

Review: “Impact of changes in climate and CO₂ on the carbon-sequestration potential of vegetation under limited water availability using SEIB-DGVM version 3.02”

General comments

In this manuscript, the authors perform simulations with the dynamic global vegetation model SEIB-DGVM to explore the impact of historical changes in climate and atmospheric CO₂ concentration on potential carbon sequestration in live vegetation. Intriguingly, they look not just at total biomass, but also “aboveground” vs. “belowground” biomass (although those terms are misleading; see below). This allows the authors to examine how plants have shifted their growth strategies over the last century to maintain a competitive edge under environmental change.

The results show that both biomass pools have increased, but with “belowground” increasing more than “aboveground” on a relative basis. Factorial experiments reveal that atmospheric CO₂ increase is unsurprisingly the dominant driver of potential biomass increase in most of the world, but temperature and other factors are more important at latitudes above 60°N. The results also show that “aboveground” and “belowground” responses to environmental change differ along an aridity gradient, as well as from each other.

The authors designed a suite of experiments well-suited to explore how plant individuals and communities have changed their growth strategies to deal with environmental change. However, the manuscript needs substantial rework. Most importantly, while the Introduction briefly mentions previous findings regarding shifts in above- and belowground allocation under environmental change, this should build up to a set of hypotheses that are then tested with the model experiments. It is also unclear why this was submitted to *Geoscientific Model Development*. Perhaps if it were more focused on comparing SEIB-DGVM biomass to observations it would fit as an evaluation paper, but the work performed is much more high-level than that. I thus think it would be more appropriate to move to *Biogeosciences*.

For these and other reasons that I will elaborate below, I think this manuscript should be ***reconsidered after major revisions*** and ***moved to a different journal***.

Specific comments

Theoretical grounding and hypothesis testing

The experiments and analyses in this paper are well designed to explore how changing climate and CO₂ concentrations, both individually and in aggregate, affect plant allocation strategies over time. However, the authors need to do a better job connecting the two. The Introduction should walk the reader through theory about *why* plants allocate photosynthesized C to different biomass pools. What’s currently there (L 70–84) is insufficient; for example, the reader should understand *why* (or at least, a theory of why) Ma et al. (2021) found what they did. This introduction should then lead into specific hypotheses about what theory suggests the experiments will show. In the rest of the paper, the

experimental descriptions, results, and discussion should continually connect back to those hypotheses. As it is, the paper feels disjointed and directionless, with the results not sufficiently linked to any sort of theoretical framework.

This is worsened by the inclusion of seemingly extraneous information on some figures. Specifically, Figs. 10–12 and A3–4 include extra plots or lines showing the results of every factorial experiment, but I didn't see these discussed anywhere in the text. I can think of reasons why it would be useful to explore the effect of CO₂ and climate drivers in each of these—for the purpose of testing hypotheses, this might not require *every* factorial experiment on each figure—but the authors did not do so. I strongly recommend augmenting the results and discussion to explore the implications of what we see from the different factorial in these figures. If not, the not-discussed results should be removed from the figures (although perhaps the full figures could be put in a Supplement).

The authors should also consider breaking their analyses up based on plant growth form. The allocation strategies of grasses are constrained relative to what trees can do, since the former lack woody stems. Combining the two growth forms in analysis muddies the interpretation of the results.

Some miscellaneous notes I made that are related to my comments here:

- L 395–403: This belongs more in the Introduction than in Results. And it should be explained *why* these changes occur.
- L 412: *Why* does fine root mass correlate with temperature?

“Aboveground” vs. “belowground”

In the last section, the authors note that they consider it a “limitation” that SEIB-DGVM doesn't separate “trunk” (i.e., wood) biomass into stem and coarse root pools. I had actually thought the authors put all wood biomass into the “aboveground” pool intentionally, and that this was an interesting and novel idea! Yes, coarse roots are of course belowground, but these experiments and analyses are designed to explore how plants change their allocation strategies to improve competitiveness in terms of resource-gathering. For that reason, the coarse roots serve an “aboveground” *purpose* that is aligned with the overall goal of wood allocation (at least in theory)—to grow taller than one's neighbors in order to outcompete them for light. (See, e.g., Dybzinski et al. [2011, *Am. Naturalist*] and other work from those authors.) Coarse roots literally support this strategy as they anchor trees into the ground, so any investment in growing taller must also have a corresponding investment in coarse roots, lest the trees become vulnerable to uprooting.

