RC Comment on bg-2021-144

Katja Klumpp

Community 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-CC1, 2021

The manuscript (https://doi.org/10.5194/bg-2021-144)" The carbon budget of the managed grasslands of Great Britain constrained by earth observations", shows the potential of uses of high-resolution frequently-retrieved satellite data (earth observation, EO) combined to biogeochemical modelling, for estimate grassland C balance under different management practices and climatic years. In brief, authors used a parsimonious process-model (C pools, and fluxes) of grassland C dynamics (DALEC-Grass) which was integrated into a probabilistic model-data fusion (MDF) algorithm (CARDAMOM). CARDAMOM generates field-specific calibrations of DALEC-Grass by assimilating satellite based LAI time series. In an earlier study the Cardamom has been validated against an extensive dataset in Scotland and South West England. In the present manuscript, the MDF algorithm was applied for a sample of 1855 managed grassland fields in GB (England, Wales and Scotland) over two years (i.e. 2018 summer heat and $\approx$ 1° C warmer than summer 2017) and by uses of UK Land Cover Map for grasslands. The DALEC –grass is driven by six meteorological drivers (Tmin, Tmax, Radiation, CO2, VPD) obtained by European Centre for Medium-Range Weather Forecasts (ECMWF), and was initialized with SOC values (0- 60cm depth) at 300m resolution of SoilGrids database. The basic agricultural practices agricultural (animal number) were obtained from the EDINA AgCensus data base (i.e. 5km grid,AgCensus, 2020). Detailed grassland management such as grazing intensity, cutting timing, are simulated by DALEC-Grass for every field using a local model calibration. In this study CARDAMON estimates were compared to biomass utilisation from the relevant literature and livestock density data from recent agricultural census data. Authors assed 4 objectives , i) detections of grassland management variability by satellite EO products (LAI), ii) estimation of subsequent C balance of managed grasslands , iii) possible indicators allowing to predicted C balance and biomass removals in grazed grassland and iv) analytical uncertainty on C balance and biomass removals of model simulations.

Results show that the CARDMOM algorithm was able to effectively assimilate the Sentinel-2 based LAI time series (overlap=80%, RMSE=1gCm− 2 , bias=0.35 gCm− 2) and predict livestock densities per area that correspond with independent census-based

data (r=0.68). According to estimates, grassland ecosystems were a sinks of C (net biome exchange, NBE) of -232 ± 94 and for 2018 was -120 ± 103 gCm− 2 y− 1 in 2017 and 2018. The lower 2018 numbers were associated to the summer drought which reduced C sinks.

The manuscript is an interesting assessment of the uses of EO-model ensemble for grasslands. The four objectives a most relevant as grassland vegetation and in particular grazing is difficult to distinguish by EO. Accordingly, this study well presents how, by fusing earth observation data and biogeochemical modelling, allows to determine C fluxes (balance) of managed grassland in GB. Likewise, the manuscript shows that the MDF framework (EO-model-algorithm ensemble) can detect biomass removals and use them to predict grassland C fluxes and balance. The manuscript is worth to be published but would need some clarification in the MM section to help the reader to get through. **The discussion part could to my opinion be a bit more lively. According to my below comments I recommend "Revisions"**

General comments .

The paper is well written and comprehensible, once the reader has understood which module/sub calculation and inputs, flows into which estimate (which inputs are predicted/extrapolated, databases, ..). Accordingly, I was wondering if a flow chart /scheme would help to guide the reader though the "model simulation"; i.e allowing to distinguish between "hard/real" data inputs from databases, (soil grid, management practices Edina AgCenesus and meteo) and those which are "elaborated" EO LAI data and how the feed into each other…

There is very little description on the Cardamom simulation outputs (the biochemical model itself) and I was somehow surprised to read "*spatial distribution of MDF-predicted GPP, REco, NEE, NBE, removed biomass and C flux into SOC…* " in the result section. I suggest to add some more details on model outputs in the MM; (E.g. L122 …to calculate primary productivity (GPP), autotrophic and heterotrophic respiration , Reco).

