April 3rd, 2023
Dear Biogeosciences Editor and Reviewers,

We appreciate your review and comments on "Exploring the impacts of unprecedented climate extremes on forest ecosystems: hypotheses to guide modeling and experimental studies". Your time spent on this peer-review process is appreciated. Below is a point-by-point response to the reviews, and a description of how the manuscript has been revised accordingly.

Key:
Gray text: comment from reviewers (e.g., RC)
Black text: response from the authors
Red text: edits or new text added to the manuscript

Comments from Referee #4: "The corresponding author have addressed my concerns and I'm happy with them."
Author response: We are pleased to hear that this reviewer was happy with the latest submitted revisions, and was happy with the changes.

Comments from BG Editor:

RC: Your overall aim, to provide a perspectives paper, is still hampered by a focus on details on the two individual models, making it hard for a reader to see the general points you are trying to make behind the results of the two models. Furthermore I find that in some places the framework and pathways to future research are not very clear.

Author response:

- Thank you for this point that there is still too much focus on individual model results, when the models are intended to be a guide to investigate unknowns and areas for improvements. At the beginning of the results section, we removed the text that validated the models to observations, and moved these results to the supplements.
- We tried to also reduce the emphasis regarding the diverging responses between the individual models that one happened to be sensitive to drought intensity, while the other model was sensitive to drought duration. We do not want to claim that one response pattern (intensity vs. duration) is "correct" over the other, but instead there are a lot of uncertainties with representing future extremes (conveyed by the increasing amplitude and shading in panel 1d in our conceptual Figure 1).

RC: Regarding H1: Why would you assume that the response to drought would be linear? We know that stomatal responses to soil moisture are not, so why should the response to extremes be linear?

Author response:

- We agree that response to droughts are not linear, and this linear hypothesis option is not necessary. We have removed any instances of a linear response, and only describe UCE impacts as varying degrees of non-linearity and threshold responses.
- Figure 1 has also been updated. We removed the null hypothesis that responses might be linear, or near-linear.

RC: Most of the results shown and mechanisms discussed refer to UCEs in water availability and drought. I think it's important to point this out in Abstract, Intro and Conclusions, because the manuscript does not address concepts applicable to for instance heat stress, cold snaps etc.

Author response:

- We updated the manuscript to more clearly point out in the abstract, introduction, and conclusions that this paper is *only* discussing UCEs related to water availability and drought, not other extremes.
- Line 41: "Here, we present a road map of how two dynamic vegetation demographic models (VDMs) can be used to investigate hypotheses surrounding ecosystem responses to one type of UCE: unprecedented droughts."
- Line 138: "While a variety of UCE-linked biophysical tree disturbance processes (e.g., fire, wind, insect outbreaks) can drive nonlinear ecosystem responses, we focus specifically on extreme droughts, which have important impacts on many ecosystems around the world (e.g. Frank et al., 2015, IPCC 2021)."

RC: You suggest that you develop an iterative framework. However, it remains unclear, in which respect or how the framework is iterative, which I read to refer to a loop of experiment - model development & testing - experiment -model development and testing and so forth.

Author response:

- We updated the manuscript to make sure we are not claim that we have developed an iterative framework currently in this manuscript, especially since we only include modeling testing and do not introduce experiments here. But that an iterative framework is suggested for future work, and is a useful tool for understanding responses to extremes.
- Line 146: "This study can help guide how the scientific community can iteratively address these questions through future experiments and modeling studies."

RC: Why are VDMs needed for this exercise? Could simpler models not also show similar behaviour? Briefly motivate why these two models were chosen to exemplify your framework.

Author response:

- We intentionally choose to use VDM for this exercise of investigating extremes due to their representation of ecosystem structure, demography, and capturing competitive responses from disturbances. Simpler models fail to mechanistically represent mortality, recruitment, and disturbance – each of which influences biomass turnover and carbon (C) allocation (Friend et al., 2014) – and thus non-VDMs are limited in their ability to realistically forecast ecosystem responses to anomalous environmental conditions like UCEs (Fisher et al., 2018).
- This is further described and justified on page 4 in the introduction, starting at line 110.

