General Comments:

The manuscript applied the existing individual-based forest gap model (FORMIND), which was developed using data from an old-growth temperate forest in the northeastern China. Authors present a novel methodology to distinguish between mature and immature trees to understand forest productivity. This approach offers a fresh perspective compared to traditional methods focusing mainly on gross primary production (GPP). The manuscript is well-organized, systematically presenting its research approach, results, and conclusions.

Manuscript is interesting and useful to international audience of the journal. However, there is room for improvement in the manuscript. The approach and conclusions are somewhat limited by the methodological framework and the absence of a comprehensive analysis of the ecological implications. Authors are suggested to address the following issues while making the revision.

> Thank you for your thorough review of our manuscript and the many helpful comments. Below, we respond to each of them and explain how we would like to address them in a potential revision.
>
> Our responses are written in blue font and indented.

The introduction provides adequate background but lacks a critical review of previous research methodologies (other process-based models) and their limitations. Also, there is a need to introduce the limitation of the current FORMIND model. Suggest to enhance the introduction.

> Thank you for this suggestion. We will add a short review of other process-based models to the 2nd paragraph after the research questions on page 3.
>
> Of course, FORMIND and other forest models have limitations, which are very important to understand to correctly interpret the results. In order to appropriately account for these limitations, it is important to be specific and explain how each individual limitation may affect the results. Otherwise, we would be at risk of leaving the reader with common sense statements about the limitations of modelling in general.
>
> We believe that the introduction, which we intend to set the frame of the study and to provide general guidance to the reader, is not the best place for a discussion of limitations with the necessary level of detail. Therefore, we provided a thorough discussion of limitations in section 4.1. We will extend this section further in response to your comments below.

Different species might exhibit significant variations in growth and carbon dynamics, even within the same maturity classification. The selection and classification of trees into mature and immature categories are not sufficiently justified. More rigorous criteria and a discussion of potential biases in these classifications are needed.

> Thank you for this thoughtful observation. We agree that a binary distinction of mature and immature trees could be difficult in field studies (as acknowledged in section 4.1; 2nd paragraph on p. 18).
>
> Our maturity definition, based on the carbon use efficiency (CUE) that an individual could attain in the absence of competition, allows us to distinguish competition (which is production-related) and individual growth bounds (loss-related) as main limiting drivers of NPP. We admit that this may not have been sufficiently clear in the paper. We will therefore

clarify and adjust the respective sections in the introduction and discussion accordingly. Furthermore, we will add a discussion of potential biases to the text.

For a potential application in field studies, we suggested and analyzed the DBH entropy as a proxy for tree maturity. However, within the context of our modelling study, we could also identify the process (i.e., individual limitations rather than competition) that drives productivity.

To account for the species- and individual-specific differences in carbon dynamics, we incorporated species-dependent and randomized growth limitations into the model.

The choice of the FORMIND model may not fully capture the complexity of forest dynamics, especially in terms of species-specific interactions and responses to environmental variables. For instance, you did not apply any space competition in the model. Suggest to compare results with those obtained from alternative models, particularly those incorporating more detailed species-specific parameters or interactions with abiotic factors.

Thank you for this suggestion. The model *does* feature space competition, but we may have caused some confusion by misleadingly using the word "space competition" for "crowding mortality", which is typically modelled as increased (stochastic) tree mortality in "full" forest patches. We did not use this stochastic phenomenological model for space competition. Instead, as a new innovation, we took a more mechanistic approach and introduced tree mortality due to strong light competition. That is, trees that do not receive enough light to fulfill their respiratory needs die.

This process has the same effect as crowding mortality ("full" forests lead to deadly overshadowing of small plants) but has a better mechanistic justification. In fact, both the species of the shadowing and the overshadowed trees are considered explicitly, since the LAI of larger plants as well as the respiratory demands of smaller plants are parameterized individually for each plant functional type (PFT). We will clarify this in the supplement section B.8 and change the wording of "space competition".

The acknowledgment of the model's limitations is a positive aspect, but the discussion lacks a critical assessment of how these limitations might have influenced the study's conclusions. Suggestions for alternative modeling approaches or supplementary methods to address these limitations would provide a more balanced view.

Thank you for pointing this out. We will extend the limitation section 4.1 to cover the effect of the limitations on the results more explicitly.

A lot of supplement information is provided with the manuscript. I'm not sure if it can refer to other fundamental literature previously published. Are the allometric relationships part of the FORMIND model? It would be better to keep concise and easier for the readers to understand.

