Reviewer: 2
Comments to the Author
I enjoyed reading this manuscript providing a deep dive into alternative ozone damage functions and their performance. The original difference between the Sitch and Lombardozzi approach principally related to whether there is a decoupling in the impact of ozone on photosynthesis and stomatal conductance, e.g. there was empirical evidence of a decoupling under water stress, i.e. sluggish stomata. In the Sitch model the global parameter “a” (eqn 4) was calibrated for that DGVM and monthly chemistry-modelled concentration fields to reproduce the empirical dose-response function for a high and low sensitivity species within a plant functional type grouping. I would anticipate implementation of Sitch et al. scheme into another DGVM would require a recalibration of this parameter (can you confirm whether or not this was done?). For example, what was really done in the runs: Li + Lombardozzi thresholds and functions and Li+Sitch thresholds and function? (any retuning?) I like the Li approach as it considers the effect of cumulative dose on phenology/leaf turnover. Given Li “adopts a data-driven optimal threshold Y” and best form of equation to fit the data, one would anticipate its best performance (especially as the combination approaches may not have been retuned)? Nevertheless, it is interesting to see the comparison of performance against multiple datasets (the comparison was thorough). Given the above I am particularly interested in the comparison of Li and Lombardozzi parameterization schemes. Overall there are some interesting elements to this study and warrants publication although the justification and explanation around inclusion of the mixed schemes should be elaborated.

Response: We sincerely thank you for your highly professional and insightful comments. We especially appreciate your in-depth discussion regarding the mechanistic differences among the Sitch, Lombardozzi, and Li ozone-stress parameterization schemes, particularly concerning the decoupling between photosynthesis and stomatal conductance, as well as the necessity of recalibration when transferring parameterization schemes across different DGVM/LSM frameworks. Following your suggestions, we revised the manuscript to clarify the scientific rationale and limitations of the mixed-parameterization experiments. Specifically, we now explicitly state in the Abstract, Introduction, Discussion, and Conclusion that the mixed experiments using thresholds and response functions derived from the Sitch and Lombardozzi schemes were conducted without additional recalibration or retuning within the Li framework. Therefore, these experiments should primarily be interpreted as structural-sensitivity experiments rather than as fully optimized implementations of those parameterizations within the Li framework. In addition, we further expanded the Discussion section to explain in greater detail the rationale for introducing the mixed schemes and to strengthen the comparison of the mechanistic differences between the Li and Lombardozzi parameterization schemes. We also incorporated experimental evidence from previous ozone exposure studies to demonstrate that ozone flux thresholds remain highly uncertain across vegetation types and species, thereby further justifying the need to conduct structural comparison experiments within a unified framework. In the revised manuscript, modifications made in response to your comments are highlighted in **red**, while revisions associated with another reviewer are highlighted in **blue** for clarity. We sincerely appreciate your valuable comments and guidance, which have substantially improved the quality, rigor, and readability of the manuscript.

Specific comments:

1. I would anticipate implementation of Sitch et al. scheme into another DGVM would require a recalibration of this parameter (can you confirm whether or not this was done?). For example, what was really done in the runs: Li + Lombardozzi thresholds and functions and Li+Sitch thresholds and function? (any retuning?)

Response: Thank you very much for this important comment. We agree that transferring the Sitch parameterization scheme into another DGVM/LSM framework would require recalibration of the empirical parameters, including the global parameter "a", because the original parameterization was calibrated under a specific model structure and chemistry-forcing configuration.

However, in this study, we did not perform additional retuning or recalibration when implementing the mixed experiments (Li + Lombardozzi thresholds/functions and Li + Sitch thresholds/functions). Instead, we only replaced the ozone flux thresholds and response-function formulations within the Li framework while keeping all other model structures and parameter settings unchanged. Therefore, the primary purpose of these mixed experiments was to serve as structural sensitivity experiments to isolate the effects of threshold definitions and response-function forms, rather than the influences arising from other structural components of the parameterizations.

To clarify this point, we have now explicitly added corresponding statements in the Abstract, Introduction, Discussion, and Conclusion sections.

The following text has been added to the Abstract:

The mixed experiments using thresholds and response functions derived from the Sitch and Lombardozzi schemes were implemented without additional recalibration, allowing structural sensitivities to be evaluated consistently within the Li framework.