RC: Since your paper is meant to be a perspectives paper, a brief elaboration of why you chose these two sites out of the many possible as representative would be warranted.

What are you trying to demonstrate with these two different sites? Why not only use one, if the disturbance simulated anyway do not correspond to the experiments at the site? I do not get the sense that the use of more than one site adds a lot of information to the manuscript that is accessible to the reader and serves the purpose of a perspectives paper. A little more explanation would be beneficial here.

Author response:

- In Section 2.2 of the manuscript, we have the following text describing the two site selections. "To exemplify how VDMs can be tools to explore new hypotheses related to UCEs we applied the models at two field sites, that were chosen due to being extensively studied and the models used here have already been run at these sites and previously benchmarked against field data (see Xu et al., 2016; Medlyn et al., 2016; Medvigy et al., 2019 for model-data validation)….. In addition, the two sites span a range of vegetation types and are in warm, seasonally dry climates that are more likely to experience droughts in the future (Allen et al., 2017)."

RC: Since your paper is meant to be a perspectives paper, the first paragraph in the Results section can be moved to the supplement. Since no data is shown to evaluate the response of the models in the light of their initial performance (and this is not a MIP-paper), I feel this paragraph is a distraction.

Author response:

- We agree with the reviewer, and these results have been moved to the supplements in the revision. A previous reviewer requested more site level validation and comparison to observations, so we included this observational data. However, we agree and think it can be in the supplements.

RC: While I understand that a in-depth discussion of the different responses of the two models is beyond this paper, it is somewhat dissatisfying that you emphasise the difference between the drought intensity and duration response of the two models without giving a clear hint as to the source of this. Either there is an explainable difference that can be summarised briefly in Section 4, or it should not be highlighted, because it remains unclear why this difference occurs. At least one could re-iterate the key differences between the models that contribute to this behaviour, rather than simply listing all possible causes that may or may not be included in the models.

Author response:

- We thank the reviewer for understanding that an "in-depth discussion of the different responses of the two models is beyond this paper" and also not the scope of the paper to be a model intercomparison. We've re-wrote parts of the abstract and conclusions to try and de-emphasize the diverging responses between the models that one happened to be sensitive to drought intensity, while the other model was sensitive to drought duration. We do not want to claim that one response (intensity vs. duration) is "correct" over the other, but instead there are a lot of uncertainties with representing future extremes (conveyed by the increasing amplitude and shading in panel 1d in our conceptual Figure 1).
- We instead try to emphasize that UCE impacts can be extremely variable, unknown interactions, and we still have some work to do in order to narrow down the potential plausible scenarios until

we have better grasp on processes surrounding plant hydraulics and stress, carbon allocation, large tree mortality, diverse community composition, etc.

- We updated the abstract to include these sentences: "The severity and patterns in biomass losses differed sustainably between well tested models. For example, biomass loss could be sensitive to either drought duration or drought intensity depending on the model approach. This is due to the models having different, but also plausible representations of processes and interactions, highlighting the complicated interactions and variability of UCE impacts still needed to be narrowed down in models."

Minor comments:

Abstract:

RC: L 41: remove reference to _two_ models. Why would your roadmap only apply to these two, and what interest would a reader of BG have if this manuscript is only relevant for these few model developers?

Author response:

Done. We have removed the reference to "two" models.

RC: L50: Complicated sentence, consider rewriting.

Author response:

Done. We have rearranged, and re-written, the sentences from line 50-54 to make more sense.

