Reviewer 2 asks us to do a detailed analysis of the comparison between the full demography and single cohort settings, test the sensitivity of PFT key parameters, improve the description of phenology, mortality, and clarify some model behavior. We can do these in a revised version. Detailed responses are below:

The authors present a model development work on vegetation demographics and seek to implement it into an Earth system model. The new model features include a greater diversity of global plant functional types and a new phenological scheme. They also compare the behaviour of this model to the eight observational locations and six MsTMIP simulations. As seen from the results within the paper it is possible to capture vegetation structure and dynamics reasonably even within such a parsimonious model at a global scale. Overall, this work is well structured and is useful in highlighting some new issues in improving the representation of the terrestrial carbon cycle within ESMs. I think the manuscript could be publishable after some major revisions. I have a few comments below.

Thanks!

Major comments:

- One of the most representative feature of this model is the full demographic processes. The authors mainly compare the differences between the simulations of the full demography and the single cohort settings of BiomeE. I think authors should separately compare simulations of these two versions of BiomeE with observations, if possible, to show the advantages of full demography in reproducing ecosystem dynamics (Figs. 6-10).

We will do more analysis for the single-cohort model in the revised version if we are asked by the editor to resubmit this manuscript.
- **Lines 157-158:** "A set of continuous plant traits are used to define the distinctive plant types. Please specify the continuous trait assignments for plant functional types, especially the differences with traditional PFT-based model.

They are in table one. We will explain it clearly in the revised version.

- **To represent the major variations in plant functional diversity,** the authors chose four plant traits as the primary axes to define PFTs: wood density, leaf mass per unit area (LMA), height growth parameter, and leaf maximum carboxylation rate (Vcmax). I would suggest a sensitivity analysis of these plant traits to different ecosystem functions, which would be very instructive for further model improvement and localization.

I agree with this suggestion. We will do the sensitivity analysis at one site to show model behavior.

- **Methods Line 184:** it is unclear to me about the assumptions in the phenological scheme. Why to define the nine PFTs as these four phenological types? Regarding the comparative advantage and competitiveness of deciduous vs. evergreen trees, are there any basic theories that evergreen species are more resistant to cold and drought than deciduous tree species? According to the global vegetation distribution, evergreen broadleaf species are usually distributed in warm and moist environments. What kinds of functional traits suggest that evergreen species are adaptive under water limitation and cold conditions?

We only defined the possible factorial combinations of drought and cold deciduous, but did not discuss who will be more competitive. It is possible that the evergreen would be more competitive in high seasonality regions (e.g., evergreen in boreal regions), though the first response of plants to harsh environments (e.g., cold or dry) is to shed off their leaves. Our simplified definition of phenology in here is to make it possible to evaluate the competitively optimal strategy in future studies. We will add a detailed explanation in this section.

- **Methods Line 254:** PFT-specific parameterisations for the mortality parameter are used, so are there different PFTs each with their own cohort structure? How are PFT-specific background mortality parameters set in the model? Are they all come from observations across different vegetation types? Related reference is missing in the main text. Since the most size-dependent mortality research focus on closed-canopy forest system, whether the “U-shaped mortality pattern” can be extended to other vegetation systems, including forbs, shrubs, grasslands, systems with open canopies and systems experiencing different risks in different environment?

The mortality is delineated by the Eq. 9. We just fixed the default mortality rate according to the general mortality patterns of trees in the forests across the world. We didn’t extensively tune the parameters. We will make the mortality settings clear and validate with observational studies.
Methods Line 374: how does disturbance history set in the MsTMIP simulations? I’m wondering whether the large inter-model discrepancy in simulating plant biomass is caused by disturbance dynamics? For clarity, can the authors be a bit more explicit about the experimental design of the MsTMIP.

According to Huntzinger et al. 2013, MsTMIP only provided prescribed land use types. It is up to the participant models for the disturbances. We will clarify it.

Figure 4d: the authors point out that model analyses are based on equilibrium simulations without explicit disturbances. But the critical height across forests shows an abrupt decrease in the 100 years of model run. What reasons made this pattern happen in the model? Is that driven by the aging-related mortality of canopy trees? Could you discuss more the underlying mechanisms behind the emergent ecological phenomena?

It is a behavior of the model. The simulated forest is transforming from even aged to the mixed aged. In canopy layer, the trees are gradually replaced by younger trees as the old trees yield their space due to mortality. We will explain it in the revised version.

Result Line 515-517: the formulation of allometry makes the tree height growth as a function of tree diameter (Eq. 5 in the main text). Since the two model versions have similar stem growth and tree size distribution, I would assume that tree height growth is stable as well. Why the full demography model shows higher tree height than the single-cohort model (Figure 11c)?

For the full demography model, the height is from the tallest tree. For the single cohort model, all the trees have the same height. We will clarify this.

Result: the authors evaluate the model outputs with the MsTMIP simulations in the Result section. The simple intercomparison would be invaluable to help determine which model behaviour is more realistic. I think it would be interesting to have a section in the discussion tracing the variability that emerges among the models and informing what modeling structural choices or assumptions lead to improved model estimates. Since this paper is a model description paper, further discussion by model developers on the potential reasons for the biases would be much appreciated.

We will add the discussions about why these models perform differently, especially focusing on the assumptions of our model, BiomeE. Turnover rates and NPP are two key factors affecting vegetation dynamics. For the land models, they treat “turnover” in a high variation.

Minor comments:

The abbreviation of the term CAI on Line 409 should be put in parentheses for the first time on Line 407.

Crown area index (CAI). Will add it to the text.
- Lines 71-75: it is unclear to me what is “the legacy of land models and the technical requirements of reversibility in model development”? Could you explain or rephrase this sentence?

We will reword this sentence. Usually, it requires to turn off the new development and make the model performs like the old one.

- Lines 225, “H is tree height” should be modified to “Z is tree height”.

Corrected. Thanks!

- In equation(10), k is ground area? Not defined.

Added explanation.

- Figure 3. How to understand the constant LAI value of KZ?

The model defined a maximum LAI. For Konza, the grasses can reach to the maximum each year, except the first year because of the low initial density. We will explain it in the new version of this manuscript.

- Figure 9. Please add units of LAI.

Will do.