

## General Comments:

Fischer et al. present a novel method for relating the net primary production (NPP) of forests to forest structure. They explore this method using the process based forest model FORMIND parameterised with data from a forest inventory plot located in an old growth temperate forest in Changbaishan, China. The method rests on the assumption that as trees grow their respiratory demands increase, resulting in a decline in NPP and carbon use efficiency with size. As a result, the proportion of immature trees in the forest predicts stand level NPP as only immature trees are putting on biomass. The authors test a number of structural metrics and find that a measure of DBH entropy is the best proxy for the proportion of immature trees and could thus be used to predict the NPP of a forest using only inventory data.

This topic will be of interest to the forest ecology community and those interested in predicting carbon sequestration using metrics of forest structure. The manuscript has a good structure and is for the most part easy to follow. The figures show the results clearly. I think keeping most of the details of FORMIND in the supplement and only describing necessary details in the main manuscript is a good approach, although I think some more high level summaries in the main text would be helpful so that readers do not have to keep moving to the supplement which is very long.

Thank you for your thoughtful review and helpful comments! Below we respond to each of your points and explain how we would like to address them in a revised manuscript.

We understand and agree that the supplement is long, potentially making it difficult for readers to find the information they are interested in. To mitigate this issue, we will add a table of contents and a high-level summary to the supplement. We will furthermore add some more details to the main text where possible.

My main suggestion is to include more discussion of how some of the assumptions in FORMIND influence results. In this manuscript, FORMIND is set up so that trees have a maximum DBH, above which they allocate all GPP to respiration meaning that NPP goes to zero. It is therefore not surprising that the proportion of immature trees predicts NPP. The authors acknowledge in the discussion that there is still debate in the literature about whether NPP declines with size. It would be good to expand on this and discuss how uncertainty in the changes of individual NPP with size influence the ability of stand structure to predict stand level NPP. They could also test this assumption using the field data by looking at a time series of growth rates for individual trees and identifying declines in growth with tree size.

Thank you; this are excellent points.

Letting the NPP go to zero with increasing tree size is a long-standing modelling paradigm of forest models and not easy to change ad hoc. Nonetheless, following your comment, we have thought of a way to assess the effect of the zero-NPP assumption and to test milder versions of it. We will run the model assuming that mature trees (though not growing) still have a positive NPP, given as a certain fraction of their last NPP prior to reaching maturity. We will vary this fraction to analyze how strong the NPP reduction of mature trees needs to be so as to lead to the effects we presented in the paper. Though breaking some of the base assumptions of the FORMIND model (in particular that NPP affects the tree dimensions), this will be a great validation experiment, which we will include into the model validation section as well as into the supplement.

Unfortunately, it will be difficult to test the low-NPP assumption directly on our field data, as tree maturity likely depends on local conditions we do not have data for. Furthermore, the

inventory data are partially subject to measurement errors (manifesting in negative or very large DBH increments) making us a little sceptical regarding such a detailed (individual-focused) direct analysis of the field data. Nonetheless, we present DBH-dependent DBH increment data on an aggregate level in SI B.7.3.

Specific comments:

It would have been helpful if the manuscript had line numbers.

We totally understand that and are sorry that we could not deliver a manuscript with line numbers due to regulations of the preprint server we used (ArXiv).

On page two the authors say that CUE is expected to decline with size because large trees have higher demand for respiration and non-structural carbon. This doesn't necessarily follow unless GPP increases with size are less than respiratory and NSC increases. Or GPP asymptotes or decreases. Do we know this to be the case in this forest?

We do not have data on the CUE of individual trees in this forest. Therefore, we modelled the CUE of trees based on the assumptions laid out on p. 5 (second-last paragraph).

Our statement was primarily motivated by theoretical arguments. Increases in GPP require new structures such as leaves, which in turn need to be maintained. The higher trees grow the smaller is the additional gain of additional leaves due to self-shading and the higher are the additional costs due to longer transport pathways. We are not aware of a study suggesting an increasing CUE for individual trees but will double check. This is a relevant point, and we would love to consider this here or in follow-up studies!

Bottom of page 4 - missing word - "*The key idea*".

Will be fixed.

Page 5 - how were the light requirements of the species known? Do those classifications align with growth and survival rates from the inventory data? It looks like in Fig S1 the two species with the fastest DBH increment rates are on almost opposite sites of the shade tolerance spectrum which is a little surprising. Was there any comparison of growth rates predicted by FORMIND and the mean growth rate per PFT from the inventory data?

Thank you for these observations and questions!

The light requirements were derived from available literature such as *Niinemets & Valladares (2006)* and *Wang et al. (2010)*. The provided values correspond roughly to the minimal required sunlight for growth of saplings: a value of 1 corresponds to a light requirement of >50% of the full sunlight; 2: 25–50%; 3: 10–25%; 4: 5–10%; 5: 2–5%. We will clarify this in the revised manuscript.

Though the inventory data are the best we have, the small sample sizes for rare species are small. This applies in particular to the species with the highest DBH increments. We agree that the combination of high median DBH increment and high shade tolerance is surprising for *Abies nephrolepis*, and it was classified as a mid-tolerant species. A different classification of the species would be possible but hardly affect the parameterization and results, as *Abies nephrolepis* is rare in the Changbaishan area.

Unfortunately, the sample size issue impact also the mortality analysis. Nonetheless, we have analyzed the mortality rates for each PFT and presented the results in SI B.9 (see also Fig. S11). There, it is visible that the mortality of shade intolerant species is mostly higher than the mortality of mid-tolerant and shade tolerant species.