Thank you for this question and comment. Our study contains several innovations that were necessary to accurately reproduce the behaviour of the Changbaishan forest. The increased accuracy came at the cost of a longer paper supplement. The supplement also contains the other values of the Changbaishan parameterization. This is required for full reproducibility of the study.

We understand that the length of the supplement poses a challenge to readers, who typically need to quickly identify the information most relevant to them. We will therefore provide a

table of contents and summary of our model extensions at the beginning of the supplement to guide the reader.

Detailed comments:

Introduction: Consider providing a brief introduction on any challenges or limitations encountered while adapting the FORMIND model to this specific old-growth temperate forest.

Thank you for this suggestion. As the parameterization is not the focus of our study but rather a “necessary nuisance” on the way to tackling the actual research questions, we feel that a discussion of the parameterization challenges might draw the readers’ attention too far away from the main topic. Nonetheless, we will provide further reasoning to indicate why a new parameterization and the new features were necessary.

Introduction: P2 second paragraph: "Nonetheless, it has proven difficult to identify clear relationships between forest structure and NPP (Chisholm et al., 2013) as several factors interact...": Suggest elaborating on the specific factors that complicate the relationship between forest structure and NPP.

Thanks, we will do that.

Method: Page 4 2.1 Field data: I have concerns about the allometry information and biomass equations provided in Supplementary A. (1) A lot of species lack allometry data and biomass equations. How did you address these species in your study? Did you use likelihood-based analysis similar to the PFTs classification? Please clarify this in the methods part. (2) The biomass equations, adopted from Chojnacky et al. (2014), are generalized primarily for North American species. Since most of the equations are empirical models, I doubt their accuracy when used directly.

We classified species into plant functional types (PFTs) to keep the number of free parameters tractable. Each PFT has a single set of “mean” allometric relationships, which we applied to all plants of the PFT – including those for which no specific allometry data were available.

In SI B.3, we explain how we fitted the allometric relationships to the available data. To mimic the Changbaishan forest as precisely as possible, we weighted each species’ data according to the species’ prevalence (basal area) in the inventory. Since the species for which data were available cover more than 96% of the basal area in the inventory, the missing data had a negligible effect on the resulting allometric relationships. Nonetheless, we will clarify that species with missing data were excluded in the process of fitting the allometric relationships.

The biomass equations referred to in SI A were only used to determine the aggregate biomass share of each PFT in the inventoried forest. This, in turn, was used to scale each PFT’s mean stem biomass proportion to the correct value. This process is described in SI B.4.5. As we applied the biomass equations from Chojnacky et al. (2014) only on an aggregate level to estimate a single parameter per PFT, and since our approach is in line with existing literature (Piponiot et al., 2022), we believe that using the generalized equations does not undermine our results. In the actual model, we only used the allometric relationships that we derived directly from the field data.

Methods: Page 6 Model fitting, second paragraph, “We fitted these 18 parameters...”: cannot get 18 parameters based on your description, please clarify.

> Thanks, it should read “26 parameters” = 4 PFT-specific parameters * 6 PFTs + 2 general parameters. We will clarify this.

Results, first paragraph, “...This contrasts with the basal area of immature trees.” please add supporting figures or statistical results.

> The corresponding statistical results are provided in the next sentence. We will combine the sentences with a colon for better clarity.

Results: “...These results are depicted in Figures 4 and 5.” Figures should be accompanied by the corresponding results in brackets. This will make it clearer for readers to correlate the interpretation with the figures.

> Thanks, we will clarify what is visible in each individual Figure.

Results: “...These results are shown in Fig. 6.” Similar suggestions as above.

> We will number the figure’s panels so as to be able to refer to them individually.

Results, Figure 7, does each dot correspond to a forest patch of 0.04ha? Please clarify.

> Correct. We will clarify this.

Discussion, Page15, second paragraph, “The proportionality can be explained by the strong connection between the individual-level basal area and GPP in conjunction with the negligible NPP of mature trees.” Consider simplifying or re-framing it for better readability.

> We will do that.

Discussion, Page 15, second paragraph, “On the stand level, however, neither the GPP nor the respiration were correlated with the proportion of immature trees (Fig. 7) ...”, in the discussion section, only this sentence refers to the figures. Please maintain consistency.

> We will remove the reference to Fig. 7 to maintain consistency.