“This manuscript by Viskari et al. presents the comparison between the established MCMC and emergent 4DEnVar calibration approaches for the MEMS v1 soil organic model using the continental-scale soil inventory dataset. The authors calibrate the model using a subset of sites where both total SOC and physical fractions were available, and validate these calibration across over 17,000 sites at a continental scale. Then they include a parameter experiment (changed from 0.15 to 0.35) to test the effects of litter input assumption on resulting parameters sets and model projections.

The authors find that MCMC and 4DEnVar calibration methods could achieve comparable performance in terms of total SOC validation with different parameter sets and different internal representation of the SOC state, while the later method is far more efficient. Changes in litter input assumptions affect both calibration methods but lead to general similar behaviors. Thus, 4DEnVar would be a potential tool for efficient calibration, yet more concrete evaluation approach for SOC modelling is needed.

This manuscript makes valuable contributions to the core issue of equifinality in the field of modelling by linking structural discrepancy between calibrating against total SOC and SOC fractions. Comparison between two methods is a practical finding for any modeler considering adopting these method and datasets.

The manuscript is generally well written. The general comments and specific comments are as following:”

We appreciate the positive feedback and the very helpful recommendations.

Line-by-line comments:

As a general comment,

“Line 9-19: It would be beneficial to mention the second object and its findings of this manuscript, i.e. how NPP litter assumptions affecting the calibration and projections, to ensure that the abstract reflects the full scope of this research.”

Thank you for highlighting our error in not mentioning the NPP experiment in the abstract. We have now added this to abstract while condensing the text in order keep the character count down. Starting from line 12, the abstract now reads:

Abstract. An abundant amount of different data is required to calibrate soil organic carbon (SOC) models to represent ecosystems at large-scale. However, due to challenges related to model state projections, this calibration becomes very computationally heavy with traditional calibration methods. Here, we test 4-Dimensional Ensemble Variational data assimilation (4DEnVar) method to parameterize the MEMS v1 SOC model using data from the LUCAS network and compare its performance against MCMC calibration. Additionally, we performed an experiment where we adjusted the litter input partition to see if the two calibration methods react differently to the change. The total SOC projections from both parameterizations showed similar improvements though the produced parameter sets differed. A thorough analysis revealed that the detailed SOC states differed from each other, but we also lacked information to determine which parameter set was closer to the truth. Furthermore, changing the litter input partition highlighted how much that assumption affects the calibration results with both methods. Our results here establish 4DEnVar as an applicable calibration method for SOC models but also highlight the need for more nuanced validation methods, as well as careful examination on how different data sets affect the model calibration.

"Line 18: as well -> as well as"

Corrected

"Line 69-80: While motivation for this study lies in the computation challenge of MCMC method, the introduction of 4DEnVar method would be expected to be introduced in a way that emphasizing on its relative computation cost with concrete references or statistics to strengthen the motivation of this work."

Excellent point and we have now reworked the paragraph to stress the computational benefit from the start. However, we were not certain what sort of concrete reference/statistics would be considered appropriate, here, as this is among the first studies, to our knowledge, that has compared the two calibration methods in this manner and even then, the benefits would be expected to be very system specific.

In order to address this request, we have added a reference to Beylat et al. (2025), where they compare the performance of a land surface model when calibrated with 4DEnVar calibration and with the traditional 4-Dimensional variational assimilation in the context of atmospheric CO2 concentrations. Hopefully this will be a sufficient concrete a reference to address this suggestion. The paragraph is now, starting from line 82:

