Ma et al. evaluate an updated version of a land ecosystem model, the Ecosystem Demography model v3, which is a DGVM based on individual-based approach. In this paper, they first introduce the new model structure briefly, and then compared the simulated results, mainly on global carbon cycle, with reference datasets. The comparison includes global cumulative/average values, and spatial and temporal variations of carbon fluxes/pools. One of new features of the evaluation process is utilization of global observation on vegetation structure (vertical profile of LAI and canopy height), which can highlight the characteristic of the new ED model compared with other DGVMs. The evaluation revealed that the global carbon cycle processes were reproduced reasonably by the model, although the soil carbon distribution and its global value are still under debated. Overall, this paper is well structured and documented, and thus I judge this paper can contribute to the journal after minor revisions suggested below.

L26, about "new spin-up": What is new compared to what?

L53: SEIB-DFVM >> SEIB-DGVM

L111-114: it sounds curious to me – this sentence tells the model has been already evaluated and calibrated globally, but this paper, in my understanding, tackles with the same issue. In addition, the reference of Ma et al. (2021) appears to target a specific USA domain, not global.

L161: In my impression, the subsection title "2.2 Model initialization" should be changed to "2.2 Model initialization and overview of experiments", because the second paragraph
refers to transition simulation, not initialization.

L165: Are 1000 years enough to obtain a real equilibrium state in your model? Considering that ED model is DGVM and that the time evolution of vegetation distribution and land carbon amount depend on each other, I would have expected more than a few thousand years are required to obtain a real carbon equilibrium. If not perfectly equilibrated, it may be one of the reasons of your model to make relatively lower amount of global SOC.

L226: This paragraph reminds me one paper: Spafford and MacDougall (2021) GMD reviewed the processes of validation processes of land models which are coupled with ESMs. Your evaluation actually covers critical variables of carbon cycle as performed in the paper, and thus this reference may help to emphasize your rational of your variable choice.

L233: I agree with the idea to aggregate the vegetation categories to compare the simulated result with satellite-derived datasets. At a same time, I wonder showing PFT distribution map with original model category would be helpful for readers and potential model users, when interpreting the simulated result. So, I would suggest putting a map with original model category in appendix.

L295, about “water fluxes” ~: In the subsections of 2.4.1~2.4.4, there seems no description on the reference dataset for water flux.

L327: Latest assessment of global SOC stock of CMIP6 ESMs have been performed by Ito et al. 2020, ERL; global SOC of your model is still within the range of the CMIP6 ESMs, but outside the 1 S.D range. As performed by Ito et al. (2020) and Arora et al. (2020) Biogeoscience, calculation of mean residence time of SOC in global/grid scale would reveal whether such lower SOC in your model is caused by turnover rate of SOC. I believe readers can obtain a benefit from such a bit deeper analysis and the discussion.

L368-374: I agree that ET is an important indicator of hydrology, and it affects terrestrial carbon cycle via soil water availability, etc. In addition, transpiration is tightly connected to photosynthesis. Such importance / purpose to evaluate ET should be addressed in somewhere.

L387, about “less agreement with the references in the tropics”: the delayed timing GPP reduction in tropics (15~0 deg, Fig. 8) seems linked to the delayed timing of LAI (Fig. 14), suggesting there is some problems with the phenology scheme. This is just my speculation, and so further insights on model biases by the model developers would be helpful for readers.
Fig. 16: GEDI L2B product shows discontinuous and large values in 10-15m height in all latitudinal bands, which makes as if ED model underestimate the LAI in the corresponding height. Is it possible to make discussion about whether this is derived from a systematic bias in the GEDI product, or whether it is a certain observational fact?

L429-430: To my knowledge, Watanabe et al. (2011) GMD and Dunne et al. (2020) J. Of Advances in Modeling Earth systems have already incorporated ED-like models (individual-based ecosystem/carbon cycle models) with land-use change impact, into Earth system models.

L459-465: In addition to the overestimation, I’m concerned about the underestimation of several variables in African savanna. When seeing the PFT distribution map (Fig. 3), the region has less vegetation, and GPP, AGB, SOC, and tree height are also underestimated. I’m wondering this model may overestimate bare ground fraction in the corresponding grids. Is this caused by LUH2 scenarios? Or fire impact? Since this paper is a kind of model description paper, further discussion by model developers on the potential reasons for the biases would be much appreciated.

L473-474: Considering that your model reasonably reproduces GPP in boreal forest/Taiga region, NPP reproduced by the model would be well reproduced as well (if autotrophic respiration is not much biased). So, I’m wondering the potential reasons for lower SOC may be caused by less dependency on soil temperature, and/or relatively lower value of base decomposition rate. These are just my speculation, and such further discussion would be helpful for readers to obtain insights on your model behavior.

Considering this model is based on dynamic vegetation distribution, the model performance evaluated here is quite good. I encourage the authors to further improvement in future.