Please provide the Genus name of *Q. mongolica* in the first mention of this species.

We will do that.

I didn't understand how growth was modeled without reading the SI. Consider some high level descriptions in the main text.

Thanks; we will do that.

Middle of page 6 - why did the authors use biomass and stem count as an indicator of the size distribution rather than the actual size-distribution?

As our goal was to correctly model the forest dynamics, driven by local competition, we desired to match the local forest composition. At the local scale (here: 20m x 20m), the forest composition can show different patterns than on the larger scale. For example, a forest may consist of 50% species A and 50% species B, but aggregation and segregation may lead to patches with, say, 75% species A / 25% species B and vice versa, meaning that the species are not well-mixed. This local structure is important for understanding the forest dynamics but difficult to cover with aggregate measures.

Since a forest's local composition can be variable, we needed to apply a stochastic framework to fit the model (which is in line with FORMIND's stochastic nature anyway). To that end, we needed to estimate the *distribution* of the local forest states. That is, for example, how likely will we find a local patch with 75% species A and 25% species B? How often will we find 50% / 50%? The joint distribution of trees and their sizes has a high dimension, making it challenging to estimate this distribution. Considering, say, 10 different size classes for each PFT would lead to a 60-dimensional estimation problem, which is practicably infeasible (unless someone comes up with exciting new sampling / estimation theory). Therefore, we considered summary statistics of the distributions of biomass and stem count at the scale of 20m x 20m. We will clarify this in the text.

Preliminary simulation experiments with additional statistics (not shown in the paper due to space limitations; this would be a full second paper) indicated that the information loss due to summary statistics is small, and the gain via improved likelihood estimates (due to a smaller-dimensional sampling space) is larger, which helps us finding the true likelihood maximum.

There are not many details of wood decomposition or soil respiration parameterisation in the methods section. The authors could consider adding a few sentences explaining this aspect of FORMIND. It wasn't really clear to me what the hypotheses were for how NEE would change with forest structure.

Thanks; we will add information on this process to the manuscript.

Page 9 section 2.5. To identify trees that have reached growth limits in field data would it not be simpler to look at growth rates and find large trees that have declining growth rates (if there are enough census intervals, or if not, those that are growing less than some quantile of the population)? Or is the idea that DBH entropy can be used in field studies with a single census?

Thank you for these ideas! In our dataset it was difficult to clearly identify declining growth rates, also due to the limited number of census intervals. If we considered all trees growing less than similar conspecifics, we would still need to distinguish the causes for the reduced growth: is it due to production-related factors (e.g. reduced light / shading) or due to internal limitations (maturity)? We will discuss this point in the text.

It is correct that we suggest using DBH entropy as a proxy in field studies, and yes, it can be used with a single census.

#### Fig 4. Why are there few forest patches with a Shannon diversity of just over 1?

Interesting question! We would like to start by observing that the Shannon diversity is computed based on discrete numbers (stem counts and PFTs). Therefore, the Shannon diversity is discrete itself and not continuous, and there is no a priori reason to expect that neighbouring diversity values should occur with similar frequency.

Now we consider the specific question why Shannon diversity values just above 1 occur much less frequently than values of exactly or a little less than one. A Shannon diversity of 1 is obtained if a patch contains exactly two PFTs with the same stem count. If the stem counts have a small relative error (which, noteworthy, requires that there are many stems), the Shannon diversity is just a little smaller than one.

A Shannon diversity just over 1 requires that there are three species: two with almost equal, large stem count and one with small stem count. In our simulations, such species compositions occurred rarely and, in particular, much less frequently than species compositions with two almost equally frequent species. Ecologically, this may be a sign of conspecific aggregation.

Results in Fig. 4 are not that surprising since the model was set up so that NPP would not be affected by large trees because large trees allocate all GPP to respiration. But there is no test of that in the field in this study. Is that a reasonable assumption? Some evidence suggests that large trees continue to actively accumulate biomass e.g. Stephenson et al. 2014

<https://www.nature.com/articles/nature12914>. Ah, I see the authors bring this up later in the discussion. If possible the authors should try to address that assumption using the inventory data from Changbaishan.

Thank you for bringing up this point! We have addressed it in our response to your main suggestion above (second response). We will evaluate the sensitivity of our results of our allocation assumption and check if we can also get some further information from the inventory data.

#### Fig 4. NEE is very low in mature patches - is this because of less wood turnover? Is all CWD in FORMIND the result of mortality or is there also a representation of branch turnover?

Thank you for this observation. A small (or very negative) NEE corresponds to large carbon emissions from the forest (which may result from *increased* wood turnover). Forests with many mature trees have a small NPP, which also manifests in a low NEE.

In FORMIND, NEE is dominated by mortality but also includes mortality of branches falling dead in response to a negative NPP (i.e., if the respiratory needs cannot be satisfied).

Bottom of page 11. Does respiration here include soil respiration? Same comment for Fig. 7c.

No, here (and in Fig. 7c) respiration refers to the respiration of alive trees only. We will clarify this in the revised paper.

**References:**

Niinemets, Ü. and Valladares, F. 2006. Tolerance to shade, drought, and waterlogging of temperate Northern Hemisphere trees and shrubs. *Ecological Monographs* 76:521–5471.

Wang X, et al. 2010. Spatial distributions of species in an old-growth temperate forest, north-eastern China. *Canadian Journal of Forest Research*, 2010, 40: 1011-1019.