

***Interactive comment on “A triple tree-ring constraint for tree growth and physiology in a global land surface model” by Jonathan Barichivich et al.***

**Response to anonymous Referee #1**

Received and published: 6 January 2021

**This paper describes the parameterization and testing of a quasi-mechanistic large-scale model of forest growth. It tests the model against two other models and against measured data. The novelty of the manuscript lies in these model tests, especially because they include stable isotopic data, which both illuminate the physiological processes that cause the growth differences and provide tests of the mechanistic basis of the model. The significance is that this attempts to link a global-scale land-surface model to three kinds of tree-ring data. If the model performs well, it may be justified to use it to describe the long tree-ring time series—potentially well beyond the range of remotely sensed, or even instrumental data.**

**The paper is mostly well-written, clearly significant, and appropriate for this journal.**

**I particularly enjoyed reading the introduction and the methods and materials, which provide access to this subject for a broad audience. The analysis represents a tremendous breadth of work. I heartily applaud the authors for building these isotopic tests into their models and appreciate the comparisons to other isotopically enabled models and to measured data. However, apart from the Intro and Methods, I found the paper difficult to read. There is so much here that the emphasis gets lost.**

**We thank the reviewer for recognising the scope and magnitude of our work and for constructively pointing to areas for improvement. As we explain below, we have taken on board the suggestions to improve the focus and highlight the key elements of our triple tree-ring constraint.**

**The abstract for a paper this complex should provide a roadmap that leads the reader to the main conclusion. It should mention not only ORCHIDEE, but also the other models, with a bit of explanation of why they were included. Also, Figure 1d-f is presented as a visual test of the models. If so, my visual impression is that MAIDENiso fails as the response surface looks quite different from that of the observations. This result should appear in the abstract. I suggest a change in the emphasis of the manuscript below. If accepted, this change should be reflected in the abstract.**

**We followed the advice of the reviewer regarding the improvement of the emphasis of the paper. In the revised abstract we now mention the other models as a brief model context for the performance of ORCHIDEE but we do not develop in their evaluation to avoid losing focus on ORCHIDEE. The poor performance of MAIDENiso is discussed in the new Discussion subsection 4.2 at the end of the model evaluation with the tree-ring triplet.**

**MAIDENiso is referred to as “specialized” in at least two places in the manuscript. The model is described briefly on page 3 L4-8, but I was left wishing for a clearer description of what makes it different. Like many of your readers, I have never used it. This will be especially important if you choose to emphasize Fig. 1d-f.**

The introduction is now more streamlined to highlight how the main growth processes are represented in current modelling approaches. It is hopefully much more clear where each model stands in the model landscape. We also added a compact description of the setup for the published simulations and processes included in MAIDENiso and LPX-Bern in the last two paragraphs of section 2.3 (Simulations)

**The simulated results are not always distinguished clearly from the empirical data. This is especially important because you are comparing the models to empirical isotopic data. In particular: 1/10-13: I presume all the “physiological” data here are simulated? If so, say so, especially in the abstract. Have there been any direct measurements of, e.g., GPP at the Fontainebleau site? The same question arises about source water below. As these are all simulated, they should be labelled as such (e.g., 13/6-9).**

We revised the text to make a clear distinction of when the description is based on simulated or observed data. The GPP referred to in the manuscript is simulated by ORCHIDEE. We did not use the short eddy-covariance measurements available for a forest near Fontainebleau in this study. Similarly, all source water is simulated and is now explicitly labelled as such.

**The LPX-Bern results are barely mentioned in the text and the only conclusion they lead to is that the model has “better isotopic forcing.” What does that mean? Does LPX-Bern use different algorithms to estimate source water and water vapour? If so, it would be interesting to see how the predictions compare. The fact that the LPX-Bern model works better than either of the others for  $\delta^{18}O$  dilutes the impact of the presentation of ORCHIDEE. I suggest, to create a clearer emphasis in the paper, to either move LPX-Bern to a supplement or to discuss it in more detail. A particularly interesting detail would be a discussion of what might be changed in future versions of ORCHIDEE and MAIDENiso to make them work as well.**

We followed the advice of the reviewer to keep the focus on the tree-ring triplet and deleted the former Figures 2-4 (the triplet with only leaf water enrichment, climate response and water use efficiency). In the revised manuscript, the former Figure 5 is now Figure 2 (Taylor diagram of model performance). This gave us space for the new Figure 3 to analyse more deeply the behaviour of LPX-Bern versus ORCHIDEE with respect to  $\delta^{18}O$ . The new Figure 3 shows a detailed comparison of the contributions of source water and leaf water enrichment to the simulated  $\delta^{18}O$  signals simulated by both models and also the correlation of each component of the forcing ( $\delta^{18}O$  of precipitation,  $\delta^{18}O$  of vapour and  $\delta^{18}O$  of source water) and the resulting simulations of  $\delta^{18}O$  in leaf water and cellulose. This Figure allows disentangling and comparing the nature of the simulated signals in both models and evaluating the consistency of the isotopic drivers used to force the models (i.e.,  $\delta^{18}O$  of precipitation and  $\delta^{18}O$  of vapour from the LMDZ atmospheric for ORCHIDEE and  $\delta^{18}O$  of source water (i.e. soil water) from the ECHAM5 atmospheric model for LPX-Bern).