As a more practical alternative to the costly MCMC approach, four-dimensional ensemble variational data assimilation (4DEnVar; Liu et al., 2008) is a novel data assimilation approach, where a model ensemble generated by varying the parameters/variable states of interest is used to determine the optimal parameter and/or state variables. It has already been used for parameter calibration (Douglas et al., 2025; Pinnington et al. 2020) and is much faster than the traditional MCMC methods. It is based on the Four-dimensional Variational data assimilation (4DVar; Le Dimet and Talagrand, 1986), where a model projection is compared with observations and the new initial state for the next iteration is generated from this information. A key difference between MCMC and 4DVar based methods is that the latter use gradient descent methods to determine the next state instead of randomly sampling. While 4DVar has initially been used more commonly for state data assimilation, for example, in weather forecast (Huang et al., 2009), it has also been successfully applied to calibrate ecosystem models (e.g. Raoult et al., 2016; Peylin et al., 2016; Pinnington et al. 2016). However, to implement 4Dvar with observations from multiple different times, an adjoint version of the model is needed which imposes its own challenges and limitations on the application (Thepaut and Courtier, 1991). The 4DEnVar method, uses the ensemble to sidestep this requirement by simultaneously running multiple simulations with different parameter sets instead of an iterative solution. The 4DEnVar method uses the ensemble to sidestep this requirement by simultaneously running multiple simulations with different parameter sets instead of an iterative solution. While to our knowledge there haven't been previous studies within the ecosystem modelling analysing the performance of the 4DEnVar to that of MCMC, in Beylat et al. (2025) the 4DEnVar method is compared to the original 4DVAR method in a very specific synthetic experiment. Within that scope the 4DEnVar was shown to be more effective than the original version, but it is only the first step in evaluation.

"Line 81-82: It is unclear for the reader why choose MEMS v1 model for this study, as this model is mentioned only at line 81. Would be nice to introduce this model already after mentioning the need to separate different SOC fractions."

We have added a brief model description as well as a justification for the choice starting from line 105:

The model in question simulates organic carbon decomposition separately for above- and below-ground carbon with pathways from surface vegetation matter to the soil pools. In the framework of the MEMS v1, the microbial pool is the central connection between the different SOC states and, crucially, along with the soil properties regulates the amount of carbon stored as long-lived MAOM compounds. For our purposes here, the model is ideal as for the most part the SOC pools are connected by first order dynamics, but the relationship between the microbial and MAOM pool is non-linear. Consequently, there is only a small number of central parameters to calibrate while simultaneously the model steady state cannot be analytically solved, requiring the more costly parameterization process.

"Line 82-84: The advantage of using the LUCAS dataset can be briefly summarized. As a side focus of this paper, I would also like to know the challenges and limitations of using large-scale datasets for model development."

We have expanded this part with an explanation the benefits of using the LUCAS dataset starting from line 115.

Because this LUCAS dataset contains measurements from thousands of plots across Europe and, thus, represents many different types of ecosystems as well as climate conditions, it allows to test a wider performance of the model calibration. One of the advantages was the level of standardisation in sample collection and analysis, the latter done by a unique laboratory, Furthermore, for a small subset of the chosen LUCAS dataset, the POM/MAOM fractioning also had been done, which provided more nuanced information for the calibration process. While Lucas is a standardised framework for SOC, was not specifically designed to assess the MAOM stocks.

As for the challenges and limitations, we have added a small paragraph into the end of section 2.1 about those relating to the dataset here. Now, from line 155, there is the following description:

While the benefit of the LUCAS dataset is its large spatial representation and inclusion of measurements from multiple different ecosystems, the execution of such a vast measurement campaign introduces different source of errors from sampling, labelling, analysis etc. Thus, it is almost more apt to be considered as a combination of several independent campaigns done with the same protocols, instead of a single consistently controlled campaign. Additionally, although locations of the measurement are known, we have the make the assumptions that the available driver data are representative for the actual conditions at the measurement plot.

“Line 85-86: From the previous context, I can’t see why can draw the hypothesis of ‘fit to the same degree’.”

We rephrased the hypothesis according to the feedback and now it reads from line 122:

Our hypothesis is that the 4DEnVar improves the model fit to a sufficient degree that, along with the reduced computational cost, it can be considered as valid calibration approach for SOC models as the MCMC

“Line 87-91: The two objectives of this study are not articulately explained. Why this ‘simple experiment’ important in the context of ‘comparison of calibration methods’, and how is the first and two objectives inform each other?”

We have made the explanation more explicit starting from line 124:

Specifically, there are two objectives for the work presented here: the first is to test if the much faster 4DEnVar calibration performs as well as the MCMC calibration and examine if there are any meaningful differences in the resulting parameter sets; the second is to conduct a simple experiment where we made a change on how the NPP litter input was partitioned. The reasoning for the latter objective is that one of the core benefits of the faster calibration method is that it allows testing how different assumptions impact the parameterizations. Because of this, if there are differences between the results of the two calibration methods, it is important to assess if the general behaviour of the parameterizations remains the same even under different assumptions.

