Responses to R. Fisher's comments on Weng et al. manuscript (LM3-PPA model)

Thanks for the comments and suggestions provided by R. Fisher, which were extremely helpful in guiding us to a revised paper that more clearly communicates the scope of the work and its importance. Our responses are below (*Reviewer comments are in italics*, and our responses are in roman):

R. Fisher (Referee)
rfisher@ucar.edu

This paper is quite clearly an exceptional study, summarizing a huge quantity of model development effort. I suspect it will become an important part of the dynamic vegetation modeling literature going forwards. The authors report the initial single-point testing of the coupling of the PPA model within the GFDL Earth system model, specifically including the fine-scale and biogeochemical processes from LM3 into the 'mathematically tractable' Perfect Plasticity Approximation model, the properties of which, regarding the elucidation of competitive and evolutionarily stable strategies, have previously been discussed by Farrior et al. (2013, in review).

Here, Weng at al. show for a site in the Eastern United states, that the model is capable of simulating carbon uptake, biogeochemical processes and succession all with some skill. They also describe a set of experiments that illustrate that the model can distinguish, by virtue of its competitive interactions driven by campy light interception, that it predicts different outcomes for polyculture/competition than for the species that maximizes productivity in isolation. This is an important result, as it further demonstrates, I think, that the use of 'many monocultures' in traditional DGVMs does not necessarily produce the correct competitive outcome.

The paper is extremely well written, and the explanations of emergent ecological phenomena are some of the best that I have come across, despite the often complex nature of the subject matter. I have few if any major objections to this paper and think it could be published in its present form. My main points are mostly suggestions that might increase the clarity of the text.

I do not agree with one other referee that the introduction is overly grandiose, since it is made quite clear at the end that this paper does not pertain to global simulations and is in fact a first step towards them. I do agree somewhat that the use of 'Earth System Model' in the title rather implies that the ESM capabilities have somehow been deployed in this paper. Clearly, implementation of the model within the code of the ESM is a hugely challenging task, and should *be acknowledged somehow. Maybe "A mathematically tractable method for scaling individual trees to ecosystems within Earth System Models" would work? Or something along those lines.*

We changed the title to "Scaling from individual trees to forests in an Earth system modeling framework using a mathematically tractable model of height-structured competition". The key changes in the title are:
"individuals to ecosystems" is now "individual trees to forests" to clarify that the paper focuses on forests
"Earth system model" changed to "Earth system modelling framework" to highlight that we have developed an approach, but have not yet deployed the new ESM. We realize that this point may still not be 100% clear from the title, but we have now clarified this point in several places in the paper to minimize the potential for confusion.

Generally, given that the PPA uses a more empirical information in the parameterization of, for example, its canopy and understorey mortality rates, than a standard BGC model might, I was often left wondering which elements of the model testing process were genuine tests of the emergent properties of the model, and which derived from the model being 'told' the answer in advance. I think that a greater transparency (or maybe, just illustration) of this to the reader wold be helpful, since I do not think that the specific outcomes of the simulations for the Eastern US are the main interesting feature of this paper, given that all such models are prone to extreme parametric uncertainty, which would most likely comprise the main difficulty in scaling up to larger regions or the entire globe.

We added two paragraphs to Section 2.2 (Model evaluation and simulation tests), to clarify how we tested the model, and which aspects of model behavior were tuned. In brief, we tuned several physiological parameters related to photosynthesis and respiration to yield realistic predictions of NPP, mortality rates of trees smaller than 0.1 m diameter, and a single taper constant that affected the diameter growth rates of all trees. However, we did not tune tree size distributions, successional dynamics, or variation in tree growth rates among species or canopy layers. Thus, these latter types of data all provide tests of emergent model behaviors that were not tuned. These points are now explained in detail in the first two paragraphs of Section 2.2. In addition, we also added a sentence to the first paragraph of Section 4.2 (Model evaluation) in the Discussion (Lines 802~804 in revised MS): "These comparisons must be evaluated in light of the tuning of the physiological model to produce observed NPP, the tuning of a single parameter affecting diameter growth, and the tuning of the elevated mortality of seedlings and small saplings."

Specific Comments and Questions
17770 L3: While it is clear that the flat top assumption is very beneficial from a simplifying perspective and therefore appears justifiable, it is of course not physically realistic. What do you lose, in theory, by adopting it? Are there any downsides that might be considered, with regard to simulating co-existence, etc., that might cause this assumption to be revisited in future versions?

We agree that the assumption of flat crown tops seems unrealistic, but it has worked well in previous empirical tests of the PPA model (see Lines 330~336 of revised MS), and it greatly simplifies calculations and speeds computation. However, incorporating the PPA into LM3 did reveal a potential problem related to the flat top assumption, which we have now made this clear on Lines 342~350:
"In LM3-PPA, the assumption of flat-topped crowns introduces a potential problem that does not occur in simpler versions of the PPA model that lack physiological mechanisms. Specifically, the NSC pool can, in some cases, be quickly consumed when a tree enters the upper canopy layer from the understory because of the sudden increase in target leaf and fine root biomasses. This increase would be more gradual with other crown shapes (e.g., rounded). To address this problem (which we view as a model artifact), we introduced a parameter to limit the rate of increase of target leaf mass (and therefore fine-root mass, given the pipe-model constraint) for cohorts that recently entered the upper canopy (see Equation A6 in Appendix A)."