The critical process and areas of improvement identified in the evaluation of ORCHIDEE against tree-ring data are now more clearly discussed in Section 4.1 and 4.2. The case of the issue of MAIDENiso is addressed in Section 4.2, but we refrained to develop detailed recommendations for this model since it was used only for comparison purposes. However we make clear that the cause of the poorly simulated growth-isotope surface is predominantly

due to the excessive carryover in simulated ring width, which decouples growth from leaf-level responses.

**Section 3.1.2: I'm not sure I understand the purpose of this long section although I've read it several times. I think it is being presented as a test of the relationships embedded in the model structures and parameterizations. If so, this seems important and the isotopic methods seem ideally suited to it. I would make this the main emphasis of the paper.**

Following the advice of the reviewer, the narrative flow of section 3.1.2 was revised and the section was renamed as "Simulated tree-ring triplet" to emphasize the triple tree-ring constraint. Former sections 3.1.3 (climate response) and 3.1.4 (20th century water use efficiency) were deleted to improve the focus of the paper.

**However, I noted that the MAIDENiso response surface looks really different from ORCHIDEE and from the data in Fig 1 d-f. I did not find this described clearly in the text. There was some description of the r-values of the partial correlations, but it is the slopes that catch the eye. The slope differences result in very different geometries across the response surfaces and this is what I would emphasize. Please note that the presentation of the response surfaces was interrupted by inferences about temperature and stomatal conductance, which I would move to the discussion. This section should end with a general model evaluation that addresses the visual impression that MAIDENiso has a problem.**

We thank the reviewer for pointing our attention to the slopes of the surface. Indeed, the simple regression slopes better quantify the geometry of the triplet and the coupling of the processes that it represents. We now labeled the regression slopes in Fig 1 from b1 to b3 and interpreted their magnitude in terms of processes. Since these inter-relationships are the basis of the novel tree-ring triplet presented in the paper we discussed the meaning of the slopes and relationships in terms of processes in the new section 4.2, which integrates the former sections 4.1 (Integrating tree-ring and carbon...) and 4.3 (Constraining model processes..).

**The manuscript also describes isotopic changes in response to climate change and CO<sub>2</sub>. Although this is an interesting application of the model, it seems to belong in another paper. This impression is strengthened by the fact that the analysis neglects recent discussion of the effect of height growth on isotope ratios (and presumably growth) (Brienen et al., 2017; Marchand et al., 2020; Marshall & Monserud, 1996, 2006; Voelker et al., 2016). If it is to remain, the height issue must be addressed and information about height growth in these trees should be added. Are these trees still young enough to be growing in height? How tall were they? It would be great to see these height effects added to some future version of the model!**

Following the advice of the reviewer, we decided to remove the analysis of 20th century iWUE and publish it as a follow up letter. The effect of tree height on carbon isotopes is not explicitly represented in ORCHIDEE. This has been shown to be particularly important in deep canopy tropical forests and certainly can account for initial trends during the juvenile period. The age of the stand in Fontainebleau is about 120 years (given in Table 1) and tree height is between 20-25 m. A newer version of ORCHIDEE introduced a better representation of forest structure and light penetration

(<https://gmd.copernicus.org/preprints/gmd-2020-29/>) and could be used to address this issue, particularly for dense tropical forests.

**The interpretation of tree-ring d18O data is notoriously difficult and the Scheidegger et al. approach, although clever, is too simplistic. Because the authors cite Roden and Siegwolf (2012) (19/13-19), I presume that they appreciate the difficulty, but they do not express it in a way that a naïve reader is likely to detect. I suggest clearly and bluntly recognizing these difficulties for the people who will follow down this path. Related to this problem is the question of how the source water and water vapour d18O were simulated for this analysis. It should be described, at least briefly. The results are contingent on how this was done and how well it worked. This is necessary in part because the source water data are emphasized, e.g., in Figs. 2 and 6.**

The uncertainties and critical processes identified for the simulation of the tree-ring variables are now explicitly addressed at the beginning of the Discussion in the new Section 4.1. We explicitly recognise the difficulty for simulating d18O in the 5th paragraph.

