Dear reviewers and editor,

First of all, in response to the general comments,

> *Reviewer1: I have now re-read the manuscript in light of my comments (and reply to them) madein the first round of review. I have not re-checked the online material since it has not changed since the last version. The paper has improved a lot and I only have a few minor comments that may help the authors to further enhance the manuscript. I would like to highlight that the discussion is very rich and interesting.*

> *Reviewer2: The manuscript gained a lot in comparison to the first version by including understandable paragraphs about the context, the advantages and the knowledge gain of the proposed method. This makes it a valuable contribution to Biogeosciences. I have only some small remarks in the text.*

we would like to thank the reviewers for their positive assessment.

We sincerely appreciate the time they have invested to improve our manuscript, and we have carefully read and considered all comments.

Most comments were asking for small changes or clarifications. We did not see any point we disagree with and will therefore largely follow the suggestions of the reviewers. A few points require some additional explanation, which is provided below. We structured our response in two parts: the first part deals with general points that were raised by both reviewers, and the second part addressed the remaining comments. All points to which we responded simply with "OK" will be included as suggested.

Sincerely, on behalf of all authors
Florian Hartig

General points raised by both reviewers:

Objective of the study

> Reviewer1: Introduction: would be good to once clearly state the objective and/or main research questions

> Reviewer2: P13101L14: Like reviewer 1 I would like to read here a clear objective of the study.

Response: The main objective of our study is to show that the technique of simulation-based likelihood approximation can be successfully applied to fit a relatively complex vegetation model, which allows us to conclude that it will probably work well for complex ecological models in general. There has been, to our knowledge, no application of this technique to a ecological model with complexity comparable to FORMIND, although our results suggest that this is possible, and not even that complicated. We therefore think it's valuable to document the process and our experience.

We will work on the text to explain this better in the revision.

Markov-Model

Reviewer1: P13103L2: "The model is Markov: : :." Reformulate (sounds a bit like technical jargon) and maybe add a sentence introducing Markov methods since you can not assume
that everybody knows that.

Reviewer2: P13103L2: 'The model is Markov' seems rather dramatic. Markov was a Russian
Mathematician, so please be more specific whhat kind of characterization you want to apply to the model.

Response: We will reformulate or erase the Markov to make the sentence easier accessible. The important information is that trees have no "memory", and their entire state is described by position, species and size

Field data

Reviewer1: P10306L22: "field data" is misleading since you also use virtual data.

Reviewer2: P13107L24: The term 'virtual field data' for the inventory created by the model is still misleading. It is tempting to combine 'virtual' (because of model generated) with 'field'
to show that these data in this case represent the 'observations' but I would feel more comfortable with 'virtual inventory' or something similar.

We will change to virtual inventory and make sure that field data is used only for data that is from the field.

Difference to previous parameterization

Reviewer1: P103110L7: Looking at Fig 5, I think it is important to note that in quite some cases, the values of Dieslich et al (2009) are outside the range found used in this study

Reviewer2: P13113L5ff: It is interesting that marginal posterior probabilities from your method is partly lower (11 parameter) and partly higher (3 parameters) than the previous estimates. Just as a interested comment: have you any idea why? There is no real pattern (like all mortality rates are lower or all parameters for pft5).

It is difficult for us to say exactly what the reason is, because, unfortunately, several things were different in the two studies. We gave details on P.13114 in the submitted discussion paper

However, although Dislich et al. (2009) calibrated to the same data, they also considered the fit of other model outputs such as total biomass and expert opinions for fixing the parameters. Expert opinion in particular would favor more pronounced
5 differences in mortality rates between mid and late successional species due to ecological expectations, although specific empirical data on tree mortality or on maximum growth rates under full light were not available. Secondly, there are significant correlations between the parameters, which allow gaining a similar fit with a range of different parameter values. And finally, we were using the model in this study at a lower temporal
10 resolution (5 yr time steps) than Dislich et al. (2009) to reduce computing time, which can affect model dynamics and equilibrium distributions, meaning that slightly different parameter values would be estimated for the same model with different temporal

resolution.

Particularly because the "objective function" was somewhat differently defined in the two studies (CD used expert opinion as well as a number of other outputs in the manual calibration), we would not necessarily expect that the Bayesian credibility interval, based on the objective function in the present study, overlaps with all values found in the study of CD et al. The main purpose here was to show that both parameterizations end up in similar areas of the parameter space.

