We would like to thank the reviewer for their helpful suggestions which we are confident will lead to an improved paper.

In “Understanding the uncertainty in global forest carbon turnover” Pugh et al use remote-sensing based turnover estimates to evaluate the performance of six TBM. Based on this evaluation the authors propose eight hypothesis which are then discussed. The study is well structured, the discussion is insightful and the hypotheses are supported by the analyses. It is clear that a lot of thinking went into this analysis which by itself is a sufficient reason to support publication of this manuscript.

In my opinion the discussion lacks one section, i.e., a critical assessment of the concept of biomass turnover and whether it is key benchmark for model evaluation or an observation that should only be used if more process-specific observations become available. Given that several model groups are replacing their turnover parameter by an explicit representation of the different mortality events, what is the future of these remote-sensing based turnover estimates?

>> This is an interesting point. Biomass turnover is a metric that captures high-level dynamics of forest carbon cycling, but we agree that the processes that drive carbon turnover are much more nuanced. Accurately simulating the response of biomass turnover to different environmental conditions likely requires breaking down the underlying processes, such as we pick apart herein. However, any constraint on overall biomass turnover is also an important constraint on models. As the manuscript is already quite long, we prefer not to add a whole section on this, but we have added the following sentences to the start and end of the conclusion:

"Biomass carbon turnover time is a high-level metric that integrates over a wide variety of underlying processes."
"This benchmarking must go beyond the emergent property of turnover time, to the underlying processes, facilitating model improvement as well as evaluation."

From a scientific point of view the manuscript could be accepted as it is. Nevertheless, the current manuscript is very dense. The manuscript could become easier to read and digest (and would therefore become more likely to make an impact) by: (1) Rewriting/expanding the equations (especially eq 2). The study does a good job in disentangling the major processes that contribute to the turnover time of biomass carbon but the equations fall short of reflecting this complexity. Either the introduction or section 2.1 could be used to refine and better formalize the definition of turnover. Ideally each of the hypothesis should be reflected in one of the terms shown in the final equation.

>> We presume that the reviewer is referring to $\tau = C_{veg} / F_{turn}$ as given in the first line of section 2.1. Reviewer 1 also commented on this point and we agree that it would be clearer to expand the equations to split F_{turn} up into its constituent components. We have added additional equations to the introduction:

"Turnover time of existing biomass can thus be calculated as,

$\tau = C_{veg}/F_{turn}$ *(Eq. 2),*

F_{turn} is the total loss flux of live biomass due to the transfer of plant tissue to dead pools of litter and soil, to harvest products and residues, or to the atmosphere via burning. It can be decomposed into,

$F_{turn} = F_{mort} + F_{leaf} + F_{fineroot} + F_{repro}$ *(Eq. 3),* *where F_{mort} is the carbon turnover flux due to plant mortality or woody carbon loss, F_{leaf} and $F_{fineroot}$ that due to leaf and fine root senescence respectively, and F_{repro} turnover due to reproductive processes (e.g. flowers, fruits)."*

We have also expanded the description in section 2.1:

"Turnover times can also be defined relative to particular turnover fluxes, such as those outlined in Eq. 3. In this case the turnover time is calculated respective to the appropriate biomass pool. I.e. Turnover time of vegetation biomass due to mortality, $\tau_{mort,}$ is defined as C_{veg}/F_{mort}, and turnover time of fine root biomass, $\tau_{fineroot}$, is defined as $C_{fineroot}/F_{fineroot}$, where $C_{fineroot}$ is the fine root biomass. F_{mort} can also be decomposed into fluxes resulting from particular mortality processes, for instance, following the conceptual groupings in Table 3,

$$F_{mort} = F_{mort,vitality} + F_{mort,disturbance} + F_{mort,background} + F_{mort,heat} + F_{mort,other} \qquad \text{(Eq. 4)},$$

although other process breakdowns can also be applied. Accordingly, a turnover time can also be defined for C_{veg} relative to each mortality process, e.g. $\tau_{mort,vitality} = C_{veg}/F_{mort,vitality}$. Turnover rates are the inverse of turnover time, i.e. $1/\tau$."

We do not think it practical to represent the hypotheses directly in these turnover equations. The hypotheses relate to how individual turnover fluxes are simulated in the TBMs, each of which has very different formulations (e.g. Table 3).

(2) Rethink fig 1. I don't get the meaning/purpose of figure 1. I think it is related to my point above, i.e., showing the diversity of processes contained in the remote-sensing based turnover observations but it did not help me. Turning this figure into a table may help. After reading the entire manuscript, I think I would have benefited more from a description of each of the terms with an example rather than the bars and arrows.

>> We think that the reviewer is referring to Table 1? This table is aimed at showing how different processes can contribute to changes in turnover and biomass both in models and the real world. It is not aimed especially at remote sensing-based estimates. We have made several modifications to the table following the comments of Reviewer 2 in order to make the examples more intuitive and better explain the meaning of the bars. We have also added the following explanation to the introduction:

"The individual mechanisms within these groupings are isolated within Table 1 so as to show how a particular perturbation in NPP, allocation, or turnover rate of woody or soft tissues (e.g. leaves, fine roots and fruits) would affect biomass or τ. Because both trees and ecosystems respond to environmental stimuli in a coordinated fashion, it is likely that many of these mechanisms will occur in concert."