As mentioned earlier, the external d18O isotopic forcings used for each model (ORCHIDEE, MAIDENiso, LPX-Bern) are now better described in Section 2.3 (simulations). The new Figure 3 also helps to understand the origin of the differences and similarities between tree-ring dO18 simulated by ORCHIDEE and LPX-Bern.

**The temporal autocorrelation and its likely causes are interesting and important, but inadequately described. I would like to see a more carefully approach to this. In particular, there are mechanisms besides photosynthate carryover that could cause it. These include, for example, root or leaf mortality or production that might influence hydraulic balance in subsequent years. Monserud and Marshall speculate on some of these (2001). Whatever the mechanism, it would be great to have these effects described by the model and I support the emphasis placed on it.**

Since the carryover is one of the critical processes to simulate tree-ring width variability, we further developed the discussion of its causal factors in the first paragraph of section 4.1 of the Discussion. Factors other than carbohydrate remobilization are now mentioned and referenced.

**It would be unfortunate if the main points of this manuscript were missed or misunderstood because of the complexity of presentation. I urge the authors to emphasize the response-surface tests of the models. If so, they might also expand the discussion of LPX-Bern and its better performance, including a comparison of the source and vapour d18O simulations. I suggest dropping the climate-change analysis for now. Especially if the height effect were included in the model, the results would be significant enough to stand alone in another manuscript. Removing them from the current one would allow the model performance results to emerge clearly.**

We thank the reviewer for all the constructive advice, which we followed in order to improve the focus and presentation of the paper. The revised manuscript highlights the response surface of the tree-ring triplet and the comparison of ORCHIDEE with LPX-Bern as global model benchmark, including a concise analysis of the isotopic forcings.

Brienen, R. J. W., Gloor, E., Clerici, S., Newton, R., Arppe, L., Boom, A., Bottrell, S., Callaghan, M., Heaton, T., Helama, S., Helle, G., Leng, M. J., Mielikäinen, K., Oinonen, M., & Timonen, M. (2017). Tree height strongly affects estimates of water-use efficiency responses to climate and CO<sub>2</sub> using isotopes. *Nature Communications*, 8(1), 288. <https://doi.org/10.1038/s41467-017-00225-z>

Marchand, W., Girardin, M. P., Hartmann, H., Depardieu, C., Isabel, N., Gauthier, S., Boucher, É., & Bergeron, Y. (2020). Strong overestimation of water-use efficiency responses to rising CO<sub>2</sub> in tree-ring studies. *Global Change Biology*, 26(8), 4538–4558. <https://doi.org/10.1111/gcb.15166>

Marshall, J. D., & Monserud, R. A. (1996). Homeostatic gas-exchange parameters inferred from <sup>13</sup>C/<sup>12</sup>C in tree rings of conifers. *Oecologia*, 105(1), 13–21.

Marshall, J. D., & Monserud, R. A. (2006). Co-occurring species differ in tree-ring  $\delta^{18}\text{O}$  trends. *Tree Physiology*, 26(8), 1055–1066.

Monserud, R. A., & Marshall, J. D. (2001). Time-series analysis of  $\delta^{13}\text{C}$  from tree rings. I. Time trends and autocorrelation. *Tree Physiology*, 21(15), 1087–1102.

Voelker, S. L., Brooks, J. R., Meinzer, F. C., Anderson, R., Bader, M. K.-F., Battipaglia, G., Becklin, K. M., Beerling, D., Bert, D., Betancourt, J. L., Dawson, T. E., Domec, J.-C., Guyette, R. P., Körner, C., Leavitt, S. W., Linder, S., Marshall, J. D., Mildner, M., Ogée, J., . . . Wingate, L. (2016). A dynamic leaf gas-exchange strategy is conserved in woody plants under changing ambient CO<sub>2</sub>: Evidence from carbon isotope discrimination in paleo and CO<sub>2</sub> enrichment studies. *Global Change Biology*, 22(2), 889–902.

***Interactive comment on “A triple tree-ring constraint for tree growth and physiology in a global land surface model” by Jonathan Barichivich et al.***

**Response to anonymous Referee #2**

Received and published: 11 January 2021

**In the present study, Barichivich and co-authors explore processes and historical changes of tree growth and tree physiology in a land surface model by simulating three tree ring proxies (namely ring width, carbon and oxygen isotopes) and by comparing them to observations in temperate and boreal sites (one specific site in Fontainebleau and a network of 5 other sites encompassing deciduous and conifer tree species). Further, the land surface model performance is compared to two other models at site and network level. Such approach and evaluation of Land surface model for long term tree growth and tree physiology variability is relevant and will certainly contribute to the understanding of carbon uptake and evapotranspiration dynamics in forested ecosystems; and will improve the predictive skills of tree/forest growth and carbon water cycles responses to projected environmental changes.**