“Line 95-102: I wonder how presentative the subset of 350 samples is, which might be a big source of validation error. Would you please show the distribution of this subset in Figure 1 and provide more information regarding the contained context.”

The subset in question was randomly drawn with the distribution of ecosystems kept constant. Thus, it should be as representative of the whole as possible. To make this clearer, from line 140, we have now added this detail:

The latter were randomly drawn from the all the measurements with the only constraint being that both datasets were similarly distributed across ecosystems with approximately 73 % being grass- or croplands with the rest being various forest types. The representativeness of the chosen 350 measurements points is elaborated upon in Lugato et al. (2021).

We have also added a panel to Figure 1 that shows the spatial distribution of the calibration datapoints. The Figure 1 is now:

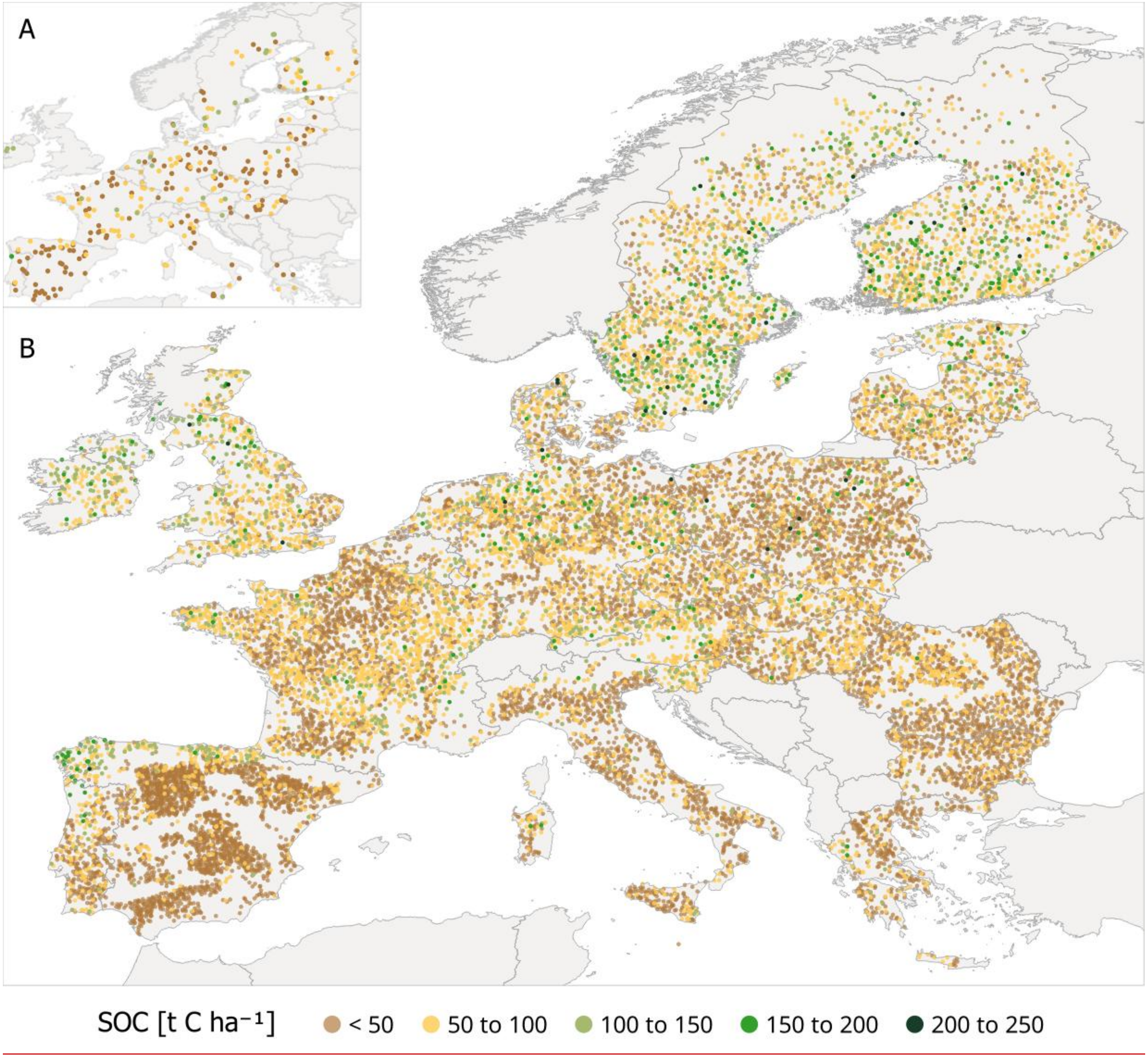

Figure 1: The LUCAS 2009 sampling points across Europe and their SOC stock used for A) calibration and B) validation.