Response to further comments

Reviewer 1

P13098L1-2: Is it not a problem of ecological modelling rather than of ecologfy and evolution? (It's a matter of taste and the latter is of course also true but I feel the former is more precise). We agree, will change to ecological modeling or a similar wording.

P13100L8-9: I do not understand was this sentence means and how it refers to the sentence before. Reformulate? We agree, this was not properly explained. What we meant is that it is easy to include explicit observer models that simulate how the data was taken in a process-based model (e.g. a virtual forester that takes a linear path through the forest, and measures every 3rd tree with a certain measurement error). We will expand this comment in the revision, as we think it's a crucial advantage of the simulation-based likelihood approximation that we can deal with simulated replications.

P13100L13-14: "this route was blocked") find another formulation OK

P13102L2: "exhibit") delete OK, thanks

P13102L25-26: In the paragraph before you said that the trees do not have an explicit position? This is confusing: We will formulate this better, they are assigned to a 20x20m grid cell, within the grid cell they have no explicit position

P13104L21: "process-stochasticity") is stochastic processes not clearer or do you mean something else? If yes, explain. It's more a question of taste, but to our feeling "stochastic process" could refer both to stochasticity in the observation, and to stochasticity in the ecological process … in statistics, it's common to use the word "process-stochasticity" to make the distinction to the stochasticity in the observation process.

P10306L17: Here the passive form "is required" seems stylistically more logical OK

P10307L5: "informal model calibration")Never heard this term before, I think you mean manual calibration based on visual assessment of model fit to data. Probably how you call it is correct, just check if this is what you actually mean to avoid misunderstandings or simply call it manual calibration since you also use that term later. OK, we will change the formulation as suggested.

P10307L14-16: So other parameters do not influence the outputs you looked at? This was not very well expressed by us. Of course, other parameters are influencing the outputs as well. What we meant to say is that from the viewpoint of the statistical algorithm, it doesn't

matter how many parameters are under calibration, except for the practical aspect that the MCMCs get slow. We will reformulate this sentence to make this clear.

P10309L5ff: It's trivial that varying more parameters as you did in V2 will widen the posterior distribution. I think you can not really compare that with the posterior distribution of simulations where less parameters have been varied. Would rather be interesting to discuss how many parameters (as in V1 or in V2?) need to be varied. I would argue the more the better but since you are aware or strong correlations between your parameters you could argue otherwise. Yes, if there are interactions between the parameters with respect to the outputs we are fitting to, it is expected that the parameter estimates get wider. Still, we find it interesting to see that this is indeed the case, to look at correlations between parameters (it's not really foreseeable which parameters estimates will widen most), and to test whether the MCMC still work. We agree that in general, all parameters that cannot be fixed through field data or other means should be put under calibration. However as we are working with a virtual example here, it was more a proof of principle, i.e. how many parameters can still be sensibly estimated with the given data.

P103110L12: "by from") revise OK, thanks

Fig 1: How do you define extreme values (beyond one SD?)? I would rather call that the full range or so. What we meant is the maximum and the minimum values, will be reformulated.

Reviewer 2

P13098L1-2: Inverse parameter estimation is a technique in many fields, not 'just' ecology and evolution. OK, will reformulate, also in response to Rev.1

P13099L6-7: Dito OK

P13099L19: Change 'parameters' to 'parameter' OK, thanks

P13101L24: Please replace 'revovery' with 'recovery' OK, thanks

P13102L2: Please delete 'exhibit' OK, thanks

P13102L25-26: Please clarify how the position of the trees and crowns is derived when trees do not have explicit positions (L20). Will be better explained, see comments to reviewer 1.

P13108L24: I would prefer 'specific' over 'concrete'. OK

P13110L18: Change 'parameters' with 'parameter' OK, thanks

P13110L22: Change 'models' with 'model' OK, thanks

P13111L25ff: Maybe the interpretation of the correlation is better understandable when you refer to them as 'higher mortality rates have the same effect on the model results, i.e. obtained biomass, as higher recruitment rates'. The same for the next sentence in which you infer from increased mortality rates to a need in increasing alo recruitment rates. But in my view, correlations just give an information on similar effects of changes

in the respective parameters (as you describe from P13112L11). We will incorporate these suggestions in the new version.

P13113L15: Is by 'reation' 'relation' meant? Yes, thanks

P13116L18: Change 'that' to 'than' OK, thanks

Table 2: Abbreviation is given for SSD but in the table SDD appears. Please clarify. OK, thanks