Dear Biogeosciences Reviewer,

We appreciate your review and comments on the following paper “Exploring the impacts of unprecedented climate extremes on forest ecosystems: hypotheses to guide modeling and experimental studies” by Jennifer A. Holm et al., Biogeosciences Discuss., https://doi.org/10.5194/bg-2022-65-AC2, 2022.

In general, we think clearly stating at the beginning of the manuscript that this study was not intended to solely be an original research paper, but rather a review of how models currently capture unprecedented climate extremes on forest ecosystems by using a handful of novel model results, will alleviate many of the reviewers' concerns. In order to fully investigate responses to such novel climate extremes that have not been experienced, we had to generate original model results which we believe is a strong novelty of this paper.

**R2 has concerns about:**

- **R2:** It is not clear whether the main focus of this study is to introduce the newest model development or review what is missing in the current VDMs and suggest future directions, or even the combination of both.

  Answer: We will rewrite the manuscript to make it clearer that the goal of this paper is to highlight what is missing in current VDMs, even after taking into account cutting edge model developments. In order to review what is lacking in current VDMs, we needed to describe the existing model frameworks and latest model developments (which might have been a little confusing). We will make this more clear in the revision.

We will also revise the manuscript to state more clearly at the beginning that this is more of a review and “guidebook” manuscript, but with a combination of an original research in order to address novel forest responses to climate extremes. We believe this combination of original research with review of current limitations of models is a strong novelty in and of itself.

2. **R2:** Lack of model validation and comparisons with the observation/data, and validating against some historic drought events at these sites.
Answer: We agree that model comparison to some site-level observations is a worthy improvement to the paper. Upon revision of the manuscript, we will include site level benchmarking observations such as biomass, stand density, and leaf area index to do basic level model validation. With regards to validating the models against historic drought events at these sites, as also pointed out by reviewer #2, it is not easy to track the effects of a drought event in reality when these events can be rare, field campaigns didn’t occur during the event, all while controlling for other factors. The goal of this paper was to investigate ecosystem responses to extreme droughts that will likely be occurring in the future, and how they are different from less intense droughts as the responses can be nonlinear.

3. R2: LPJ-GUESS showed large swings in LAI at EucFACE site, and R2 is wondering if this is reasonable?

Answer: In the revised manuscript, we will work to provide observational data of leaf area index (LAI) at the EucFACE site to compare against the large swings from the LPJ-GUESS model.

4. R2: Question about why these two specific study sites are selected? (And why no site level tuning was done).

Answer: We choose these two sites (the Australian and Costa Rica sites) mainly because the models have already been run at these sites in previously published papers (Xu et. al., 2016; Medlyn et. al., 2016), thus making model application possible with a “built in” reasonable degree of validation. These sites also contain multiple measurements and information that allowed for previous papers to validate the models. These sites were also chosen because they span an interesting range of vegetation types (a temperate-subtropical transitional forest in EucFACE and a seasonally dry tropical forest at PaloVerd), and are in warm, seasonally dry climates that are more likely to experience droughts in the future. In the revised manuscript we will include a description of why the sites were chosen, reference the publications where model validation has already occurred, and include this in Section 2.2 where we already initially described the sites. We will also emphasize that the purpose of this paper wasn’t to do site comparisons between many different sites, but just select a few for hypothesis testing.

5. R2: Would like to see a tighter connection between the simulation results and discussion section which describes what is lacking in VDMs, by providing more detailed descriptions of the two models.

Answer: In the revised manuscript, we will improve the text about tighter linkages between concepts, hypotheses, and model outcomes. For example, we want to make better connections that large losses in carbon were likely linked to uncertainties in how we currently represent plant hydraulics, non-structural carbohydrate storage, or phenology diversity, which we then reviewed in the discussion section. We will also revise the manuscript to emphasize that a goal of this paper is to demonstrate how to use these VDMs in order to help generate future hypotheses about UCEs. Therefore, we used the models and sites as conceptual “experimental” tools to investigate the given hypotheses. In the supplemental material we will do a better job of referencing previous papers that have done more through investigations of specific model processes within each of the models, and listing which processes or mechanisms were evaluated and why.

We are glad to hear that R2 thought the paper was “clearly written and will be an interesting topic to the readers of Biogeosciences.” As well as, “In the discussion, what is missing in the current VDMs are thoroughly reviewed”, which as we listed above in our answer to concern #1 is the goal of this paper.