**The authors conduct thorough simulations and analyses and the study is well designed. There are a few major points that can be addressed or explained better to clarify the results and their implications and highlight the relevance of the study presented here.**

**1. The introduction can be refocused into the potential of existing tree ring data to evaluate LSM and why is such work relevant to specific global change questions. The authors mention that but never make the case for it. What knowledge will be gained in term of processes by simulating tree ring attributes and comparing them to observations and output of other models. How do the three models differ which will contextualize the results and the discussion of their performance, specifically ORCHIDEE which is the major one being evaluated.**

We thank the reviewer for the constructive advice. Following this comment we have revised and extensively rewritten the introduction to better explain the need for improvement of land surface models and why tree rings are especially suited for doing it at long timescales. The different modelling approaches are briefly explained and provide the context for process descriptions and the use of MAIDENiso and LPX-Bern for model-to-model comparisons.

**2. The results can be structured to better follow the study design. Site level (Fontainebleau) comparison of ORCHIDEE, MAINDENISO and observations and then the other sites where LPX-Bern model outputs are also used to compare with observations and ORCHIDEE. LPX-Bern is briefly described and then appears again in the discussion. In this regard the methods can clarify the forcing of all three models.**

We followed the advice of the two reviewers regarding the focus of the paper and organisation of the results and discussion sections of the manuscript. The new Figure 3 addresses the detailed comparison between ORCHIDEE and LPX-Bern and the external d18O isotopic forcings used for each model (ORCHIDEE, MAIDENiso, LPX-Bern) are now better described in Section 2.3 (simulations).

**3. The Discussion relies heavily on descriptive results and does not highlight the physiological processes (beyond the use of carbohydrates and even so, this point needs more careful consideration) that can potentially explain the model-data comparison (or mismatch). In this regard, uncertainties in of tree ring proxies and modeling assumptions (iWUE Farquhar model, leaf water enrichment model, source water  $\delta^{18}\text{O}$  forcing) are not addressed or discussed.**

The discussion was rewritten to recognise uncertainties and highlight the physiological interpretation of tree-ring data and the unique mechanistic value of combining growth and isotopes in the tree-ring triplet to constrain models. The uncertainties for each variable are explicitly discussed in section 4.1 (Uncertainties and critical processes...) and the physiological interpretation of the tree-ring triplet is discussed in section 4.2 (Constraining model processes with the growth-isotope tree-ring triplet).

**4. The references can be more updated in terms of recent efforts in using tree rings to benchmark process-based models but also to reflect the appropriate papers describing the mechanistic links between tree physiology and isotope variations in tree rings (specifically the O isotopes, beyond the review of McCarroll and Loader 2004).**

The revised introduction and discussion incorporate more updated essential references, some of which were suggested by the reviewer.

Detailed Comments:

**Pg1, Line 20: Their responses to what? Increasing atmospheric CO<sub>2</sub>, changing climate, disturbances?**

We revised the clause for clarity and added "...their simulated responses to environmental changes..."

**Pg2, line 12: A suggestion would be to change adapt and perish as follows: how trees perish or adapt to environmental change is still limited.**

Suggestion taken.

**Pg2, lines 14-15: additional references are relevant here specifically when using tree rings to either parametrize or evaluate mechanistic physiological models:**

- Lavergne, A. et al. Modelling tree ring cellulose  $\delta^{18}\text{O}$  variations in two temperature-sensitive tree species from North and South America. *Clim. Past* 13, 1515–1526 (2017).

- Belmecheri, S., Wright, W. E., Szejner, P., Morino, K. A. & Monson, R.K. Carbon and oxygen isotope fractionations in tree rings reveal interactions between cambial phenology and seasonal climate. *Plant. Cell Environ.* (2018).

- Lavergne, A. et al. Historical changes in the stomatal limitation of photosynthesis: empirical support for an optimality principle. *New Phytol.* 225, 2484–2497 (2020).

In this part we refer to some relevant references for empirical studies using tree-rings to infer ecological changes. Their application for modelling is introduced later. We added the reference of Lavergne et al 2020.

**Pg2, lines 19-20: These references correspond mostly to mature trees exposed to elevated CO<sub>2</sub>. The present study investigate historical records and model simulations of tree response to gradual increase of atmospheric CO<sub>2</sub>. As such, this ought to be highlighted as well.**

To reflect this point of the reviewer, the sentence was revised as follows "...to study the range of historical responses of mature trees to gradual global change or manipulative experiments"...