The following text has been added to the Introduction:

Notably, the mixed parameterization experiments using thresholds and response functions derived from the Sitch and Lombardozzi schemes were not recalibrated within the Li framework and should therefore be interpreted primarily as structural sensitivity experiments.

The following text has been added to the Conclusion:

Notably, the mixed experiments incorporating thresholds and response functions derived from the Sitch and Lombardozzi schemes were conducted without additional recalibration; therefore, the differences reported here primarily reflect structural sensitivities rather than fully optimized implementations of those parameterizations within the Li framework.

2. Overall there are some interesting elements to this study and warrants publication although the justification and explanation around inclusion of the mixed schemes should be elaborated.

Response: Thank you for this insightful suggestion. Following your comment, we substantially expanded the discussion better to justify the inclusion of the mixed parameterization experiments and to elaborate on the mechanistic differences between the

Li and Lombardozzi schemes.

Specifically, we added a new discussion section comparing the ozone flux thresholds adopted in the Sitch, Lombardozzi, and Li schemes with evidence from previous ozone exposure experiments across different vegetation types. The revised manuscript now discusses how experimental studies suggest substantial uncertainty in ozone sensitivity among needleleaf forests, broadleaf forests, shrubs, grasses, and crops. For example, some experiments indicate stronger ozone sensitivity in loblolly pine than in yellow-poplar, while others show that crops such as soybean can be substantially more sensitive than woody broadleaf species. These observations are not always fully consistent with the threshold assumptions adopted in existing parameterization schemes.
We also clarified that one important motivation for introducing the mixed schemes was to evaluate the structural sensitivity associated with threshold selection and response-function forms under a unified Li framework. In addition, by comparing the Li + Lombardozzi experiment with the original Lombardozzi scheme, we were able to isolate further the role of cumulative memory effects from threshold and response-function differences alone.

The following text has been added to the Abstract:
Experimental evidence from previous ozone exposure studies also suggests that the ozone flux thresholds currently adopted across different parameterization schemes remain highly uncertain across vegetation types. For example, experiments comparing loblolly pine (needleleaf) and yellow-poplar (broadleaf) showed stronger ozone sensitivity in the needleleaf species (Tjoelker et al., 1991), whereas studies on Magnolia denudata and Cotinus coggygria indicated that broadleaf trees could be more ozone-sensitive than shrub species (Xu et al., 2021),. Other controlled experiments further suggested a sensitivity gradient of soybean > tobacco > poplar, implying that crop species are generally more sensitive to ozone than many broadleaf woody species (Dai et al., 2019). However, the Lombardozzi parameterization scheme applies a uniform threshold of 0.8 nmol m^{-2} s^{-1} to all five vegetation categories, which may oversimplify species- and PFT-dependent detoxification capacity. Similarly, the original Sitch scheme was developed under limited observational constraints, lacking dedicated experimental evidence for shrub and grass vegetation types; therefore, shrub thresholds were assumed to be equivalent to those for broadleaf vegetation, while grass thresholds were assumed to be similar to those for crops. In contrast, the Li scheme assigns higher ozone sensitivity to needleleaf forests than to broadleaf forests, a result that is not always consistent with the experimental evidence summarized above. Taken together, these inconsistencies indicate that ozone flux thresholds remain insufficiently constrained and still require further recalibration across vegetation types and model frameworks.
This uncertainty was one of the primary motivations for introducing the mixed parameterization experiments in this study. By combining the Li framework with alternative threshold and response-function settings from the Lombardozzi and Sitch schemes, we aimed to isolate the effects of threshold definitions and response-function forms from other structural components of the parameterizations. Importantly, we did not perform additional retuning or recalibration of the empirical parameters, including the global parameter originally used in the Sitch scheme, when transferring these thresholds and response functions into the Li framework. Therefore, the mixed experiments should be interpreted as structural-sensitivity experiments rather than as optimized implementations. In addition, the comparison between the Li+Lombardozzi and the

original Lombardozzi schemes enabled us to distinguish further the role of cumulative memory effects from that of threshold and response-function choices alone. Our results suggest that differences in long-term ozone accumulation and canopy memory can substantially alter the magnitude and spatial pattern of simulated ozone damage, even under similar threshold settings.