We would like to thank Referee #1 for the interest taken in our study and the helpful suggested points of improvement. Attached is a pdf with our answers in red italic type for easier readability. Unfortunately, we can not attach the revised manuscript here, making tracking line changes a bit difficult; but if requested we can send a version with the correct line numbers, and red type for modified sections.

General Comments

A minor suggestion would be to more frequently remind the reader in the technical sections (e.g., 2.1.2) that additional details are available in the appendices. Given the manuscript's length, this would help reader navigation. Similarly, a clearer, earlier statement regarding the different time-steps (e.g., daily, hourly) and the multi-layered soil structure used in the water modeling would help orient the reader from the outset.

Reply: A statement summarizing these aspects of the model has been added at the beginning of section 2 (L202-205). Additional references to the appendices have also been added in the relevant sections.

The presentation of the evaluation results could, however, be improved. In several figures, font sizes are quite small, and some plots appear stretched. For long time series of daily data (e.g., Fig. 10), the fine temporal scale is lost, making them difficult to interpret. Presenting more of this data as scatter plots (as in Fig. 13) could enhance clarity. To improve readability and focus given the paper's length, the authors might also consider moving some detailed results to the supplementary materials, while retaining the key findings from each evaluation level in the main text.

Reply: We have moved figures 12 and 15 to appendices to reduce the paper's length while focusing on the most important and readable results. Regarding the time-series in Figure 10 and 11; while it is indeed difficult to finely interpret the variations between days or even two months, they do make visible the variations between summer and winter months, and particularly whether the model is able to capture ETR peaks and soil draining in summer; furthermore, they allow us to make distinctions between simulation results from earlier and later years, which as important as deviations in stand structure can accumulate over time (e.g. Hesse).

Specific Comments

Introduction:

- The introduction provides a very detailed account of species mixture effects (e.g., lines 70-95), which could be shortened. On the other hand, a broader context of other forest modeling approaches is a bit lacking.

Reply: We have attempted to remedy this by inserting, near the end of the introduction (L161-174), a paragraph that zooms out and replaces the development of PHOREAU within the broader context of 'vegetation dynamic models'.

- **L105-110:** The text mentions "identified two main shortcomings in forest models." It would be helpful to briefly elaborate on *how* these specific shortcomings were identified (e.g., through literature review, previous modeling experiments, etc.).:
  Reply: These weaknesses in predicting tree mortality and regeneration were identified through a review of litterature on gap models, as well as it has been highlighted in former studies (Bugmann and Seidl 2022 – cited in the manuscript): these processes were rarely evaluated directly, but rather integratively with predicted species distributions and site basal areas. The paragraph has been amended to reflect this (L106-109).

Model Description:

- **L159:** Typo: "plaform" should be "platform". *Amended*

- **L216 (Eq. 1):** There appears to be a layout issue in the equation. It should likely read $2*H\_max,s - b\_s * e^{(...)}$ rather than having the allometric parameters in the denominator's exponent. Please verify the formula. *Amended*

- **L232:** Suggest inserting the word "species": "shade intolerant *species* having...". *Amended*.

- **L251-265:** The concept of "crown ratio reversion" needs clarification. Does this mechanism allow the base of the living crown to move downwards again, effectively re-greening parts of the stem that were previously bare? This should be clarified here and in the appendix. *Yes this is the intent of the new mechanism. The paragraph has been clarified (L312-314).*

- **L277:** The symbol for the clumping factor appears to be missing from the parentheses: clumping factor ( ). *This is very strange, the Ω symbol is visible in the uploaded pdf on our end.*

- **L333:** While "symplasm" is defined by contrast, a brief explanation of "apoplasm" (the continuum of cell walls and extracellular spaces) would be beneficial for non-specialists. *Description has been added (L395-397).*

- **L368:** Typo: "depending on depends". *Amended*.

- **L382:** It would be useful to briefly explain what a "semi-implicit solver" is and give an indication of the runtime difference it makes compared to an explicit solver. *Some*

further details have been added in the text (L446-449); but readers should refer to the Ruffault 2022 paper for a comprehensive description of the pros and cons of each method.