(3) Thinning the results section. In my opinion the model comparison is the least developed part of the manuscript and I even doubt whether it is essential. If the definition gets better developed, it might be possible to derive the hypothesis from the definition and then discuss these hypothesis in the light of scientific literature. This would change the type of study but it could increase the impact of this study. If you decide to keep the model comparison, please, better justify the model experiment.

>> Our study is predicated on understanding why TBMs differ in their estimates of turnover time. Although Friend et al. (2014) identified divergence in turnover time projections for the 21st century, we are not aware of a global synthesis of baseline carbon turnover times nor a thorough assessment of the reasons for their divergence (as described in the second paragraph of the introduction). As such

we consider the model intercomparison to be a fundamental component of the study. It is consistent with the Model-Data Synthesis approach advocated by Medlyn et al. (2015), that TBMs can be used to identify hypotheses for further testing by observations. Although all of the hypotheses could have been formulated without the model intercomparison we would have had no indication of their importance for driving differences in turnover time estimates. We have revised paragraph 2 of the introduction to include the following sentences:
"Relatively little attention has focused on the representation of τ and its drivers in current vegetation models, with some but not all relevant dependencies represented in different models."
"The divergence that can be traced to TBM structure and parameterisation (Nishina et al., 2015) has not been closely analysed in terms of the contributions of specific underlying processes, interactions and driver dependencies, or their basis in knowledge from real world ecosystems."

(and add revision numbers for each of the models).

>> We will add revision or version numbers for each of the models to Table 2.

It would have been much easier to compare the models if a run with a prescribed PFT distribution was used as well.

>> Prescribed PFT distributions are only possible for a subset of models here (listed as average-individual in Table 2). Cohort- and individual-based models compete PFTs directly against each other within the same stand both in terms of vertical competition for light and horizontal competition for water and, in some models, nutrients. As such, their PFT distributions are emergent outcomes of the model simulation and cannot be fixed without fundamentally altering simulated stand structure.

How can you justify the comparison of data with management to simulations without management? How meaningful is this given that management is a major driver of both the growth and the mortality components of turnover?

>> Our study concentrates on comparing between TBMs (please see below), which made simulations using a consistent protocol. We have made comparison to satellite-based turnover times as an independent comparison, recognising that they also include management (line 243).

Including management in the simulations would have broadened the scope of the study from the topic of natural vegetation dynamics and greatly complicated the interpretation of an already complex response. We agree that management is an important driver of forest biomass turnover, however, to keep the task to a manageable level of complexity we explicitly exclude management from the scope of the study (please see lines 131-133).

Given the complexity of the processes described by turnover but the simplicity of the observations (i.e. a single number), the model comparison remains superficial in the sense that it is hardly possible to label some of the model behavior as "very unlikely". In the end this section takes up a lot of space for very little information (although I liked Fig 2 a lot. It is an informative way to show both models and data – note that this is the only figure that shows the observations).

>> Figures 1 and 2 both show the satellite-based method for estimating turnover time, although we would hesitate to call these data observations, as satellite-based NPP products derive NPP using a modelling approach and large-scale biomass products rely on empirical modelling approaches to extrapolate in space. We have included observations or alternative approaches where possible, but note that large-scale observational constraints on biomass turnover time are extremely limited. It is this lack of observations which has led to such a diversity of approaches and outcomes in TBMs. As

new products become available in the future it should indeed become possible to identify unlikely model behaviour. In the absence of these products, we have focused on decomposing the reasons behind the model responses (encapsulated in the hypotheses) and outlining how uncertainty might be reduced in the future.

Maybe the bulk of the comparison could be moved to the supplementary materials?

>> Much of the comparison is already shown in the supplementary materials. We prefer to keep most of the existing figures in the main text as they each underline important results relating directly to the derivation of the hypotheses. We have modified the selection of figures in the main text slightly, following suggestions of Reviewer 1.

References

Friend, A. D., Lucht, W., Rademacher, T. T., Keribin, R., Betts, R., Cadule, P., Ciais, P., Clark, D. B., Dankers, R., Falloon, P. D., Ito, A., Kahana, R., Kleidon, A., Lomas, M. R., Nishina, K., Ostberg, S., Pavlick, R., Peylin, P., Schaphoff, S., Vuichard, N., Warszawski, L., Wiltshire, A. and Woodward, F. I.: Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO_2., Proc. Natl. Acad. Sci. U. S. A., 111(9), 3280–3285, doi:10.1073/pnas.1222477110, 2014.

Medlyn, B. E., Zaehle, S., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hanson, P. J., Hickler, T., Jain, A. K., Luo, Y., Parton, W., Prentice, I. C., Thornton, P. E., Wang, S., Wang, Y., Weng, E., Iversen, C. M., Mccarthy, H. R., Warren, J. M., Oren, R. and Norby, R. J.: Using ecosystem experiments to improve vegetation models, Nat. Publ. Gr., 5(6), 528–534, doi:10.1038/nclimate2621, 2015.

Nishina, K., Ito, A., Falloon, P., Friend, A. D., Beerling, D. J., Ciais, P., Clark, D. B., Kahana, R., Kato, E., Lucht, W., Lomas, M., Pavlick, R., Schaphoff, S., Warszawaski, L. and Yokohata, T.: Decomposing uncertainties in the future terrestrial carbon budget associated with emission scenarios, climate projections, and ecosystem simulations using the ISI-MIP results, Earth Syst. Dyn., 6(2), 435–445, doi:10.5194/esd-6-435-2015, 2015.