Because all wood biomass—even coarse roots—is in the “aboveground” pool, I think that that and “belowground” are misleading as labels. The authors should rename them to something that better reflects the theoretical purpose of allocating to these different pools. Based on my background and understanding, I'd probably go with something like “light-gathering” and “soil-exploiting” (fine roots are the only belowground pool, and they enable uptake of water and N), respectively. However, the authors should choose labels that are appropriate to whatever theoretical framework and hypotheses they choose to lay out.

While I think this is a powerful and meaningful distinction, its novelty makes it difficult to compare SEIB-DGVM outputs to previous results from the literature. However, that's not necessarily a problem either. Some other DGVMs also don't really distinguish between aboveground and belowground wood in terms of allocation strategy: for example, LPJ-GUESS only makes the distinction (using a global constant for all trees) for the purposes of fire fuel calculations, wood harvest, and transfer of killed biomass to litter/soil pools. The authors should use the literature to classify wood into truly above- or belowground pools in post-processing using even something as simple as a global constant, then augment their evaluation results and experimental discussion to compare to previous findings.

Note that my thoughts here make less sense if SEIB-DGVM actually models tap roots and does so as part of the woody pool. If tap roots aren't modeled at all, this missing process should be mentioned at L 515.

Methods (description) issues

The MODIS NPP evaluation methods are complex and should be moved, along with the other NPP evaluation methods, to a new Methods subsection. (This would also resolve the current problem where the authors start describing the methods, then talk about the results, then finish talking about the methods.) There are a number of issues with the description here (and implications for discussion):

- The authors should clarify exactly what steps they used, in what order, to isolate “undisturbed” land (more on that in the next bullet point) in the MODIS data. Were cells excluded from the NPP dataset before or after aggregation to 0.5°? This should have occurred at the native 500-m MODIS resolution.
- What MODIS land cover classification scheme was used?
- Exactly what land cover types were considered “disturbed?” (The authors only list “undisturbed” types.) Clearly, urban and crop land should have been, but including all grasslands is going to include a lot that is grazed by livestock. To some extent and in some regions, livestock has simply replaced wild grazers, but that's not universally true qualitatively or quantitatively.
 - The authors should mention how including land grazed by livestock might affect their estimates of “potential” biomass. What kind of bias does this result in, and where might it be strongest? Maps of grazed area and grazer density can illuminate the latter.
 - Consider filtering the 0.5° gridcells in the analysis based on some threshold of pasture fraction, which can be obtained for example from the LUH2 land use data. That's actually at 0.25° resolution, so perhaps this filtering could happen at an intermediate step before aggregation to 0.5°, potentially enabling more 0.5° cells to be included.
- Minor notes:
 - L 304: This says MODIS NPP comparison period starts 2000, but Fig. 3 caption says 2001.
 - L 308–310: Simplify (this can be one sentence), and clarify why you're saying this (because MODIS data include used land).

In Fig. 4 and related text, it's not specified over what time period the SEIB output was averaged. This is critical to understanding how it compares to previous findings. SEIB value should *not* cover more than, say, 30 years; then you should exclude literature values outside that range. Or consider instead a scatter plot, with each literature value vs. SEIB *at the time the literature value refers to*.