"Line 105-109: I wonder what is the intention to include agricultural soils into simulation, where the-steady state assumption might not hold true?"

We simply chose to include all the representative ecosystems in Europe into the calibration. Especially since even when using SOC models in agriculture, most of them have also been calibrated with steady state assumptions since we very rarely have a reliable detailed initial state that could then be used as a starting point when calibrating time series data. Although in this particular environment, Europe has been similarly cultivated for so long that the steady state assumption is more reliable here than in many other regions.

Furthermore, even in grasslands and non-commercial forests, the steady state assumption is fraught now due to climate change that is affecting all soils in the world.

Thus, there is no such thing as a steady state, but it is unfortunately assumption we still have to rely on.

This is still a valid comment, though, and we have added a line to highlight this as an additional uncertainty source in section 2.5 as we explain the steady state approach itself there. The addition from line 365 is:

It should be noted that with agricultural soils and commercial forests are expected to have a large variability in litter input over a given time window, which does raise challenges for the steady state approach. We are still including those data points in the analysis here as this is intended as a general calibration across European ecosystems and there is no additional data to constrain those specific ecosystems, but this is expected to be an additional uncertainty source.

“Function 8-11: What is r^eco?”

Apologies for us forgetting this definition here and our gratitude for pointing it out. Now the term is explained on line 226.

Finally, the r^{eco} represents the fraction of NPP that is assumed to have been removed from the system due to economic activies (harvest, grazing, etc.)

“Line 162: Wrong equations reference?”

Yes, corrected to 8 to 11.

“Line 212: Wrong equation reference?”

Equation reference corrected to 12.

“Line 224: Wrong equation reference?”

Equation reference corrected to 14.

As the primary, my sincere apologies for the wrong equation references here. There was a large restructuring of the manuscript at a late stage of preparation, and these were unfortunate remnants from that. Our gratitude for reading with such attention to notice these errors.

"Line 268-270: Hard for me to follow the description of the twin experiment. What are the baseline parameters for the parameters you calibrate in this model? How do you generate the perturbed parameter set?"

We expanded the paragraph on line 343 in order to clarify the raised points:

After having set up the algorithmic framework for both calibration methods for the selected LUCAS data points, the first task was to complete twin experiments. In those, we randomly drew a value for each the parameter chosen for calibration from the uncertainty distributions assigned for them in Table 1. Synthetic observations were generated with the model using the new parameter set. Then, we performed the calibration with both tested methods using these synthetic observations with their associated uncertainties set to be 1 % of those synthetic observations and still using the same prior distribution established in Table 1. This allows us to check if both methods were able to find the correct parameter sets in a situation where the true answer was known. For the 4dEnVar, the additional importance of these tests is to assess the ensemble size dimension required to consistently estimate the correct parameter set. This was accomplished by repeating the twin experiment multiple times with different ensemble sizes and choosing the ensemble size where the calibration always found the correct parameter set. The repetitions were necessary as, because the 4DEnVar ensemble members are randomly drawn, there are potential situations where with a given ensemble size it retrieves the correct parameter set four times in a row, but then fails on the fifth time.

"Line 299-304: Please provide the full name of the MODIS product. And not clear for me that why the NPP dynamics are not expected to meaningfully affect the modelling results."

We added the product name to the paragraph starting from line 397. Additionally, the reason we assumed that the NPP assumption wouldn't impact the results was the total annual NPP input stays the same. That is now made explicit in the text.