**Pg2, lines 25-30. This statement is misleading. Using a concept such as "cursed" imply an inherent unsuitability of ring width proxy for growth reconstructions. This is not true if the sampling strategy is adequately designed for that purpose. Indeed, the ITRDB repository includes trees collected mainly for climate reconstructions and it is well known that when using the same data for inferences of growth and specifically productivity, the data will reflect the growth dynamics and sensitivities of old, mature, climate sensitive individuals. It is not clear what is the point being made by the authors here? Why not test then model assumption based on collection specifically made for growth/productivity reconstructions? There are a few existing records (ecological sampling methods applied in Flux tower sites for e.g.).**

This paragraph introduces tree rings as a means to infer long-term changes in growth and physiology, highlighting their advantages and potential pitfalls. The magnitude and reach of the biases is still a contentious issue in the community as is reflected in this part of the text. The text was revised to better reflect that the issue applies mainly to the tree-ring data archived in the ITRDB and not inherently to all tree-ring data. We added a closing sentence stating that "The effect of sampling biases on long-term growth trends can be effectively minimized by using appropriated sampling designs for productivity reconstructions (Gough et al., 2008; Nehrbass-Ahles et al., 2014; Dye et al., 2016). "

**There is a great potential to tap tree ring data to benchmark LSM. The introduction can make a stronger case for the use of both ORCHIDEE and MAIDEN iso. Why compare both models and what information or improvements can be gained from using then ORCHIDEE.**

We agree. This is the aim of the paper. We revised the introduction and now we better explain why these models were used to provide a wider context for the few tree-ring enabled models.

**P4, Line 25, it is not clear whether the soil hydrology was modeled using an older version compared to the multi-layer cited after. If so, what is the motivation for this choice. Otherwise, it unnecessary to cite/describe what is not used.**

Point taken. The mention to the new multi-layer soil scheme was deleted.

**P7 line 18, Where does the assumption of the effective path length of 8 mm comes from? How is this universally applied to different tree species/locations? See Roden et al. 2015, - Roden, J., Kahmen, A., Buchmann, N. & Siegwolf, R. The enigma of effective path length for 18O enrichment in leaf water of conifers. Plant. Cell Environ. 38, 2551–2565 (2015).**

Thanks for pointing out this important source of uncertainty for simulating dO18 in global models limited to a few forest PFTs. LPX-Bern and ORCHIDEE have a similar representation of the Peclet effect and isotopic mixing during cellulose production but the Peclet parameterization differ with respect to the highly uncertain and variable L (path length). A value of 3 mm for all PFTs was used in LPX-Bern, which is almost three times lower than our L value. Our value of 8 mm was obtained by Risi et al., (2016) from tuning the model against seasonal isotopic observations in a few mid-latitude sites in Europe. The



impact of this high parameter value in ORCHIDEE is evident when compared with LPX-Bern because ORCHIDEE simulates a much stronger imprint of source water and dampening of the leaf enrichment signal in tree-ring  $\delta^{18}\text{O}$ . This source of uncertainty and its consequences are discussed in the 6<sup>th</sup> paragraph of section 4.1.

**P11, lines 9-13. Why was this approach used to evaluate the relative contribution of source water versus evaporative enrichment, this is a statistical inference and will not reflect the mechanistic relationship between cellulose and leaf/source water  $\delta^{18}\text{O}$ . For tree ring observations, a more adequate test would be using a proxy forward model (Evans et al., 2006) to evaluate how recorded  $\delta^{18}\text{O}$  in tree ring cellulose compares to the modeled one using input of source water from observations (when available) or from the LMDz; and how varying source water and relative humidity (or VPD) affect the results. In the model world, these parameters are also used to simulate tree ring  $\delta^{18}\text{O}$  and could similarly be compared first to the observations in order to evaluate the relative role of source versus leaf water evaporative enrichment.**  
-Evans, M. N. et al. A forward modeling approach to paleoclimatic interpretation of tree-ring data. *J. Geophys. Res. Biogeosciences* 111, (2006).

We opted for this statistical approach because it allows a simple partitioning of the  $R_{\text{sq}}$  due to source water and leaf water enrichment in simulated  $\delta^{18}\text{O}$ . This was not attempted for the observations because the real isotopic  $\delta^{18}\text{O}$  forcing in source water is unknown. We are aware of Evans' model, but adding another model to our comparison is out of scope.

**This also brings the point of potential uncertainty related to the assumptions of the péclét effect and the fraction of oxygen atom during cellulose synthesis. These were shown to vary along aridity gradients and intra-seasonally (Cheesman and Cernusak 2016) and with cell-size (lumen area, Szejner et al., 2020).**

-Szejner, P., Clute, T., Anderson, E., Evans, M. N. & Hu, J. Reduction in lumen area is associated with the  $\delta^{18}\text{O}$  exchange between sugars and source water during cellulose synthesis. *New Phytol.* 226, 1583–1593 (2020).

