Comment on gmd-2021-287

Anonymous Referee #1

Referee comment on "Climate and parameter sensitivity and induced uncertainties in carbon stock projections for European forests (using LPJ-GUESS 4.0)" by Johannes Oberpriller et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-287-RC1, 2021

The manuscript by Oberpriller et al. with the title "Climate and parameter sensitivity and induced uncertainties in carbon stock projections for European forests (using LPJ-GUESS 4.0)" studies the sensitivities and uncertainties of a state-of-the-art dynamic vegetation model, LPJ-GUESS, with respect to its parameters and climatic drivers. Although I wished to see more species (afterall the title says "European forests'') simulated, the study already has a quite comprehensive analysis protocol and it reports useful results as is. As the authors argue, the model has changed substantially since its last UA/SA and it is being used widely without these sensitivities and uncertainties being assessed properly. I would have also liked to see more links back to the development efforts, e.g. were (will) the new changes to the model structures (be) justified at the expense of introduced uncertainties? Maybe such aspects of the discussion could be fortified. But overall, I think this is a highly relevant study for the community, it is clearly written and I recommend its publication. I only have minor suggestions (numbers refer to lines):

24: Regarding "predictive uncertainty increases with temperature" I wonder if this is sort of an artefact of choosing relatively northern species. Could it help readers if simulated species are explicitly mentioned in the abstract already? It may also help with the connection to the "stress-gradient hypothesis" (i.e. how you define "harsh environment" may change from species to species).

25: dominates what?

31: This sentence reads a bit unclear, maybe just focus on the latter part (differing in future carbon projections, although note that models also differ for the observable past; Bastos et al., 2020 https://doi.org/10.1029/2019GB006393)

33-35: Could also include initial condition uncertainty (Dietze 2017 https://doi.org/10.1002/eap.1589)

46-47: Although, isn't SA most effective when magnitudes of uncertainty & variability are known, i.e. if one varies something by 1%s, whereas its natural variability is by 10%s, SA could be misleading.

68-75: Are these all for NPP? It would be nice to say which output variables' sensitivities and uncertainties they contribute to, if not NPP.

83: For other models, this may be a good citation here https://doi.org/10.5194/bg-17-2681-2020

2.2. Simulation setup: Could authors add what they did for identifying soil types as well? I believe that's another input to the LPJ-GUESS model.

152-153: Are these (thinning intensity and cutting intervals) the only management parameters that arised from the new management module mentioned in lines 77-81?

179-180: Could authors clarify what it means that they performed simulations as monospecific and mixed stands? I assume they ran LPJ-GUESS with all three species for all the 200 sites and let the model simulate whether the stand is monospecific or mixed, right? But the sentence reads as if they prescribed the model to simulate monospecific and mixed stands (i.e. ran LPJ-GUESS with one PFT only for monospecific sites).

212: Reads better if "respectively" comes before the numbers.

214: Now this note reads like authors pre-defined which of the sites are monospecific and which are mixed.

217: "model output" == is this the average over all years?

217-218: I think it may help clarify if authors state how many linear regressions they fitted in total in the end.

222: Random forest appendix section is A1.3, not A1.2

222-224: Just curious. Did the authors end up discarding any of the sites altogether? I.e. Were there any sites where none of the three species were sufficiently established?

Section 3.1 mean sensitivities: It might be worth noting some opposite behaviour in the sensitivity results as well, e.g. TSB and GPP are negatively sensitive to temperature except for *F. sylvatica*. For NBP, something similar occurs for lambdamax, respcoeff, turnoverroot, krp and emax, amongst the important parameters.

Fig2: I think the alignment of the x-axis labels read better for Fig 2 (centered) than Fig 1. Also please consider sticking to the same order of parameters (x-axis) in both figures. Why is the mixed (*) symbol so much faded for NBP and TSB drivers? Finally, for sensitivity it makes sense (like authors say SA/UA have different interpretations) but isn't it a bit uncommon to visualize "negative contribution" to uncertainty? I mean, the caption says negative relationship not negative contribution, but still with the y-axis it reads confusing.

Fig 3: Radar charts (or whatever you would call them) look almost identical to the eye, having more grids (inner circles) might aid the eye, or could plot an "average" polygon with a solid black line on each for relativity. Alternatively/in addition, for each process/driver the authors could add the highest percentage to the label, better yet if they use the color code (e.g. all "Drivers (X%)" would be orange for mixed). I'm not sure how these would affect the readability but currently the information to ink ratio is rather low on this figure. Finally, "mortality" label seems to have strayed away from the BOR chart.

296: Where did pH come from? This is the first and only time it appears in the manuscript text. It also appears unexpectedly on Fig4. Please clarify.

Fig 5: It could help if panel a bars were also filled with the panel b color scheme (orange, pink, red, light green etc.) for process groupings.

310: I guess one of the interactions in the parentheses was supposed to be radiation-temperature.

334-337: I think this is a very useful sentence in terms of guiding future development efforts. I wonder if authors could dedicate a short paragraph in the discussion (or fortify this one and the next) to summarize recent changes in the model structure and whether

corresponding parameters stuck out in their SA/UA (it could also be important to state if they didn't show up). For example, nitrogen parameters showed up (and they were novel) but among them *nrelocfrac* contributed more. Management module was new but parameters didn't show up in the SA/UA. Do we learn anything from these?

337-338: I'm confused. On lines 71-72 authors said high sensitivities to water-related parameters were found: " Additionally, LPJ-GUESS showed high sensitivity to [...] water-related parameters (minimum canopy conductance not associated with photosynthesis, maximum daily transpiration, Pappas et al., 2013; Zaehle et al., 2005)." Please clarify.

358-359: I thank the authors for explaining the potential cause of the negative effect, but while temperature affects TSB and GPP negatively, how does it affect NBP positively in the model?

360-361: Not just in magnitude but also in direction?

367: Random forest results could be mentioned in the results section.

372-373: Does the finding "nitrogen-induced uncertainty decreases with increasing temperatures" correspond to the general statement of "limiting factors change across environmental conditions"? Or did the authors mean to cite a more specific ecological principle / hypothesis here?

373-377: Since the authors emphasize the stress-gradient hypothesis in the abstract and conclusion, I wonder if they can elaborate more and clarify the reasoning here to convince the reader. I had to read these sentences multiple times and it is still not clear to me how it follows from "decrease of uncertainty contributions of structure-related parameters on the temperature gradient" to the stress-gradient hypothesis which states where the physical environment is relatively benign (harsh) competitive (facilitative) interactions should be the dominant structuring forces. Water, mortality, establishment and photosynthesis parameters' uncertainty contributions also increase on the temperature gradient, which seem to indicate more competitive interactions to me.

378: Unless it was an artefact of the analysis protocol (i.e. lines 341-342).

385: After conducting the analyses, (I know they say their results are robust to these

choices but) were the authors happy with the uncertainty characterizations they initially came up with? I.e. do they still think these would be their best guess at this point or were there any parameters/drivers in particular that they wished they had varied/treated differently, for future studies?

388: Roux & Buis et al. 2021 could be a good citation here
https://doi.org/10.1016/j.envsoft.2021.105046

405: gradient-stress -> stress-gradient

420 & 753: I guess I would avoid language such as "most important countries"

424: Wasn't there any reparameterization in Smith et al., 2014 accordingly?

425: "the productivity of trees in managed forests did not fit to the reported inventory data" where/when was this shown?

445: Was the title supposed to say GPP and NBP? Likewise, check line 765 (Fig A4 caption)