For Net Primary Production (NPP), first the average annual NPP over the decade 2000-2010 is extracted from the MODIS product MOD17A3 (Running et al., 2004) grid cell overlaying each LUCAS point. Then, a standard sine function is used to distribute the NPP across the year in order to produce the daily litter input. This approach was used instead of an averaged MODIS NPP annual time series as the NPP reflects the time when the atmospheric carbon is allocated into vegetation, not when the vegetation becomes litter input. Hence, we simplified the time series here and, since the total annual NPP remains the same, it is not expected to affect the modelling results to a notable degree.

"Line 363: I think the baseline parameters are not presented here? Also I wonder if you could provide uncertainty measures in this table? 0.3s5 -> 0.35."

Thank you for pointing out the baseline parameter line, a remnant of an older version of the manuscript. We have also corrected the 0.35 typo. The table header on line 478 is now:

Table 3: The expected parameter values produced by the different calibration methods. The first value is for f_{doc} 0.15, the second for f_{doc} 0.35.

As for the uncertainty measures, we have already presented those in the Figure 2 with the parameter distributions. We have also now added the standard deviations in Supplemental Table 2. Originally, we tried to include them in table 3, but in our opinion the resulting table was far too cluttered.

"Line 376-377: I wonder if you could provide qualitative statistics for the error distribution. As the difference between two methods with f_doc=0.35 can't be intuitively recognized."

Another oversight on our part and we are grateful for this review as well the others to drawing attention to it. We have now added the error statistics to the manuscript and expanded the Results section to discuss these results. From line 480:

To examine the impact of the new parameter sets, Figure 3 presents the differences between the measurements and model projections across all the validation sites, while in Table 4 we show both the Root Mean Square Error (RMSE) and mean error (ME) representing bias in regard of the validation dataset for each parameter set. While the 4DEnVar parameter sets produces a somewhat symmetric error distribution around zero in both calibrations, with the higher f_{doc} there is a slight apparent tendency towards positive errors. In contrast, the MCMC error distribution shows a notable lean towards positive errors for the lower f_{doc}, while with the higher f_{doc}, the bias is much reduced. Since the SOC errors here are calculated as the measurement minus the model projection, this means that positive errors reflect the parameter set systematically underestimating the SOC projections. It is notable that with the higher f_{doc}, the RMSE values for the two parameterizations are very closer to each other even with the larger positive bias of the 4DEnVar method.

	f_{doc} 0.15	f_{doc} 0.35
MCMC	42.5 / 27.4	31.3 / 7.4
4DEnVar	29.8 / -1.9	32.0 / 14.2

Table 4: The error statistics for the different parameterizations with regard to the validation dataset. The first value is for the root mean square error (RMSE) and the second for the mean error (ME). The unit for all the values is t C ha^{-1}.

"Line 401-403: The spatial pattern in Figure 5 seems to be coincided with the contexts in Line 404-408, and conclusions in 422-425. I wonder if you could discuss a bit more the regional patterns and related driver data limitation, model structure limitation, impacts in model performance evaluation in the discussion section. To make it more concreate in depicting the challenge in modelling with large-scale dataset."

An excellent suggestion that adds depth to our results in this work. We have expanded the discussion on two parts on the regional impacts.

First, when discussing the role of limited soil moisture dynamics hinder the model in question, we added a sentence on line 614 to connect those to the biases we see in the Iberian peninsula:

This could be a partial explanation for the Iberian peninsula error biases visible in Fig 5 as the soil moisture dynamics are much more complicated in arid climates vulnerable to drought (Almendra-Martin et al., 2021)

Then, in the NPP assumption section of the Discussion, we wrote a new paragraph starting from line 672 where we touch on the spatial element of that issue:

Adding to the challenges discussed above is that the various assumptions are not expected to be spatially homogeneous even in the same ecosystem type. For instance, the Nordic countries, especially Sweden and Finland, are dominated by economic forests where the NPP-to-litter pathway is heavily impacted by the growth stage as newly growing forest will have a large NPP, but not a corresponding amount of litter due to mortality. This could be connected to bias seen in the northern Europe in Fig 5. Another example would be agricultural ecosystems as climate conditions affect which crops will be dominant in a given region. The type of crops naturally affects its traits as, for instance, the root depth distribution, which in turn is expected to impact the soil carbon stocks (Fan et al., 2016). These various components could be a reason why when analysing global soil databases, there is a weak statistical relationship between NPP and SOC despite that dynamic being well understood (Luo et al., 2021).