-Cheesman, A. W. & Cernusak, L. A. Infidelity in the outback: climate signal recorded in  $\Delta^{18}\text{O}$  of leaf but not branch cellulose of eucalypts across an Australian aridity gradient. *Tree Physiol.* 37, 554–564 (2016).

As explained above, these sources of uncertainty and their consequences for ORCHIDEE are now discussed in the 6<sup>th</sup> paragraph of section 4.1.

**Replace carrying over with carryover and specify when first mentioned that it is a carry-over of carbohydrates from previous year or season. This could be discussed in more detail: the dynamic of stored versus recently assimilated photosynthates throughout the growing season (conifer vs deciduous) and under climate extremes (droughts) or other disturbances.**

Thanks for pointing the mistake in the wording. It was corrected. The carryover and its causes is described in more depth than before in the first paragraph of section 4.1 in the Discussion.

**In the results section, when comparing ORCHIDEE and Maiden iso models, it is not clear whether the input data for running both models were the same (e.g. Source water, meteorology, etc)?**

The external  $\delta^{18}\text{O}$  isotopic forcings for the already published simulations of MAIDENiso and LPX-Bern were different. This is now better described in Section 2.3 (simulations).

**P12, lines 25. The effect of stomatal conductance on isotopic discrimination should also be recorded in leaf water enrichment. It is obviously not the case since  $\delta^{18}\text{O}$  does not show a linear relationship with ring width. How would you explain this decoupling of isotopic responses to a reduction in stomatal conductance?**

The dual carbon-oxygen isotope relationship is significant ( $r=-0.40$ ,  $p<0.001$ ), therefore stomatal conductance is recorded in leaf water enrichment. The non-linear and somewhat weaker relationship between ring width and oxygen variations points to a decoupling of growth and leaf responses, which arises from the growth memory present in the observations as growth would use a mixture of older and new carbon. This is discussed in the 3<sup>rd</sup> and 4<sup>th</sup> paragraphs of Section 4.1.

**Atmospheric  $\text{CO}_2$  and  $\delta^{13}\text{C}$  data used for calculation and simulation of  $\delta^{13}\text{C}$  should be revised to the most updated datasets. This is discussed in details in Belmecheri and Lavergne 2020 with suggested recommendation for historical data.**

**-Belmecheri, S. & Lavergne, A. Compiled records of atmospheric  $\text{CO}_2$  concentrations and stable carbon isotopes to reconstruct climate and derive plant ecophysiological indices from tree rings. *Dendrochronologia* 63, 125748 (2020).**

McCarroll and Loader (2004) dataset has long been the standard in the community. As Belmecheri and Lavergne pointed out, the main impact of the different sources occurs on the absolute values. Although we appreciate the suggestion of the reviewer, changing the  $\text{CO}_2$  and  $\delta^{13}\text{C}_{\text{atm}}$  data would not change the interannual variability of carbon discrimination and  $i\text{WUE}$ , which is the main focus of our model-data comparisons. Therefore, we don't believe this change to be necessary given our focus on correlations.

**P14, lines 12. What is the evaluation metric for “well simulated” The model simulate less than 40% of the observed variability with 70% of unexplained variance. How do the authors assess that the model performance is good? Figure 5 is a great visualization for model performance summary. The result section can rely on this Figure for a consistent description of model performance and model-data comparison. For instance, “moderate” category is never cited in the result text.**

Although this is subjective and relative to the field, well simulated or well reconstructed in a tree-ring context is usually anything higher than 25-30% of the total variance. This corresponds to a correlation  $r \geq 0.5$ , which qualitatively in our Fig.2 (former Fig. 5) falls in the moderate category. We understand the point of the reviewer and we are more explicit with the qualitative adjectives of performance, linking them to the Fig. 2. We added the following sentence in the caption of the figure “Four qualitative areas of performance in terms of the magnitude of correlations or simulated variance are indicated as a visual aid”

**P14, lines 14. How could the model parametrization be biased towards temperate and deciduous forests since PFTs are informed in the model and climate drivers should reflect the forest/tree growing conditions (knowing that Maiden iso was calibrated for Fontainebleau site, but for ORCHIDEE?). In this case, which parameters are thought to be biased towards temperate forests.**

Given that this sentence is not extremely insightful and was not correctly understood by the referee it was deleted. This sentence was intended as a general concern and did not refer a specific process or a set of parameters. It expressed a general concern about ecological research and thus mechanistic modelling based on the outcomes of this ecological research. Much of our ecological understanding is based on research in temperate ecosystems, this understanding dominates textbooks and university courses and as such enters our land surface models. In the absence of knowledge specific to the boreal and tropical biomes the models fall back on our understanding (and parameters) of temperate ecosystems. This may have resulted in a bias of these models towards the temperate zone. A clear examples of the temperate bias can be found, for example, in how LSM formalize croplands (one rotation per year, homogeneous system, short growing seasons, fallow in between rotations) which represents intensive cropping in the temperate zone but does not represent at all cropping over most of Africa.