17772 L17: Does the NSC target scale with biomass, leaf area, etc?
The NSC target scales with target leaf mass (crown area * target crown LAI * LMA), as specified in Eq A2.4 in Appendix A. To make this clear in the main text, we added the following on Lines 401~403: "…and a species- or PFT-specific NSC target, which scales with target leaf area and tracks a plant's phenological state (Eq A2.4 in Appendix A)." We also added additional clarifying text below Eq A2.4 in Appendix A.

17773 L7: Does mortality of the entire cohort occur immediately if the average NSC for that cohort drops below zero, and if so does that not cause unrealistically 'thresholdy' behavior in the model? Perhaps the cohorts are large enough in number that killing a whole cohort does not produce a large deviation to the model state? Otherwise, I would imagine that some kind of continuous function relating NSC to mortality might be appropriate.
Yes, the cohorts will immediately die if their NSC approaches zero. But this rarely happens in LM3-PPA in its current form, except under prolonged drought (which was not simulated in the current paper). Even in the understory, we rarely observed starvation mortality in our simulations, which may simply reflect our parameterization of mortality, which attributes high rates of mortality in small size classes to "background mortality" (Fig. 1a), with "starvation mortality" occurring in our model only if NSCs drop to zero.

In our parameterization, target NSC (NSC*) is around 3-4 times the size of the combined target leaf and fine-root masses (Eq A2.4). So, trees usually have sufficient NSC to produce their target amounts of leaves and fine roots (Eq A2.6). When actual NSC is lower than NSC*, stems stop growing (Eq A2.10), and the NSC is only consumed by maintenance respiration. If a tree doesn't have carbon to sustain maintenance respiration (NSC = 0), it will die. This is the logic of the model formulation behind these equations. In our simulations, starvation (NSC=0) is restricted to the most-suppressed understory cohorts. Because these inevitably have small biomass and do not shade taller cohorts, the loss of an entire heavily-suppressed cohort has little effect on other predictions of the model. We inserted the following (Line 418~422 in the revised MS) to briefly clarify these points: “Because the target size of the NSC pool is assumed to be several times the size of the combined target leaf and fine-root masses (see Eq A2.4 in Appendix A), trees rarely die of carbon starvation unless they experience prolonged drought (which was not simulated in the current study) or have chronic negative carbon balance due to shading.”

17773 L22: Does the PPA apply to all the vegetated areas, including cropland and pasture? It isn’t clear here.
In this paper, we only include trees, although we have developed parameterizations for grasses, forbs and shrubs that will be published in forthcoming papers on a global implementation of the model. We explain these points in a new Discussion section: “4.6 Future challenges”.

Although beyond the scope of this paper, the model can currently be run with the old LM3 structure for croplands and pastures and the new multi-cohort structure in non-agricultural tiles. The final version will have the multi-cohort structure in all tiles, but this will have no effect in croplands which actually have only one cohort.

17773 L10: How is mortality a function of size in the understorey? I think this needs fleshing out or deferring to the appendix.
We added text in the Mortality and Disturbance section that explains this and gives the equation (see Lines 426~429 in revised MS). This equation was in Table 2 (the column of μ_{U0}). We deleted this column and put this equation here.

17766 L30: This paper does a very good job of carefully explaining the theory behind the PPA, but I think that, as written, this particular section risks alienating large fractions of its potential readership, and will decrease its impact accordingly. The ED model code -actually- discretizes the cohort and patch properties, and thus is realized as a relatively simple set of differential equations determining the growth of different biomass pools and mortality of the different cohorts, etc. It is therefore much easier to understand than the more abstract descriptions of this theory that exist in the literature.
This is something of a style point, but I find that the use of this type of language in model description papers, borrowed from more physical disciplines, does more harm than anything else in making demographic model theory less accessible to those who might wish to understand it. This barrier of understanding is a genuine problem, and is most acute when trying to communicate how models abstract the real world to scientists from fields that do not make use of such principles and to land surface scientists who do not habitually conceptualize ecological processes in this way. Is it possible to rephrase this section to reduce the likelihood of losing the audience at this rather critical point?
We added considerable text to the Population Dynamics sub-section in Section 2.1 (Lines 253~261) to make it more user-friendly. The text was obviously unclear before because the

algorithmic simplicity that the reviewer describes in the ED model is precisely what is specified in the three population dynamics equations (1-3). We do feel strongly that these three equations need to be in the main text, because they are the foundation of the method.

17774 L8: Given that this paper doesn't pertain to managed forests, this seems like a little too much detail, and also raises the question of why age-since-disturbance dynamics are operational for managed forests and not for the natural vegetation (my interpretation of how this operates at present). I think discussing the managed forest component is confusing if it isn't presented at all in the simulations.