RC: L61: It would be preferable if the abstract would focus more on what this framework contains than individual model results that are only a guidance to develop the framework

Author response:

This is a good point, and we agree with the reviewer. Since this was not intended to be an in-depth model comparison paper, we removed one of the two times we mentioned in the abstract that the models reported either a stronger response to drought intensity vs. duration. Instead, we would like to emphasize that these different results are both plausible based on our current understanding of processes in these models. I.e., we can't narrow down the mechanisms (listed in the discussion) until more experiments can be done to explore the unknown UCE interactions/impacts.

RC: L104: drop reference to IPCC here. IPCC only assess existing science. CMIP6 is the framework in which these model runs were done. Ciais 2013 is not a recent IPCC assessment. Consider citing IPCC AR6 or IPCC SRCCL here.

Author response:

- Done.

RC: L348: Integrated carbon loss. I am struggling with the term, because it's a complicated mixture of intensity and duration of the response. A loss should simply have a unit of

kg/m2, but this is more a severity-of-impact index, and I recommend to call it that (or something comparable. Note that the unit varies in text and tables between kg C m-2 yr and kg C m-2 yr-1. Please check and clarify.

Author response:

- We understand that there might be some confusion around the integrated carbon loss term, because as the reviewer points out it's a combination of disturbance intensity level, and the duration of the recovery period, integrated over a time period. Not just a one-time carbon loss. Therefore, we have changed the name from integrated-carbon-loss to "severity-drought index" when reporting the carbon loss from drought events only. And changed the name from integrated-carbon-change to "severity-climate index" when quantifying the difference of the effects of climate change plus droughts compared to just drought alone.
- We have also corrected the discrepancy in units to always be kg C m-2 yr, not per year.
- We updated Figures 1-3 so that the axes now have the correct metric term of either severity-drought index, or severity-climate index (kg C m^{-2} yr). The figure captions have also been updated to try and explain this severity metric more clearly.

RC: L456: Use of SO and MO not clear here, please re-iterate

Author response:

- We agree that the beginning of this sentence was not clear, and also probably not needed. We updated the sentence to go straight into explaining the VDM model responses with regards to SO and MO patterns.
- Line 449: "When comparing VDM responses to increasing drought severity and its interactions with warming and eCO_2 (related to conceptual Fig. 1d), ED2 showed a more consistent MO response during UCEs and with additional warming and eCO_2 ....."

RC: L789: Check grammar: does one "do lists"?

- Corrected.

RC: Table 3/3 unclear what "ratio-based optimality partitioning theory is meant to refer to.

Author response:

- In term of carbon allocation strategy in plants, ratio-based optimal partitioning theory is allocation to plant organs based on the most limiting resources (McCarthy and Enquist, 2007), which is described in the text on line 574. Based on other reviewer comments we have suggested different optimality theories to potentially explore in models as ways to decrease the variability and unknowns in model results.

RC: Figure 1: The shading in 1d are non-trivial, but also not well explained. Please clarify in figure legend or text.

Author response:

- Thank you for pointing this out about Figure 1d. We have updated the methods text (line 206) to include: "Additionally, more climatic variability from unprecedented eCO_2 levels and warming will lead to unknowns in how ecosystems are affected in the future (i.e., the widening, and downward shape of the shaded areas compared to historical, Fig. 1d)."

We updated the figure caption to include: "We conceptualize how oscillations between SOs and MOs could be amplified and the widening of the shaded areas represents increased variability in how unprecedented eCO_2 levels and temperatures will affect ecosystems in the future compared to historical."

RC: Figure 4-5 are only referred to once in the discussion. I don't really see the need to include them in the main text

Author response:

- Figures 4-5 are used to help explain the wide range of variability when trying to represent demographic ecosystem response to UCEs. These figures help to make our point that there are different potential patterns that still need to be investigated when dealing with complex extremes, and which specific responses need to be explored more (i.e., disagreement between models) or not explored (i.e., better agreement between models). For example, both models indicate that large trees will be impacted the most by extreme droughts, agreeing with observations (reviewed in Discussion). But the models disagree on the recovery of different species composition. We hope because of these points that Figure 4-5 can remain in the main text.