- **L531 (Eq. 13):** There may be a typo in the denominator. The text reads P_gSBB, which seems inconsistent with the parameter P_gs88 defined on L525. Please check. *Amended.*

- **L545:** The growth reduction factor GR_crown is used here but has not been fully explained. A brief definition is needed. *Explanation has been expanded (L605-608).*

- **L553:** The citation Hammond et al., 2019 appears to be missing a closing parenthesis. *Corrected typo.*

- **L620:** The rationale for weighting light availability by daily mean temperature needs more justification. This method heavily weights hot summer months when growth may be limited by other factors (like drought). A temperature response curve that is primarily limiting at the cool end of the spectrum might be more ecologically realistic.
  Reply: *This explanation was indeed misleading: it is in fact not average temperature but rather GDD that is used to weight light availability. Furthermore the disregarding of drought-stress in the formulation is a result of the fact Phenofit was linked with ForCEEPS before SurEau : further developments should indeed take full advantage of drought-stress prediction. This limitation was aknowledged in the test (L681-684).*

- **L651 (Eq. 19):** Notation for leaf unfolding and coloration intervals is slightly inconsistent between the text (Uls, Cls) and the key below the equation (UI_s, CI_s).
  Reply: *It seems that this is a formatting issue from the GMD site because we have exactly the same notations on our end (now Eq. 24).*

- **L695:** A word appears to be missing in "the fine root area of a tree in a determines...". *Amended.*

- **L749:** There is an extraneous character (a hyphen) after the period at the end of the paragraph. *Amended.*

Results, Discussion & Figures:

- **L765/Figure 4:** This figure is difficult to interpret due to very small font sizes and hard-to-distinguish color-coding. Furthermore, the claim that it shows "acclimatization" over 1500 years seems more likely to reflect changes in stand structure and species composition rather than plastic adaptation within individual long-living trees. Please clarify.
  Reply: *Due to concerns about the readability of the figure and its overall quality, we have decided to move it to supplementary information (W17). It is indeed the acclimatization*

effect and not structural changes that are responsible for the overperformance of simulation B (with a moderate drought event before the main drought event) compared to simulation A, and we had confirmed this by looking at simulation results obtained without the new root plasticity module, where simulation B does not overperform simulation A. Unfortunately this is not apparent in the figure, and we have updated the caption to reflect this.

- **L880/Figure 6:** There appears to be an inconsistency in the visualization. For instance, for the Puéchabon site, the caption states "3 patches of 100m²", but the grey grid lines on the ground seem to depict a different arrangement (e.g., 4x4 grid). This should be checked for all subplots.
  Reply: Indeed gridlines are a purely visual artifact of the Capsis vizualisation, the caption has been amended to reflect this.

- **L914:** The paper states that longer time-lapses "would have mechanically improved simulation results". This is counter-intuitive, as one might expect simulating longer periods to be more challenging and prone to error accumulation. Could the authors please clarify what is meant by "mechanically improved" results in this context?
  Reply: This statement has been reworked and nuanced (L922-926). In fact, when looking at simulations results, the length of the simulation had no significant impact on prediction error (Fig. W11e). By longer time-lapses we meant that focusing on yearly patterns is a harsher test (and thus more relevant) than doing so on decades or more. In fact, while longer simulations are indeed more prone to error accumulations, they are also more forgiving to errors in predicting single-year deviations from the norm caused by extreme climatic events. It is this second effect which we are more keen to evaluate for PHOREAU, in view of its application to predicting the effects of climate change. Of course, the best situation would be to use long time-series with fine temporal resolution as reference to evaluate a model's predictions. Yet, such data are still scarce for trees physiology, except in some key sites such as the ICOS ones.

- **L1019:** Typo: "crown Al ratio" should likely be "crown ratio" or similar. *Corrected error*

- **L1059:** There is a minor date discrepancy for the Hesse site thinning. The main text mentions a cut in 2005, whereas Appendix Q lists thinnings in 2004 and 2009. This could be harmonized for clarity. *This has been amended.*

- **L1250/Figure 17:** This figure effectively illustrates the model's performance across ecological gradients. A very nice visualization.
  Reply: We thank the reviewer for this statement!

- **L1375 (and elsewhere):** The citation 'Allen, Macalady, Chenchouni...' is very long. This format occurs multiple times (e.g., L54) and could be consistently shortened to 'Allen et al.' for readability. *Yes this was very strange! Amended.*

References:

- **L2231:** The reference for Bréda, Soudan and Bergonzini is listed with "(no date)", which is unusual and could be clarified. *Amended.*