EGUsphere, referee comment RC2
https://doi.org/10.5194/egusphere-2022-49-RC2, 2022

Comment on egusphere-2022-49

Anonymous Referee #2

Referee comment on "FORCCHN V2.0: an individual-based model for predicting multiscale forest carbon dynamics" by Jing Fang et al., EGUsphere, https://doi.org/10.5194/egusphere-2022-49-RC2, 2022

Review of "FORCCHN V2.0: An individual tree-based model for predicting multiscale forest carbon dynamics" by Fang et al. (egusphere-2022-49).

This manuscript by Fang et al. describes the phenology and growth processes for the forest carbon model FORCCHN2, and they evaluate the model at 78 forest sites in the Northern Hemisphere. The paper presents a model that addresses carbon-cycle science questions relevant to the scope of EQU, and the model methods are presented in a reproducible manner. In addition to providing the source code and detailed description for the model, the authors have implemented a module to allow for a seamless integration of the model into a variety of software languages, allowing for user to run model predictions more conveniently with high efficiency.

General Comments

I feel that this manuscript warrants publication based on its reproducibility and presentation quality; however, I have some reservations regarding the significance and quality that I feel need to be addressed.

1) The substantial contribution to modelling science needs to be clarified, as the authors themselves state that the methods are compiled from previous versions with the only new methods being a software module rather than scientific concepts or ideas.

2) They state they apply the model on a global scale; however, they only evaluate the model at 78 Northern Hemisphere sites, which it misleading. While they present maps of outputs across the Northern Hemisphere, none of these large-scale outputs are evaluated against other satellite, modeled or derived products (i.e. fluxes from FluxCom or satellite-derived biomass). Additionally, several of the methods use the hard-wired date of January 1 for exchanges, which is likely not suitable for global use, particularly in the Southern Hemisphere.

Detailed Comments

1) I find calling the model a "tree-based" model to be misleading. First off, with the recent upsurge in literature on machine learning models, my first assumption was that this was a model using machine learning techniques from data, rather than a physical-based ecological terrestrial model. Second, individual trees are not actually being represented. While they do calculate tree height, diameter at breast height, and biomass as well as gap fraction between trees, these are all empirically calculated from LAI using similar methods to other terrestrial models. According to their documentation, it does not appear that each tree is allowed to grow separately, such as in a dynamical model, or that the model is even self-consistent between carbon uptake and growth, but instead is prescribed growth following plant functional type (PFT) equations. Additionally, the authors state that they use PFTs; however, their input data for forest type is a biome map rather than a PFT map. Given that input, it does not seem appropriate to call the model a "tree-based" model when it uses forest biomes such as mixed forest that incorporate not only a wide variety of species but also a variety of forest functional types (such as both deciduous and evergreen) to represent the forests being simulated.

2) In the abstract (line 12) they state that the model can predict yearly phenology, but I'm not sure what that means. Phenology is the seasonal changes in vegetation, so do they mean that the model is capably of predicting inter-annual variability in phenology? This needs to be clarified.

3) It is not clear to me how GPP is incorporated into the growth. The equations for leaf and fine roots growth in the supplemental material do not include any carbon from GPP, does that carbon get allocated out to the carbon pools to make the model self-consistent? If so, how is the carbon allocated? If not, then are the processes really separate and the carbon is not conserved in the model?

4) Why are tree height and DBH only updated annually? Given that the model runs on a daily timestep, why can't all the pools and processes be updated daily? Additionally, where does the change in basal diameter and tree height increments come from and how are these then related to daily growth? Also, how well do these increments match up with annual GPP and what is the allocation of the GPP to these pools?

5) Why does tree death only occur annually? If the NSC pools are updated daily, can't they become insufficient at any point in time? How realistic is it to have the tree mortality occur on 31st December and what are the impacts on the carbon cycle, especially given the claim that this model could be run globally and include forests that would be in the middle of their growing season? In section S1 line 179 states "photosynthate has been allocated to the growth of canopy height and basal diameter", how is this allocated? And how does this tie back to daily growth?

6) In equation S4, which of the (0.5, 1.0) is used in being subtracted from rh and what does that depend on?

7) In the S1 discussion on autotrophic respiration, it states in line 33 that "In Equation S1, t_resp represents..."; however, equation S1 is the GPP equation. Which equation is meant here and where is t_resp used?

8) Why is the NSC updated only once per year, in Eq 2? Why not continually update this? What impacts does the sudden jump cause, particularly in respiration, and are these realistic? Additionally, what impacts are caused by the NSC active pool then being initialized to zero on the first day of the year? Wouldn't this cause a disruption for actively growing forests?

9) In Table S5, where are the allocation parameters used? They are not given symbols and do not appear in the equations from what I could tell.

10) In section 3, the inputs include soil and geography data, what are those and how do they come into play? In the equations in the supplemental material, I only saw LAI used as inputs, how is the soil initialized and what is the geography data used for?

11) Section 3 line 128 states "We can choose four output results", yet I only see 2 listed.

12) Watch verb tenses throughout the entire manuscript and choose a consistent tense for the entire manuscript. It started as present tense, but then switches to past tense half-way through section 3.

13) FLUXNET2015 is used for the site evaluation, which includes numerous options of ER and GPP. Is one of these pre-calculated options used? Is so, which one? If not, why not? And how was the selection of ER and GPP determined? The various methods each have advantages and disadvantages and can lead to substantial differences in the flux

estimates. Additionally, why was the ER set to the night-time NEE? Are the daytime contributions added to this since you are using daily fluxes? I think that the ER being used does include both the night-time and day-time contributions; however, lines 148-151 in section 4 are hard to follow between the three different statements for what ER is.

14) Figure 2 is very tiny and hard to read. Additionally, it appears to be in alphabetical order and has labels for a) through e) that appear arbitrary? It would be more helpful to have these sorted by forest type and/or show the average fluxes per forest type.

15) For Figure 3, what is model efficiency? Also, the colors need to be labeled, which I believe are GPP (green), ER (blue), NEP (tan).

16) In section 4, line 160 states that the model has the best performance in capturing GPP dynamics, what evidence lead to this statement? Later in line 168 it is stated that "predicted ER performed lower than GPP", what is meant by "lower", do you mean that it does not perform as well or has higher errors?

17) Many forests in FLUXNET2015 are annual sinks of carbon due to stand age and regrowth. Since the model uses LAI as input, how well does it do at capturing this? And how well does the model then match or capture the annual growth of biomass and change in fluxes as forests mature?

18) The input data used is described in section 5 lines 183-187, was this the data that was also used for the sites? Is so, this should be much earlier in the methods section. If not, what was the input data at each of the sites?

19) Figure 4 has unreadable font. Additionally, these large-scale carbon fluxes for the Northern Hemisphere should be evaluated in some way to claim that the results are reasonable, and in addition ideally the carbon pools should be evaluated as well. I believe further evaluation is necessary before the conclusions can state in line 210 that "FORCCHN2 was able to predict satisfactory carbon dynamics."

Technical Corrections

1) Section 2, line 71 should be "participating *in* the autotrophic respiration"

2) In the supplemental material, all of the "Where" words do not start a new sentence and should be "where"

3) In S1 line 106 should be "leaf growth *is* based on the assumption"

4) Section 2 line 81, remove first sentence of the paragraph or reword it as it doesn't make sense. (Major control equation of each individual tree.)

5) Section 3 line 112 should be "adapt model runs *to* their"

6) Section 3 line 118 should be "First, *we* installed and loaded"