The MM section (L133ff) misses some description on how C balance is estimated, and on how Cardamom accounts for/ estimates (or not) C sequestration??, as some of the results are difficult to understand (e.g. L314 ; "The rate of C inputs to soil does not account for the loss of C from the SOC pool"). Along the manuscript I missed some explanation on the difference between C sequestration and NBE??

-As well as how to get from one term to the other ect. (E.g. Why harvest is not removed in the NBE Table1, L371ff), as to my understanding NBE= NEE-harvest+manure. In short, C balance, used terms and NBE vs. SOC changes (C Sequestration), needs clarification in MM and not only in the abbreviations!.

The same was for manure. Where did Manure come from and how Cardamom accounted for Manure (C/and N) ? EDINA database? (see L368)

The beginning of the discussion section (4.1) is a bit sparely being a description of key results, instead of a discussion on possible improvements; assumption/hyps which have been too "shallow/severe" (e.g. … in the MDF, RF,) in DALEC-Grass.

– Personally I would add the section of Limitation of the approach and opportunities (e.g. SWOT) here. As well as suggestions, what can be modified and what we can learn? However, these sections are standalone at the end of the manuscript, and I wonder if they should/can be moved to the corresponding sections (at the top of the discussion instead of 4.4 uncertainty and 4.5 limitation), which would make them more complete for the reader.

Having said this, the section on C balances need more discussion on the limits of the study, the usefulness for national inventories (i.e. NBE vs. SOC changes see section future work), ….

I am not quite sure the cited studies (L390ff) were interpreted in the right direction. Accordingly, it is very important that authors indicate throughout the manuscript how to read their numbers with respect to C source and sinks and negative/positive signs, respectively

Specific comments

L 87 may be cite : Pique et al 2020 Remote Sens. 2020, 12, 2967; doi:10.3390/rs12182967

L 134ff: "At each time step the algorithm reads the vegetation reduction information and decides whether to simulate the corresponding … " this is not quite clear and I wonder of a flowchart will help? What is the time step? Do you mean management practices comeing from Edina AgCenesus . If yes please refer to section 2.1.5

L209 a set of 2108 fields (Fig, ??). number missing

L175 "and a temporal resolution of 10 days (?). " information missing?

L 199 personally I find the title misleading, I suggest "2.2.1 Sampling of grassland fields from landcover map"

L212ff "To assess the effectiveness of the LAI assimilation process we quantify the level of fit between MDF-predicted and EO-based time-series using …" Until now I did not get that Cardamom estimates LAI (see L 141 and L241) put is used this as an input. Seems I have missed a point. Can authors please clarify. (eg in a scheme?)

L234 "The estimated SHAP values are normalised (0-1) to be comparable to r^2 ." So 1 would be very good ? and what is the number for low fitting (ie limit of SHAP)

L235 and L 331ff "RCR is equal to the size of the MDF-predicted 95% confidence interval divided by the corresponding…" please help the reader to get the number in the right way. eg RCR is 42 ± 9% for LAI, means the uncertainty of LAI is 43% so very high? Or very low? With respect to which best value?

We L252 "mirrored the census-based livestock density data well" please add the name of the DB (EDINA?) so that reader can follow

L280 ff please help reader to get if a number is a sink or source of C??? (… NBE of fields dominated by cutting was 38 ± 70 $gCm^{-2}y^{-1}$ while fields with more grazing had a NBE of -126 ± 95 $gCm^{-2}y^{-1}$ ..)

L363 please mentation that these a flux tower measurements

L393-398 move to the corresponding section L

L370 there are number of other papers having used NBE estimate potential C sequestration Soussana et al 2011, Zeemann et al , Hoertnagel et al. Merbold et al

L416 "conclude that management is more important than climate in terms of the C

balance of managed grasslands in GB." --do authors have a citation which confirm/underline this interpretation

L435-44 suggest to move to the L355ff? keep only L441 - 448

Table1 can authors add grazing and cutting

Table 2. Normalised SHAP, is 0.3 acceptable? Please add some indication to help the reader

legend Table 1. please help reader to get if a number is a sink or source of C???