**P14, lines 15-17. If autocorrelation is removed from observations, does it improve model/data comparison since ORCHIDEE simulates poorly carbohydrate carryover.**

Doing prewhitening to the observations does improve slightly the model-data correlations in the French sites where autocorrelation is stronger. We have added two sentences about this possibility in the 1<sup>st</sup> paragraph of section 4.1 of the Discussion: “The carryover of ring-width observations can be removed through statistical prewhitening (Cook, 1985). A data-model comparison using prewhitened observations results in slightly higher correlations with simulated ring-width variability in the three French sites (e.g., increase in  $r$  from 0.50 to 0.56 in Fontainebleau) but not in the Finish sites (not shown), where autocorrelation was lower (Fig. 6a).”

**P14, lines 19-20. Could the amplitude discrepancy between observed and simulated  $\delta^{13}C$  be explained by post photosynthetic fractionations?**

The Farquhar model referenced in the paper and used in the present study describes the isotopic discrimination at the leaf level (See Frank et al. 2015). It is not clear from the methods that ORCHIDEE takes it into account nor that the measurements of tree rings were scaled to the leaf level.

**-Frank, D. C. et al. Water-use efficiency and transpiration across European forests during the Anthropocene. Nat. Clim. Chang. 5, 579–583 (2015).**

The caveats of using this simple formulation for carbon discrimination in land surface models were explicitly described in the last paragraph of section 2.1.2 in Methods. As the reviewer points out, it assumes that post-photosynthetic and mixing processes are negligible, which admittedly is a major simplification. However, as we mention it, it is still a useful model of carbon discrimination to constrain the environmental response of land surface models with tree-ring data as demonstrated in earlier studies (e.g., Bodin et al., 2013, Keel et al., 2016).

The possible causes for the underestimation of the interannual variability of carbon discrimination were also explicitly addressed in the former discussion section 4.2 L15-23. We stated that the amplitude of simulated carbon discrimination in Fontainebleau was very sensitive to soil depth and maximum photosynthetic capacity ( $V_{cmax}$ ) parameters. Therefore, beside the possible effect of post-photosynthetic processes, the underestimation of the variability is in part due to these parameter values in the standard parameterization of the model PFTs. This is the message for the reader in the revised discussion (4<sup>th</sup> paragraph of section 4.1).

The sensitivity of the simulated tree ring  $\delta^{18}\text{O}$  depends on the selected months of model outputs. In the methods, the authors describe selecting May-August. Was this informed by knowledge of the growing season? The authors stated that using this window ensures a standard time window to compare all sites and isotopes (although different time windows for C and O isotopes did show different levels of agreement between observed and simulated isotopic variations). While the choice of a standard time-window can be justified for the reason outlined by the authors, it is not clear that such justification is beneficial when there is a loss of statistical agreement between data/model and if this choice is not informed (even roughly) by the tree's growing season. An important consideration for cellulose  $\delta^{18}\text{O}$  is the timing and duration of cell-wall thickness during which most of the cellulose is deposited. This will determine the isotopic ratio recorded in tree rings and the time window not only vary by latitude/altitude but can be narrower than anticipated (See Cuny et al. 2015).

-Cuny, H. E. et al. Woody biomass production lags stem-girth increase by over one month in coniferous forests. *Nat. Plants* 1, 15160 (2015).

In addition, because a fraction of leaf water will exchange with xylem water, C isotopes will carry more signal of previous carbohydrates compared to O isotopes (dampened carryover signal).

As we stated in the methods, the choice of a fixed window is a necessary compromise to ensure comparability between the isotopes across the sites. Although not explicitly said in the text, an exploration of different windows showed that a sufficiently wide window such as May-August better integrates the seasonal responses of simulated photosynthesis to soil moisture and temperature along the climate gradient, which are then comparable with the seasonally integrated observations. Shorter or longer seasons did not change substantially the correlations between simulated and observed carbon discrimination, but the oxygen isotope was sensitive to May. Michelot et al. (2011) showed that in Fontainebleau the growing season (photosynthesis) of oak starts in mid-April, reaching a peak in transpiration and photosynthesis in late-May to early-June, which coincides with the transition from earlywood to latewood. This helps explaining the sensitivity to May and also informed our choice for the May-August window. To address the concern of the reviewer we added the following sentence "Late-May and early-June correspond to the seasonal peak in transpiration and photosynthesis of oak in Fontainebleau, which is closely followed by the transition between earlywood and latewood (Michelot et al., 2011)."