"Line 444-453: looks more like results for me, and these are not brought up in the results section."

We agree and have moved the comments about the cost function to the Results section to line 508.

As for this paragraph here, based on comments from another reviewer, we have expanded it to discuss in more depth equifinality and how the results here contribute to that challenge. From line 574:

What is striking, though, is how much the parameter sets produced by the two calibration methods in both litter distribution scenarios differ from each, even with the higher f_{doc}, they perform approximately equally well with regard to the total SOC measurements in the validation dataset at the given time. Equifinality, a situation where there are multiple parameter sets that produce similar model outputs, is a known issue in ecosystem modelling in general (Sierra et al., 2015; Marschmann et al., 2019), but the surprising element here is that the calibration method itself determines the resulting parameter set. Generally, Twin experiments are efficient first pass to test for equifinality and the challenge can be addressed by reducing the amount of parameters being calibrated, but here there are questions how much those efforts can be relied on in assessing equifinality.

"Line 466-492: I wonder if you could provide a bit more information on the potential solutions to solve the prior impacts on calibration. So that this part would be more useful for practitioners considering using 4DEnVar model for similar applications."

This is a really good request and something we should have addressed in the original version of the manuscript, even if there are naturally no easy answers.

In response, we have brought up the importance of providing uncertainties with the measurement sets as well as putting more effort in producing the parameter prior distributions as there is often a tendency to just give uniform distributions. We also point out that when using multiple datasets for calibration, it is important to first confirm if they are compatible with each other.

Finally, while we have only done a single calibration cycle with the 4DEnVar in the study here, the results from a single calibration can technically be used as the prior range for the next one. The benefit of doing this is that by repeating the process multiple times until the results themselves do not change is that it allows addressing the concern that the initial prior range is too far away from the actual value.

To provide a hypothetical example of this, let's say we have a parameter where the optimal value is 2, but because our prior is too badly defined, our expected value is set 6. With one calibration cycle, the estimated value can end up around 4 just due to the sampling of the prior distribution. Yet by repeating the calibration in this situation, the next posterior distribution would be nearer to the correct value.

Why we have not done this, though, and why it is not as straight-forward as it might seem is that this repeating cycle reduces the impact of the prior range which goes against Bayesian philosophy. Furthermore, with each iteration cycle the uncertainty range is reduced which, in turn, results in unrealistically high trust in the results. Thus, while the repeated calibration cycles is a possibility, it needs to always be implemented with

Additionally, there is also a practical solution in having the 4DEnVar algorithm to sample outside the prior distribution that would help with some of the issues, but since that is more of a question of the parties creating the algorithms than the user, we do not mention that here.

The new paragraph on line 634 reads:

Naturally this underlines the overall importance of providing reliable measurement uncertainties along with measurements themselves, but that is not something a model user can simply produce by themselves. When implementing the calibration, based on the results here we would recommend of initially looking through the calibration data and confirming that all the values are sensible for the model/system being calibrated. As a more practical solution, it is possible to repeat the 4DEnVar calibration multiple times by using the previous posterior distributions as the priors to the next cycle. This way it is possible to ensure that the resulting parameter set is not simply because the prior had been set too far from the correct value and thus partially reduce the impact of the assigned prior distribution. However, the downside of repeating the calibration cycle in this manner is that not only does it reduce the impact of the prior, but each iteration reduces the resulting uncertainty distribution. Thus, the final parameter distributions would be artificially too confident. While the repeated calibration is a worthwhile tool in certain circumstances, it always needs to be implemented with great care and consideration.

"Line 500-511: Current discussion mainly focuses on the challenges we are facing right now. It would be beneficial to put it in a broaden ecological context, and discuss about the ecological outcomes of changing the f_doc parameter. This might help to assess whether the chosen f_doc parameter is ecological reasonable."

Reasonable request, although it is important to note that within the context of the model implementation here, the f_{doc} value actually govern the litter amount directly deposited into the soil. Thus, it reflects more what is the division between the above- and below ground biomass distribution for different plant species and how we reflect that in our modelling work affects many different facets of the results.

