

Geosci. Model Dev. Discuss., referee comment RC1
<https://doi.org/10.5194/gmd-2021-190-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Review of Yang et al. (gmd-2021-190)

Anonymous Referee #1

Referee comment on "CARDAMOM-FluxVal version 1.0: a FLUXNET-based validation system for CARDAMOM carbon and water flux estimates" by Yan Yang et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2021-190-RC1>, 2021

Review of Yang et al. "CARDAMOM-FluxVal Version 1.0: a FLUXNET-based Validation System for CARDAMOM Carbon and Water Flux Estimates"

General Comments

Yang et al present a consistent validation framework (including a suite a model evaluation metrics) for a terrestrial biosphere model data assimilation system (CARDAMOM) that can be used to systematically test different versions of the model in addition to various DA system configurations (observation record length etc). The introduction really nicely lays out the motivation for developing a rigorous DA-validation framework and the analyses presented provide a useful demonstration of how that framework can be used to answer the key questions posed in lines 90-95. This is important work that all TBM DA groups need to be routinely performing, and as such this paper should serve as a guiding framework or benchmark that I believe will be of wide use to the TBM modeling community.

I have very few comments or suggestions – this paper could be published as is. The following are requests for minor clarifications or questions based on curiosity.

Minor specific omments

Line 104-105: Am I missing something here because "one solely trained by satellite and inventory datasets, and the other constrained by 50% of data at each FLUXNET site" is not

the same as what is presented in the CARDAMOM Analysis 1 and 2 (lines 136-139)? [Updated comment] This actually only becomes clear once we've read the first paragraph of section 2.2. It might be useful to make this more clear at lines 104-105 so readers know what is coming up.

Line 119: I am wondering why the authors didn't use the FLUXNET site met forcing but the ERA-Interim grid instead? I wasn't sure based on reading lines 119 and 155 if this was the case, but it became clear in Section 4.1. Why use a coarser scale forcing when that could introduce biases and the gap-filled site met forcing are readily available (although I appreciated this caveat is acknowledged in the discussion)? (In fact, the FLUXNET 2015 site data have been gap-filled using ERA-I based on the method of Vuichard et al. (2015)). Other caveats related to the different spatial resolutions (and need to aggregate to 1km) for the satellite data used in the assimilation could also be addressed in Section 4.1.

Line 150: might be good to put the original spatial resolution in Table S1.

Fig. 7: really interesting that there seems to be a dip in model performance with record lengths of 2-3 years, especially for the C fluxes. Any thoughts as to why that might be? Is it again related to the types of sites/land cover types with that record length?

Line 36: Minor comment but I think there maybe the latest GCB budget paper might be a better reference here than Le Quere et al. (2020) here given this paper is focused on the impact of COVID-19 on the global C sink (latest GCB budget is Friedlingstein et al. (2020))?

Line 39: If the authors would like an updated version of Friedlingstein et al. (2014) see Arora et al. (2020).

Line 41: I'm not sure I'd use the Reich or the FLUXCOM references for "improve modelling of the Earth's climate system and reduce uncertainties" – especially in the context of terrestrial biosphere modeling as these references refer to a scaled up flux data product?

Certain references are given as the discussion (preprint) version of the manuscript and not the final accepted paper and some references are missing.

Table S4: I am guessing for the initial conditions you mean "at time $t=0$ " – i.e. you only update these at the beginning of the assimilation window? Also, I'm probably missing something but what does "at T, P " mean?

Figure 1: I guess the tables of metrics are just for the prediction window, or for the whole timeseries?

References

Arora, V. K., Katavouta, A., Williams, R. G., Jones, C. D., Brovkin, V., Friedlingstein, P., ... & Ziehn, T. (2020). Carbon-concentration and carbon-climate feedbacks in CMIP6 models and their comparison to CMIP5 models. *Biogeosciences*, *17*(16), 4173-4222.

Friedlingstein, P., O'sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., ... & Zaehle, S. (2020). Global carbon budget 2020. *Earth System Science Data*, *12*(4), 3269-3340.

Vuichard, N., & Papale, D. (2015). Filling the gaps in meteorological continuous data measured at FLUXNET sites with ERA-Interim reanalysis. *Earth System Science Data*, *7*(2), 157-171.