We moved this most of this section (**Subgrid-scale heterogeneity**) and the subsequent section (**Land use change and ED gap approximation**) to Appendix A and added a brief explanation of the tile dynamics to the main text (Lines 438~444 of the revised MS): "Like LM3, LM3-PPA model is implemented on a flexible grid, whose cell size can be specified independently of the atmospheric model's grid. LM3-PPA also includes LM3's dynamic tiling scheme for land use, stand-level disturbance, and subgrid-scale heterogeneity (Shevliakova et al., 2009). As explained in Appendix A, the tiling scheme can be used to implement the ED approximation for canopy gap dynamics (Moorcroft et al., 2001), but this feature was not used in the simulations presented in the current paper."

17775 L15: In this discussion of ED, I think it would be good to state more clearly what is happening in the PPA at present - i.e. 'we assume that there are no stand replacing disturbances', or that 'we do not run the model to old-growth equilibrium, on account of the absence of disturbance dynamics'. Otherwise it isn't clear what once might stand to gain from adding in ED-like dynamics. Also, why does this section only seem to propose using gap-phase disturbance dynamics for secondary vegetation and not for primary vegetation?

We removed this section from the main text to avoid confusion because these issues do not affect the model presented in this paper. However, we include the discussion of ED in Section 5 of Appendix A because the design of any model intended for an ESM is constrained by the need to address issues of sub-grid-scale heterogeneity. We have now clarified in the revised MS that in the current work, we do not implement stand-level disturbances (Lines 435~437) or the ED gap approximation (Lines 441~444). In addition to the main text, we have also clarified these points in the new Section 5 of Appendix A.

We did not mean to imply that gap-phase was relevant for secondary but not primary vegetation. The tiling scheme can be used to implement land use and various forms of disturbance, including the ED gap approximation (see response to previous comment, above). We believe that the ED gap approximation would likely make the biggest difference when modeling old-growth forests (Lines 932~946 and 547~550 of the revised MS).

17775 L18: Why choose these N American species in particular? What is the general purpose of this illustration?

The Wisconsin, USA site was chosen because there are multiple data sources in the region that facilitate model calibration and testing; and because there are clear, well-understood patterns of succession in this region among the three selected species, which allows us to evaluate (as a proof of concept) if LM3-PPA can correctly predict forest succession. We have added two paragraphs at the beginning of Section 2.2 (Lines 455~480) to explain the rationale of the model evaluation, including the choice of region and focal species.

17777 L20: It isn't clear to me here where the extra carbon goes, if it is not allocated to dbh increment, does it go into roots, or storage, etc?

As a result of parameterization of allometry and allocation, the growth rate of DBH is independent of DBH when crown area is proportional to $DBH^{1.5}$. In experiment H2, we stop the expansion of crown area when DBH exceeds a critical value (0.8m), but we maintain the same rules of allocation of NPP and the same allometry of DBH vs. wood biomass. Thus, carbon in excess of maintenance goes exclusively to stem growth and reproduction. Because the GPP of a tree is proportional to its crown area, GPP in experiment H2 no longer increases with DBH beyond a DBH of 0.8 m. As the tree continues to grow, the same GPP must be "stretched" around an ever-growing stem, and so the DBH growth increment must decrease as DBH increases beyond 0.8 m.

17778 L12: I think you should state the purpose of this comparison with the BGC models here? Is the intention to have the BGC processes mirrored exactly, or to more generally test the implications of modeling succession on carbon accumulation?

We added text to the beginning of Section 2.2.3 (Lines 560~563) to clarify the purpose of the comparison. The first sentence of this section now reads: "To explore how incorporating individual-level competition and successional diversity into land models affects carbon accumulation in vegetation and soil, we compared the LM3-PPA predictions to those of a CENTURY-like standard biogeochemical (BGC) model…"

17781 L24: Is the understorey aspen mortality parameterized, or emergent from the properties of the carbon starvation model?

Carbon starvation rarely occurred in our simulations, perhaps because of the way we partitioned mortality between starvation (NSC = 0) and "background" (i.e., high mortality rates in small size classes are attributed to "background mortality" in our parameterization; Fig. 1a). High understory mortality of aspen in our simulations is due to a combination of (1) its high background per-capita mortality rate (which is parameterized: see Table 1 and Eq 8 in the revised MS); and (2) growth suppression due to shading, which is an emergent property of the dynamics of size structure in the model, and which maintains understory cohorts in small size classes where they suffer high rates of mortality rates. We have clarified these points on Lines 645-652 of the revised MS.

17792 L20: What is the difference implied here between computational and mathematical tractability? If alternatives are computationally intractable, that implies that they are impossible to compute with given computational resources, which would be hard to demonstrate. Maybe computational efficiency rather than absolute tractability per-se, might be a better term to use?
We agree with the reviewer's suggestion, and the new version of this sentence now reads: "Because of the tractability of the PPA, the coupled LM3-PPA model is computationally efficient (relative to existing alternatives to modeling height-structured, individual-level competition within ESMs) and retains close linkages to mathematically tractable special cases (e.g., constant climate)."