

Review of Merganičová et al., “Biogeochemical model Biome-BGCMuSo v6.2 provides plausible and accurate simulations of carbon cycle in Central European beech forests”

General comments:

This paper describes a method of parameterization and application of the Biome-BGCMuSo v6.2 model for European beech forests and evaluates its accuracy in simulating aboveground wood, soil, and litter carbon stocks. In general, the paper is well-written and offers a useful methodology for setting parameters in biogeochemical models. It is very challenging to model carbon stocks correctly on a continental scale, and the authors have produced reasonable results for a large area in eastern Europe spanning multiple climatic zones. However, there are a few places where the manuscript could benefit from clarification and/or expanded discussion. I would like to see a section that clearly outlines the advantages and disadvantages of the parameterization methodology. It is currently discussed but scattered in various parts of the manuscript. It would also be good to focus the discussion more on general lessons—what processes are most important for accurate simulations, and what’s missing or inadequately represented in the current version of the model— and priorities for future model development. There are also some minor corrections needed as specified below.

Specific comments:

Abstract

Line 28: “model wide-ranging applicability” sounds awkward; rephrase to something like, “applicability of the model over a wide range”.

1. Introduction

The first few paragraphs (lines 52 – 71) contain a lot of generalizations and model development history that are not particularly helpful in introducing the topic of the manuscript. I suggest you cut most of this and focus specifically on the challenges of simulating carbon dynamics in forest ecosystems and different approaches for calibrating parameters.

Line 74: “intrinsic variability” isn’t quite right here, you are describing uncertainty and range that comes from mostly external factors that are not intrinsic to the system

Lines 77-79: awkward sentence, rephrase

Lines 86-87: replace “generality” with “wide applicability”; elaborate on the “advanced calibration designs”—what needs improvement, and what advantages does your approach offer?

Line 89: replace “addressed” with “used”

Line 91: “these elements”—specific what these are (e.g. carbon, nitrogen, etc)

Line 97: delete “the” at end of line

Line 105: not sure what is meant by “phenotypical plasticity”, please explain/elaborate

2. Data and Methods

2.1 Model: This section would benefit from the addition of more details about the model, such as the timestep (daily) and the main inputs (meteorology and soil information). While the next subsection 2.2 mentions the datasets used, it's not clear from the text that these are required to drive the model.

Lines 116-117: It's not clear what is meant by a "two-leaf" representation

Line 128: "storages" should be singular, i.e. "storage"

Lines 130-132: Here you are implying that there are many crop-specific parameters that were added, but are these needed to simulate forests?

Lines 150- 157 and Figure 1: The sites selected for calibration, validation and testing mostly cluster in a specific area of climatic space. How representative are they of the whole European beech population? I know that data availability is a necessary driver of site selection, but something should be said about the limitations of the calibration. Having all countries represented does not seem relevant, as political boundaries are not climatic ones. It might have been better to include more northern sites or sites at the extremes.

Lines 175-177: Why was the relatively coarse-resolution dataset E-OBS used, rather than a newer and finer-resolution reanalysis product such as ERA5? This will affect the quality of the simulation

Lines 223-224: "Under a successful simulation we understand ..." should be rephrased to "By 'successful simulation' we mean the ecosystem existence was maintained from spinup to the end of the normal run ...".

Line 226: "maintain ecosystem existence" is an odd phrase—I assume that unsuccessful simulations involve complete mortality of plants only. I think it would be better to say "ecosystem stability". It's also not entirely clear what you mean by "potential causes"—are these input data, parameters, other model constraints?

Line 245: Why were AbgWC, SoilC and LitterC chosen as the output variables to examine? Why not fluxes, e.g. net primary production (NPP) or net ecosystem production (NEP)?

Lines 252-259: This paragraph is not clear. Provide more explanation of the global SA for readers not familiar with reference Verbeeck et al 2006 and "least square linearization method". Provide a link to RBBGCMuSo package in code and data availability section.

Line 282: Are estimation errors truly independent from each other? If they are not, what impact would this have on your parameter estimation?

Lines 279-292: This paragraph is a little hard to follow. Perhaps move the figures into the main text. Were there any cases where there was no overlap between parameters/sites? It is not necessarily a given that there exists a parameter set that leads to plausible simulations at all sites.