We have expanded this paragraph, and splitted into two, to discuss the ecological representation of that variable. Now, from line 654, the new paragraph reads:

What complicates future work is that coefficients associated with litter input are challenging to calibrate simultaneously with parameters associated with SOC decomposition, as their influence on the SOC overlap too much. It is important to note that while the focus in this experimentation has been the f_{doc} value, what it actually represents is the assumption of dividing NPP between upper- and below ground biomass as it reflects the amount of litter deposited directly into the soil. This is a central assumption that has to be included in some manner in SOC modelling and is represented by the plant species traits assigned to the surface vegetation. This highlights why better understanding of the vegetation qualities of the ecosystem being modelled is important for calibrating even simple SOC models.

As for even attempting to calibrate the NPP/litter coefficients simultaneously would first necessitate determining which exact coefficients would be calibrated. For example, in our case, there is first the question how well the MODIS NPP product represents reality for different systems. Then, part of that NPP is removed to represent economic activity before it is distributed to the four MEMS initial pools based on the three coefficients. Any of these three parts can be altered to change the final NPP input to the soil in different ways, but there is really no certainty at the moment what is the correct manner to better regulate the NPP based litter input. This complicated relationship in the surface vegetation driving litterfall and the SOC state has been shown in prior work such as in Raczka et al. (2021). There when they used remote sensing data to constrain their model state, while this improved their modelled aboveground biomass and carbon exchange accuracy, it also caused their modelled SOC accuracy to decrease because they were only using the aboveground data for both systems.

"Line 552-554: I wonder if this sentence should be stated more carefully, by specify the NPP assumption, while MCMC resulted in lower J under certain circumstance."

Good point, although this statement was in regard of the validation dataset, thus we did not weight this from the J perspective. We have now rewritten the sentence from line 724 onwards as:

In our work presented in this article, we have shown that 4DEnVar parameterization produces the approximately same RMSE for the validation dataset as the traditional and more cumbersome MCMC DEzs algorithm when the soil litter input is increased and actually outperforms in this metric the MCMC with the lower litter input.

“Line 557-559: The conclusion of the second object of this work is rather vague, and didn’t explicitly summarize the ecological meaning. Otherwise, it looks like just a repetition of the first object’s finding, without drawing importance of the second object.”

More than a valid criticism and we have expanded the conclusions here to both be more concrete as well as stress how this reflects the role of ecosystem related assumptions. On line 730:

We also conducted a simple experiment to assess the impact of changing how the soil litter input is distributed among different litter pools. These results showed that while the litter input adjustment did impact the calibration, the general model behaviour produced by the two calibration methods remained similar. This implies, if it holds true with further testing, that the differences between the behaviours of the two calibration methods are not dependent on the driver data. Another facet of these results is that it confirms how large of an impact ecosystem related assumptions have on the resulting calibrations.

“Line 562: Might be nice to add one sentence at the end and return to the broaden context of SOC modelling goals.”

We considered this a good suggestion and added the following starting from line 740:

This will make it more pragmatically possible to assess how various assumptions impact ecosystem model results as well as better include those uncertainties in future projections as the various drivers are altered by climate change.

Added references:

Almendra-Martin, L., Martinez-Fernandez, J., Gonzalez-Zamora, A., Benito-Verdugo, and Herrero-Jimenez, C.M.: Agricultural Drought Trends on the Iberian Peninsula: An Analysis Using Modeled and Reanalysis Soil Moisture Products. Atmosphere, 12(2), 236, 10.3390/atmos12020236, 2021

Beylat, S., Raoult, N., Bacour, C., Douglas, N., Quaife, T., Bastrikov, V., Rayner, P.J., and Peylin, P.: Towards the assimilation of atmospheric CO_2 concentration data in a land surface model using adjoint-free variational methods. *Geosci Model Dev*, **18**, 7501-7527, 10.5194/gmd-18-7501-2025, 2025

Fan, J., McConkey, B., Wang, H., and Janzen, H.: Root distribution by depth for temperate agricultural crops. *Field Crops Res*, **189**, 68-74, 10.1016/j.fcr.2016.02.013, 2016

Luo, Z., Viscarra-Rossel, R.A., and Qian, T.: Similar importance of edaphic and climate factors for controlling soil organic carbon stocks of the world. *Biogeosciences*, **18(6)**, 10.5194/bg-18-2063-2021, 2021