Other comments on methods:

- It should be clarified somewhere what is meant by “potential” carbon stocks. Presumably this means “in the absence of human land use.”
- L 153–155: Is tree growth daily or monthly? If both, I guess those are different growth processes? Please clarify.
- L 169–176: This paragraph purports to outline advantages of SEIB-DGVM compared to other models, but I think it should just be deleted.
 - L 169–171: I interpret this to mean that SEIB-DGVM includes size-mediated competition for light, but it's not the only DGVM that does this. See, for example, LPJ-GUESS (Smith et al., 2001, *Glob. Ecol. & Biogeog.*) and LM3-PPA/LM4 (Weng et al., 2015, *Biogeosciences*). See also Fisher et al. (2018, *Glob. Chg. Biol.*) for a review of such “vegetation demographic models.”
 - L 171–172 made me think that the simulations would start with PFT composition and structure derived from observations, but later it seems that this is not the case (“SEIB-DGVM simulations begin with seeds of selected plant function types planted in bare ground. The plant functional types are favored for establishment by the environmental conditions in each grid cell.”). Please edit this sentence to clarify that such inputs *can* be used.
 - L 172–174: Unclear what this is trying to say. Is it that SEIB-DGVM can't do land use? If so, that's not an advantage—in models that have land use, it can be disabled for potential vegetation runs if desired.
- Sect. 2.3 (description of relevant processes in SEIB-DGVM): It should be clarified what of this is new to SEIB-DGVM in this paper vs. what was already there.
- L 226: It's a bit surprising that plant demand doesn't actually enter into the calculation of water limitation status. The assumption seems to be that plants are always stressed, to some extent unless, the soil is fully saturated. I guess this is more of a comment about the Discussion: The authors should discuss the implications of this. It would seem to contribute to a bias of SEIB-DGVM towards greater fine root allocation.
- L 277–279: It's very unclear what this test of “detection trends” actually is.
- L 518–524: Unclear. Did you not include N deposition at all? Wouldn't that mean that you've *underestimated* CO₂ fertilization? Please elaborate the N deposition methods (or lack thereof) in Methods and clarify this text.

Results interpretation issues

There are a number of places in the Results where I thought the authors' interpretations were either incorrect, confusing, or insufficient:

- L 338: Slower? Slower than what? CO₂ change looks *faster* than biomass C change.

- L 415: Reference to Figs. 8–9 here is inappropriate, as those maps don’t show attribution.
- L 419–426: Short-term variation is a completely different thing from long-term trend; it’s unclear why they’re being lumped together here (where it says the bit about temporal compensation).
- L 427–429:
 - If there’s no long-term trend in precipitation and radiation (as asserted at L 422–423), how can they induce a long-term change?
 - Precipitation effect appears to *not* be compensatory for “above-ground” biomass, which is most biomass! (Again at L 489–490.)
- Figs. 10–11 and related text:
 - Regression tests for trends in mean and standard deviation across AI bins would be useful. I’d suggest trying a linear fit for Fig. 10 and a quadratic fit for Fig. 11.
 - L 438–439:
 - It seems to me like an increasing trend in this difference would indicate that more water-limited areas experience *more enhanced* C growth, not less.
 - ...Although the trends are very small!
 - This sentence is very confusingly written. “Fluctuations” I think might be the reason. This connotes year-to-year changes rather than a trend, which is what the figures are actually looking at.
 - Discuss: Why is there a (slight) increasing trend for AVBC but a (slight) unimodal pattern for BVBC?

Other specific comments

- Throughout
 - Use of “integrated” and “integral” throughout is confusing. Do you mean “total,” as in AVBC+BVBC?
 - “carbon-stocks” should be “carbon stocks” throughout (no hyphen).
 - “Regions” and “regional” to me imply land masses or geopolitical boundaries, but it is often used here to describe latitude bands. It would be better to use “latitude bands/zones” and “zonal” instead.
- L 49–51: It’s unclear what the difference is between “direct” and “indirect” effects. This idea of “two mechanisms” is not ever returned to, so I suggest just simplifying this sentence to remove the distinction.
- L 73–75: Abrupt transition to talking about models was confusing.
- L 76: “negative response to climate” is vague.
- L 80–82: This sentence is vague (what is “oversensitivity”?) and seemingly unsupported.
- L 182: What is “stock” biomass? This is missing from Fig. A1 and its caption.
- L 215: How frequently? Annually?
- L 238: “adjusted”? What is the usual method?

