Reviewer #2

Breil et al present a quantification study of afforestation impacts on regional climate in Europe in the late 20th and early 21st century. This topic is of interest due to afforestation effects masking increasing temperatures from climate change during this period. The novelty of the work lies in explicitly quantifying this impact.
Reply: Thank you very much for reviewing this manuscript, for your assessment and your helpful comments. Detailed answers to your comments can be found below. Changes in the revised manuscript will be implemented with tracked changes.

While the core concept and method of the manuscript are sound, the current presentation reads more like a textbook chapter on climate than a current research article. This is due to a high level of repetitiveness throughout the text as well as weakly developed greater context, mostly in the discussion and conclusion sections. The authors also present results from Wilcoxon Rank Sum Testing without including this technique or research goal in their methods.
Reply: The high level of repetitiveness (and this redundancies) was criticized by both reviewers. Therefore, we removed parts of the text where a repetition actually occurred:
Lines 42-47
Line 244
Lines 425-429

However, we think that this apparent repetition of results arises from the fact that afforestation in Europe has similar effects on the surface energy balance in summer as well as in winter. By dividing the results into local and non-local effects, this impression was possibly reinforced. However, the weighting of the individual afforestation effects varies in the different regions of Europe. Addressing these processes in each region and in each season is key in order to be able to explain the effects of afforestation on the surface energy balance conclusively. We thus left most of the process description as it is to ensure that all afforestation effects are described in a comprehensible way.

Moreover, we have placed the results of our study in the revised manuscript in a broader context and expanded the discussion according to your suggestions (see comment below).

Furthermore, the Wilcoxon Rank Sum test is now included in the method section in the following way:
"The statistical significance of the temperature changes in AFF in comparison to REF is calculated with a Wilcoxon-Rank-Sum-Test, a non-parametric statistical test analyzing the differences between two paired datasets." (Lines 141-143)

Possible additional context:
There are a variety of topics that come to mind to give the paper greater context and interest. Given the results, what might the impact of additional afforestation in Europe be? In other regions? How might the results be different if different vegetation types were modelled (ie., evergreen broadleaved trees)? How did the strength of temperature effects vary with size of afforested region?
Reply: Thank you for your helpful suggestions. We have extended the discussion and incorporated your suggestions as follows:

Regarding the impact of additional afforestation and the size of afforestation:
"These general effects of afforestation on the surface temperatures in summer seem to be independent of the afforested area, as shown by the results of coordinated model

intercomparison studies with idealized afforestation scenarios. For instance, Davin et al., (2020) and Breil et al., (2020) show that afforestation would have the same local temperature effects if the whole European continent would be afforested." (Lines 406-410)

Regarding the dependency on the afforested region:
"A general reduction of the surface albedo, an increased release of turbulent energy into the atmosphere and a resulting local cooling in summer are also described by Burakowski et al., (2018) for North America. This indicates that the results of this study may be representative for afforestation in the mid-latitudes and transferable to other regions." (Lines 437-440)

Regarding the impact of the forest type:
"In addition, the results of this study are only valid for evergreen needleleaved trees and deciduous broadleaved trees that are characteristic for the mid-latitudes. Other tree species, like for example evergreen broadleaved trees or deciduous needleleaved trees can of course have other effects on the local surface energy balance and consequently induce other remote effects. The described afforestation effects in this study could therefore be both, stronger and weaker." (Lines 474-478)

Introduction
L42, Replace biogeophysical with physical, as that is what is meant
Reply: We would like to keep the term "biogeophysical" as this term is commonly used to explicitly distinguish between the effects of afforestation on the exchange of energy, water, and momentum (biogeophysical) and the effects on the carbon cycle (biogeochemical).

L42-45 are nearly identical with L73-75.
Reply: We agree and removed lines 42-47 from the introduction (see also first point above).

Methods
Local vs Non-local, defining all non-afforesting grid cells as non-local seems like a generous definition for the regions (as opposed to say, just the surrounding grid cells within a particular afforesting patch). Please justify this choice.
Reply: If only the surrounding grid cells of an afforested area would be considered for investigating non-local effects, one would assume that non-local effects only have an impact at the regional scale. However, atmospheric feedbacks resulting from local afforestation can also have remote effects, for example due to changes in atmospheric moisture content, cloud cover, downwind precipitation, or more generally in terms of the atmospheric circulation.
Since the changes in forest cover fraction are the only difference in the simulation setup between REF and AFF, changes in non-forested grid cells must be caused by these indirect non-local atmospheric feedbacks. This includes changes in all grid cells that are beyond the afforested areas and its surroundings. It is therefore common practice in model studies that all changes in grid cells which are not afforested are considered as non-local effects (e.g. Winckler et al., 2017, Winckler et al., 2019, Pongratz et al., 2021).

Results
L363-375, The text here can be reduced and figures moved to supplementary materials.
Reply: In the revised manuscript, the text is shortened and figures are moved to the supplementary materials.

Discussion

L400 – 411, Be more specific about the effect seen in other studies (ie., list their values in a comparable way to your work). How did these other studies deal with local vs non-local effects?

Reply: The effects in the cited studies are now specified in the revised manuscript and the previous text is replaced as follows:

"In contrast, the small local warming effect in winter is quite astonishing, since it is generally assumed that afforestation is associated with a pronounced warming in the mid-latitudes in boreal winter, as for example shown by Lejeune et al., (2017) for North America. Using Land-Use and Climate, Identification of Robust Impacts (LUCID) models and Phase 5 of the Coupled Model Intercomparison Project (CMIP5) models, Lejeune et al., (2017) provided evidence that the snow-masking effect of forests (e.g. Essery, 2013) is clearly pronounced in North America. In combination with slightly increased evapotranspiration rates, winter temperatures of forests are about 0.3 K (LUCID) and 0.4 K (CMIP5) higher than those of other vegetation forms. However, the snow-masking effect is less pronounced in Europe than in North America, as shown by Asselin et al., (2022) within the framework of an idealized afforestation experiment for Europe and North America. They could show that snow-masking reduces the surface albedo on both continents in a similar way, but the reduced surface albedo effect on the surface temperatures is in North America much stronger than in Europe. For the same latitude, European climate is warmer than the climate in North America, and snow cover in winter is consequently restricted only to higher latitudes, notably central and northern Europe. There, insolation is low in winter and thus, the albedo effect on surface temperatures is small. The same conclusions were drawn by Strandberg & Kjellström, (2019) from regional climate simulations with an idealized afforestation scenario for Europe." (Lines 413-427)

As already mentioned above in our comment about the definition of local and non-local effects, it is common practice in model studies to define changes in grid cells with no afforestation as non-local effects.

Conclusion

The climate masking point should be expanded and moved into the main body of the discussion. Conclusion should be saved for reiterating your main messages from the discussion, not introducing new synthesis of the results.

Reply: We agree and moved this part of the conclusion in the discussion section as you suggested (Lines 485-497).

References:

Pongratz, J., Schwingshackl, C., Bultan, S., Obermeier, W., Havermann, F., & Guo, S. (2021). Land use effects on climate: current state, recent progress, and emerging topics. Current Climate Change Reports, 1-22.

Winckler, J., Reick, C. H., & Pongratz, J. (2017). Robust identification of local biogeophysical effects of land-cover change in a global climate model. *Journal of Climate*, *30*(3), 1159-1176.

Winckler, J., Lejeune, Q., Reick, C. H., & Pongratz, J. (2019). Nonlocal effects dominate the global mean surface temperature response to the biogeophysical effects of deforestation. *Geophysical Research Letters*, *46*(2), 745-755.