycle. Essentially temperature and precipitation affect both isotopes similarly. We have left out “c” from the below equation as it is a stand-in for carbon content.

$$k_i(\theta, c) = \frac{\alpha_i}{J} h(d)\left(1 - e^{\gamma_i P}\right) \sum_{j=1}^{J} e^{\beta_{i,1} T_j + \beta_{i,2} T_j^2}$$

> Line 126: how were the parameter “grid” and increment refocused? Was this done in a systematic way?

Yes, and multiple times with different initial states. The supporting material contains one run with the starting point at the origin (we end up with same results regardless of starting point). We run the model with all combinations from the “grid”, where each parameter was varied similarly; then we choose the local optima as the new middle point for the grid, decrease the value increment and run everything again etc. The text has been modified to clarify this.

> Figure 1: What do the color gradients represent? Likelihoods, presumably? In the panels situated along the diagonal, does the vertical axis on each panel show the likelihood? What do the vertical lines represent – parameter values at maximum likelihood? This caption needs to be expanded to clarify.

We removed the “heatmaps” from this figure as they did not provide any important additional value. The y-axis in the diagonals were the likelihoods. We have now substituted this to L/max{L} as this “normalisation” does not affect the shape of the distribution. Additionally, we have “zoomed out” to give a better view of the shape of the distribution.

> Figures 2-3: Why does d13C change over time in the default case? The default parameters are identical for 12C and 13C, correct? In this case, shouldn’t the 12C:13C ratio be preserved in all transformations, and the d13C value remain the same over time?

We have added an explanation on why this happens to the manuscript. The changes in d13C are driven by differences in the initial litter 13C content (and therefore dd13C) and differences in the matter flow between the pools.

> Methods section: Please include details about the computing methods. How were these procedures implemented? What computing environment was used (e.g., Python, R, Matlab)? Were any R packages used to assist with fitting?

This information has been added to the end of methods section: “All experiments were run on a 8-core laptop utilising RStudio version 1.4.1103. We used the R interface of Yasso (see code and data availability) in addition to R.utils version 2.10.1 (no other libraries were needed).”

> Lines 131-132: I believe there are formal methods for evaluating collinearity between parameters. Computing a “collinearity index” might be useful for determining whether the parameters are identifiable (although such indices still reduce to qualitative rules of thumb). There are methods in R for this sort of analysis (package “FME” might be useful).

Parameter (multi)colinearity is not usually reported. It means that we interpret the parameters as independent variables (predictors) and the likelihood as the dependent variable and model their relationship *via* linear regression; then colinearity means that there (nearly) exists a linear mapping between two parameters and multicolinearity means that one predictor is close to a linear combination of the others – so we can (nearly) exactly predict one parameter value from the rest. We have now increased the x-axis range to 0.2 for which we show the shape of the distributions – each distribution consists of 201 datapoints. Collinearity manifests as too large variances in these distributions, which does not appear to be the case. We calculated variance inflation factors (VIF), but their interpretation is not straightforward and depends on what range we give for each parameter. Additionally, we do not have all combinations for the presented 201 points / parameter (these were produced for optimal values; the grid would consist of $201^4=1.6*10^9$ datapoints) as refocusing the grid etc. enabled the use of a radically reduced number of points. Therefore, these indices are already biased towards areas of high likelihood. The VIF indices for the whole data (in repository) were at maximum a bit over 10, which urges caution. However, the biases are presented in this data as well, which is the reason we have decided not to include colinearity reports in this manuscript. Hopefully, this explanation satisfies.

> Lines 151 – 152: Here the emphasis is on incorporation into ESMs, not MRV for soil carbon sequestration.

The manuscript now focuses more on ESMs.

> Lines 145-146: So depth and time have been exchanged? Is this based on an assumption that the peat is accreting linearly? How was the conversion between depth and time parametrized? Why 10 year intervals, why not 20 or 50 years? More justification/expanation is needed here.

We have added explanations on this to the measurements part of the manuscript, where we now give the peat column layers and ages in a table and justifications for these in the text.

> Lines 167-169: I do not follow this reasoning. Is the non-ideal finding that the parameter for the N pool is positive? How does the lack of depth resolution explain this?

This was meant to underline that the positive N-pool related parameter value goes against the initial hypothesis. We have modified the text accordingly.

> Lines 179 – 180: The results presented here indicate that calibration of 13C parameters to 13C data improves accuracy and predictive power for 13C. However, they do not show how this improves the skill of the model with respect to bulk C pools or fluxes. What can these results tell us beyond 13C fractionation?

This is correct and all mentions to direct improvements in simulating bulk C have been removed. We have added discussion and references to how 13C can be used as a tracer, e.g., to examine in detail carbon cycles and fluxes. The examination of management practices is also alluded to and mentioned as a probable future research.

Hilasvuori, E., Akujärvi, A., Fritze, H., Karhu, K., Laiho, R., Mäkiranta, P., Oinonen, M., Palonen, V., Vanhala, P., and Liski, J.: Temperature sensitivity of decomposition in a peat, Soil Biol. Biochem., 67, 47–54, https://doi.org/10.1016/j.soilbio.2013.08.009, 2013.

Viskari, T., Laine, M., Kulmala, L., Mäkelä, J., Fer, I., and Liski, J.: Improving Yasso15 soil carbon model estimates with ensemble adjustment Kalman filter state data assimilation, Geosci. Model Dev., 13, 5959–5971, https://doi.org/10.5194/gmd-13-5959-2020, 2020.

Viskari, T., Pusa, J., Fer, I., Repo, A., Vira, J., and Liski, J.: Calibrating the soil organic carbon model Yasso20 with multiple datasets, Geosci. Model Dev. 15, 1735–1752. https://doi.org/10.5194/gmd-15-1735-2022, 2022.