- Table 1: Please replace the heading “CO₂ fertilization” with “CO₂ concentration” for consistency (referring to environmental conditions rather than plant processes).
- L 282: “We defined”... not really, right? Isn’t this the same as used in Chen et al. (2019)?
- Fig. A2:
 - What is “no value”?
 - Please increase weight of font in legend.
- L 332–333: Is this just repeated from Methods? If so, delete. If not, elaborate (and move to Methods).
- L 337–338: How was 2.44 calculated? Mean annual range? Be more specific.
- L 340–341: Specify R-squared and p-values for AVBC and BVBC as well.
- Fig. 5
 - a:
 - Is the inset plot just the pink line? Clarify this in the caption.
 - Please use a color other than pink, as it’s hard to tell from the red.
 - Please change “Dynamic of biomass carbon” to match the clearer label on the inset plot (“Biomass carbon”)
 - Please add units to Y-axes in inset plot.
 - Consider just removing the inset plot. It doesn’t really add much except potential for confusion. This would also allow zooming in on the biomass Y-axes to provide better visibility.
 - Caption: “during the first decade, ~~the averaged value~~ (1916–1925, red line) and the last decade ~~averaged value~~”
- L 365: “further supports.” Further? Where was this mentioned before? Should be mentioned in Sect. 3.2, where it becomes obvious that AVBC will dominate because it’s so much higher.
- L 366–367: “the proportion of the total change in carbon-stocks is small (3.08 ± 0.14 Pg C)” —what does this mean? “Proportion” makes me think you’re talking about a fraction, but the units are PgC.
- Fig. 6:
 - Colors on map are fine, but cells should be gray (or otherwise distinguished) if p-value is not significant. And then the inset bar graphs should not have colors, because that’s confusing with the map colors; distinguish increasing vs. decreasing trends instead with text.
 - Note different color scale for (c) in caption.
 - Increase resolution so that pixels aren’t blurred.
- L 386–393: What of this is coming from Fig. 7a/c? It’s not referred to anywhere in the text.
- Fig. 7: Include labels on figure for above- vs below-ground.
- L 419: Start a new paragraph here at “Previous”, to provide some separation between pure results and discussion.
- Figs. 8–9:
 - Gray out pixels without a significant trend.
 - Increase resolution so that pixels aren’t blurred.

- I don't think these are actually discussed anywhere except L 415, which I think is inappropriate because they don't actually deal with this aspect of attribution. Add some discussion of them in the Results.
- Might be more useful to replace these with mapped versions of Fig. 7, showing the fraction of the trend contributed by each factor in each gridcell.
- "Modelled AVBC enhanced magnitude" throughout is unclear.
- Figs. 10–11: Specify what a negative vs. a positive value means on the Y axis.
- L 432–435: Remove citations from this sentence. Add actual discussion (in a new paragraph after your results in this subsection) of how your results compare to the literature.
- L 435–438: Combining these sentences would increase readability.
- L 442–446: What does "drivers attributed to increase (A/B)VBC changed" mean?
- L 457–465: This should be in the Introduction.
- L 470: What is "terrestrial water"? Where did you do this?
- L 502: What are "indirect factors"?
- L 502–505: Second part of this sentence seems unrelated to the first.
- L 505: "lowers"? Relative to what?
- L 516: "in factorial simulations"? What does that have to do with anything?
- L 517–518: Why is this a limitation?

Technical corrections

- L 199: Second and third commas should be semicolons.
- L 200: Comma should be a semicolon.
- L 205: "are" should be "is"
- L 213: Tilde should be an en dash.
- L 253: Should be "functional"
- L 265: "trend" is there twice
- L 327: "form" should be "from"
- L 348: Tropicall
- L 382: Comma should be a semicolon.
- L 405: Specify Fig. 7a
- L 423: "variant" should be "variation".
- L 453: Delete "that"; "spatial" should be "temporal"