Section 2.5.1: It is not clear from the text what the DT method is or why it is being used here. It seems like a repeat of the steps described in previous sections. Is this an alternative method for selecting parameters? Or does it add to the previous analysis? A figure might be helpful to explain it.

Line 329: what "environmental conditions" did you test? I see it is in the results section but you should also mention specifics here.

3. Results

General: This section would benefit from a figure with a map showing the values of the 3 target output variables over the whole domain

3.1 Parameter sensitivity analysis: the text of this section could be expanded to explain in more detail why some parameters that seemingly have a high sensitivity in Figure 3 (e.g., GRC) were not chosen for calibration, whereas some that were chosen (e.g., Sseff) have a relatively lower sensitivity. Also, I find it surprisingly that all three output variables are highly sensitive to the same parameters. Is there a structural explanation for this?

Figure 4: The spread/distribution in site-specific optimized parameters sometimes cluster far away from the multi-site optimised value. This deserves more discussion.

Line 405: describe and/or provide a reference for the “Wilcoxon signed rank test”

Lines 435-436: You say that CLEC and MSC are significantly correlated. Does this have implications for your method or results?

Figure 7: most sites consistently increase or decrease, but CZ4 suddenly declines around 2015 and then recovers somewhat. Is there a reason for this?

4. Discussion

Line 495: Why would the use of parameters in previous model versions would prevent you from changing them in this one?

Lines 510-511: “Already” should not be used to start the sentence; rephrase

Line 522: insert “and” before “frequently”

Lines 537-540: The discussion of Bayesian calibration techniques feels too brief and out of place here; either discuss in more details the advantages and disadvantages of the technique compared to yours, or don't mention it

Line 543: what do you mean by “hybrid” approach?

Section 4.3: This analysis is good but it covers several different topics (parameter covariance with other parameters, parameter correlation with site characteristics) that should be separated into their own section. I would also suggest making the analysis more focused on the implications for model development beyond this specific application. You touch on this at the end, but do you results suggest that processes are missing and/or represented incorrectly in the model? What are priorities for future model development?

Lines 576: It's unclear how you simulated a varying CLEC. Please describe in more detail (perhaps in Methods section)

Lines 578-584: covariance/interdependency of parameters is an important point and should be discussed thoroughly

Line 607: Using a global database would seemingly contradict the need for a site-specific parameter! I think the lesson/recommendation is to measure these values routinely at research plots.

Section 4.4: This section is too long. I suggest breaking it further into subsections by output variable or process

Line 610: what do you mean by “large scale”? 100s, 1000s of kilometers? Different climatic zones rather than distances?

Line 613: are these statistics among simulated sites? This is where it would be helpful to have a spatial plot

Line 645: Are there other models that do include the process of soil aggregates? Has it been shown to improve model performance? Would this be a priority for BiomeBGCMuSo future development?

Lines 655-669: again what does this suggest about processes that are missing or inadequately represented in the model?

Line 670: should be “Similar” rather than “Similarly”

Line 686: delete “the” before “human influence”

Lines 739-748: This paragraph gets at what should be the key point of the paper, which is what can we learn about missing or inadequate representation of processes that are necessary for a more accurate simulation. This should be elaborated.

5. Conclusion

The conclusion is a little repetitive. I would like to see a discussion of priorities for future model development.

Code and data availability:

The input files used to run the model should be made available if possible, along with the optimization code

Supplement:

Inconsistent labeling of figures—in the text they are referred to as Fig A1, A2, etc but in the supplement they are labelled as Fig S1, S2, etc. This is also confusing with the labeling in Figure 2 to represent the steps in the optimization method (S1, S2, S3, S4). I recommend choosing a different notation in Figure 2 to avoid confusion

Figure S3: increase the size of the axes labels and ticks, it is not legible in its current form

Figure S6: I don't understand why there is more than one simulated and observed point for each parameter set. What is varying? Is it over time?

Figure S14: Missing reference in the caption

Figure S19: Missing reference in the caption. I suggest moving this to the main text as there is quite a bit of discussion around it.