Like other LSMs, ORCHIDEE (r898) does not consider the intraseasonal dynamics of non-structural carbohydrates. Thus, it does not capture the fine-scale temporal asynchrony between photosynthesis and carbon expenditure in wood growth and respiration as is seen in the observations (e.g., Cuny et al 2015). However, the wood formation process is highlighted in the conclusion as an area of improvement for the representation of tree growth in the model.

**Pg 15, lines 5-10. Intuitively, GPP is expected to correlate with D13C, yet it is not the case here and D18O correlating better with GPP is rather intriguing. What is the mechanistic link to explain evaporative enrichment correlation with GPP. If this is driven by stomatal conductance so should D13C.**

This point was addressed at the end of the Discussion in former section 4.4. Now it is addressed in the 1<sup>st</sup> and 2<sup>nd</sup> paragraph of section 4.3. D13C does correlate significantly with GPP as expected (shown in former Fig. 8 and now Fig. 6), but the correlation with  $\delta^{18}\text{O}$  with GPP is more consistent spatially because of the synergistic effect between source water and leaf enrichment.

**Pg16, Lines 25. This can be easily tested by removing auto-correlation from observed TRW prior to comparison with modeled TRW. As a side note (not a criticism to the present study). RW simulation can be tested using the Vaganov–Shashkin (VS) model to simulate TRW and compare ORCHIDEE performance to the VS model (similar approach to comparing ORCHIDEE and MAINDEN iso). This may shed more light into the poor ORCHIDEE performance in reproducing high frequency TRW variability.**

We thank to the reviewer for pointing this out. As explained in the response above we added two sentences about the impact of using prewhitened observations in model-data comparison in the 1<sup>st</sup> paragraph of the Discussion.

We are very familiar with the use of the VS model and could have certainly been an alternative to MAIDENiso, however, it is based only on sink-driven growth and does not simulate stable isotopes. In addition, it would work only for the two coniferous sites. The semi-empirical alternative VS-Lite model applied to any species but is not process-based.

**Pg16, Lines 35. This assumption of drought legacy recovery has been tested for triple- proxies such as this study and the outcome depends largely on the detrending methods used for TRW. It also depends on the frequency of droughts. Hence, the results obtained by ORCHIDEE simulation might not be due only to a poor performance of the model.**

- Szejner, P., Belmecheri, S., Ehleringer, J. R. & Monson, R. K. Recent increases in drought frequency cause observed multi-year drought legacies in the tree rings of semi-arid forests. *Oecologia* 192, 241–259 (2020).

The authors make a case of the 1976 drought to discuss limitation of the processes represented in the various models. The impact and recovery from the 1976 drought could be further elaborated using Superposed Epoch Analyses on simulated and observed tree ring proxies.

SEA would be useful to statistically quantify the mean pattern of response to extremes in the model, but the result of such an analysis is obvious given the large differences in persistence between simulation and observations. It would also be redundant beside the autocorrelation analysis presented, which we believe illustrates the issue.

**Pg 17, lines 25-30. This still contrast with an increase of D13C documented in atmospheric CO<sub>2</sub> (Keeling et al., 2017). There is no discussion about the discrepancy between continuous increase of simulated D13C and the “pause” of D13C tree ring measurements since the 1980.**

Keeling, R. F. et al. Atmospheric evidence for a global secular increase in carbon isotopic discrimination of land photosynthesis. *Proc. Natl. Acad. Sci.* 114, 10361–10366 (2017).

The analysis of water use efficiency was discarded following the advice of reviewer 1 to improve the focus of the paper on the tree-ring triplet. The ongoing reduction of the rate of increase in water use efficiency in recent decades has been documented in recent papers and will be the subject of a forthcoming paper with the model.

**P18, lines 5. Are the simulated historical trends of  $\delta^{18}\text{O}$  consistent with other observations (paleoclimate studies) in other forests, climatic regions?**

We only focused our analysis on the interannual variability of  $\delta^{18}\text{O}$  and the analyses of trends is interesting but would be out of the scope, since it would add another dimension to the already long paper.

**P18, lines 15. How about the mechanistic representation of source water vs leaf water enrichment contribution in LPX-Bern model? If the forcing for the three models are not the same (specifically for source water  $^{18}\text{O}$ ), the comparison of model performances is then biased.**

This is fully addressed in the new Fig. 3. Although the isotopic forcings of precipitation and vapour are different, the simulated isotopic signatures of source water and leaf water enrichment are significantly correlated between the two models because of the strong integration effect of soil water and the effect of the atmospheric forcing (VPD).

**P18, lines 30. Could it be that the sensitivity of stomatal conductance to factors other than  $\text{CO}_2$  be misrepresented (soil moisture for example).**

Yes. This is discussed in the 3<sup>rd</sup> and 4<sup>th</sup> paragraph of Section 4.2.