Referee comment on "The carbon budget of the managed grasslands of Great Britain - informed by earth observations" by Vasileios Myrgiotis et al., Biogeosciences Discuss., https://doi.org/10.5194/bg-2021-144-RC1, 2021

The authors presented a modelling framework to simulate the carbon budget of the managed grasslands of Great Britain for the year 2017 and 2018. The framework is consist of a biogeochemical model DALEC-Grass, a model-data fusion (MDF) framework CARDAMOM. The authors introduced two sets of earth-observation (EO) data to constrain and optimize the modelling framework. They further applied the framework to infer grassland management across the GB, and further predicted the carbon budget of the managed grasslands of Great Britain for the year 2017 and 2018. The topic is interesting, and the proposed framework is useful, especially for inferring grassland vegetation management across a region. However, the current description of the Materials and methods are not clear, which prevent making further assessment of the quality of this study. Here is the list of my major concerns:

1) It is not clear why and how the two sets of remote-sensed LAI were used in the study. I would think both are EO-based data. It seems that CGLS LAI were used as input to DALEC-Grass, while Sentinel-2 LAI were used in CARDAMOM to optimize parameters of the DALEC-Grass model. As one set of LAI was used as input, it will not be surprise the modelling framework can give reasonable LAI against LAI from another dataset. I would suggest the authors to clarify the reason and the necessity of using the two sets of EO-based data.

2) After going through the Materials and methods section, it is still not clear e.g., how the different models/frameworks were connected; how and where the EO-based data were used; how the model parameters were optimized; how the C fluxes were estimated. A flowchart represents all the inputs, model connections, and outputs with step-by-step procedures will be very helpful.

3) It is not clear how grazed/cut events (as a critical result of this study and an important
component of the C budget) were identified, and grazed/cut biomass was simulated. These are the most interesting and important part of this study. This part of the methodology may need more details. An example of i) the variations of original EO-based LAI, modelled LAI, associated C fluxes, ii) how exactly the grazed/cut events were identified, iii) how “mostly grazed” and “mostly cut grasslands” were differentiated is necessary.

4) The components of C budget were only very briefly mentioned. It is not clear how each component was estimated. Especially for manure, I can not find how it was estimated (or derived from another dataset).

In addition, it seems that the manuscript was not carefully checked before submission. There are plenty of “i.e. …”, “?”, “??” in the text (e.g., L80, L91, L175, L209, L231, L298, L374 etc.) that looks like unsolved comments from the authors, and the manuscript is not taken seriously at all. Such mistakes should not be presented in a submission for peer review.

Specific comments:

- How the sampling of grassland fields can result in only 1-5 simulated fields per cell?
- What are the Metropolis-Hastings (MH) method and the Simulated Annealing (SA) algorithm? What is the difference between them.
- For Fig. 2 and 4, it will be very useful to show not only the absolute values/biases, but also fraction of bias or (mean of MDF-predicted – census) / census, and maybe discuss the reason of bias. For Fig. 4, it might give insights on the mismatch due to the different years of prediction and census.
- How the mean C fluxes across the GB were calculated? Area weighted? If so, how? Whether the selected points are representative for all grassland grid cells?
- It would be necessary to provide the maps of rough grazing, permanent and temporary grassland, and the maps of resulted management type (e.g., grazed only field or grazed + cut field), grazed, and cut biomass for users to understand the management intensity.
- It is strange that NEE/NBE were negatively related to both GPP and REco.
- As the uncertainty for LAI is nearly half of mean LAI, the robustness of the prediction should be further discussed.
- “Mostly grazed” and “mostly cut grasslands” were not explained before results section.
- Paragraph started from L409: It seems that the second assumption is not an assumption but observed phenomena. The logic of the discussion here is hard to understand. Why the C source/smaller sink caused by drought in 2018 can infer
management is more important than climate?
- All the abbreviations will need to be explained in the main text in addition to the “Abbreviations” in the beginning.

Minor remarks:

L12: For RMSE and bias for LAI, should they be m2m-2?

L15: NEE and NBE for 2017 and 2018 respectively should have 4 values rather than 2.

L244: CARDAMOM or CARDAMOM-DALEC-Grass?

L265: “temporary” grassland?

L279: “ny” should be “by”?

L272: Gridded spatial pattern in SI would be very helpful to understand the gradient of first cut across GB.

L294: What is met driver?

L302: Why such interesting results were not presented?


L313-314: It is hard to understand the logic of this sentence.

Figure 7: Given the same r for A-B and B-A, half of the figure (i.e., a tri-angle of the r
map) will be enough.

L356-360: The authors argued two reasons for the under-representation of cut-only grasslands. But it is not clear from the text, how these two reasons will cause the under-representation. The last sentence of this paragraph is very hard to understand.

L367: Who is the first author of “et al., 2011”?

L384-387: The logic is not clear here. Grazing is also continually reduce aboveground biomass, and theoretically, the same as cut ones, the model should also allocate more C to aboveground tissues. Then how cut sites have lower root-shoot ratio? Is there any evidence?

L426: Again, the maps of grazed only and cut+grazed grid cells were not given. People can not see what the authors mean without such maps.