Reviewer #1

The paper uses CCLM in order to investigate what effects afforestation in the year 1986 to 2015 had on the climate in Europe. I think the paper has a clear idea and research agenda and properly executes on it. From my point of view an improvement is needed in the validation.
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.

The paper discusses effects on temperature and precipitation but only looks at the model performance with respect to temperature. I think a discussion similar to temperature but looking at precipitation is needed. Even though the authors look at averages and extremes of the two variables, the validation only looks at averages. I think one can live with that but, again, a look at precipitation (averages) is needed.
Reply: Thanks for pointing this out. We agree with your assessment that a model validation with respect to precipitation is important and have now included a validation for precipitation in the revised manuscript. The discussion of the model performance is extended in the following way:

"First, we analyze the capability of CCLM-VEG3D to reproduce the general climate conditions in Europe. Figure 2 shows the differences between the reference simulation (REF) and the ERA5-Land reanalysis (Muñoz-Sabater et al., 2021) for (a) the yearly mean 2 m temperatures and (b) the yearly total precipitation sums during the period 1986-2015.
A warm bias is simulated over most parts of Europe in the reference simulation, extending from Southern Europe over Central Europe to Eastern Europe. However, these deviations to ERA5-Land are in the same range as the biases of other RCMs, as demonstrated by Kotlarski et al., (2014). Regarding Northern Europe and the British Isles, the simulation results agree well with the reanalysis data.
Total precipitation sums are underestimated in CCLM-VEG3D in southern and western Europe, but overestimated in eastern and parts of northern Europe (shown as a percentual deviation in Fig. 2). This is also true for the mountainous regions of the Pyrenees and the Alps. On the other hand, the simulated precipitation sums agree well with the reanalysis data over large parts of Central and Eastern Europe as well as of southern Scandinavia. Thus, the results of CCLM-VEG3D reflect the already known precipitation pattern of regional climate simulations with CCLM (Kotlarski et al., 2014).
Therefore, although a certain model bias for the simulated 2 m temperature and the total precipitation sums is found, the simulation results of CCLM-VEG3D are comparable with the results of other RCMs (Kotlarski et al., 2014) and we conclude that the model is generally able to reproduce the general climate conditions in Europe." (Lines 178-195)

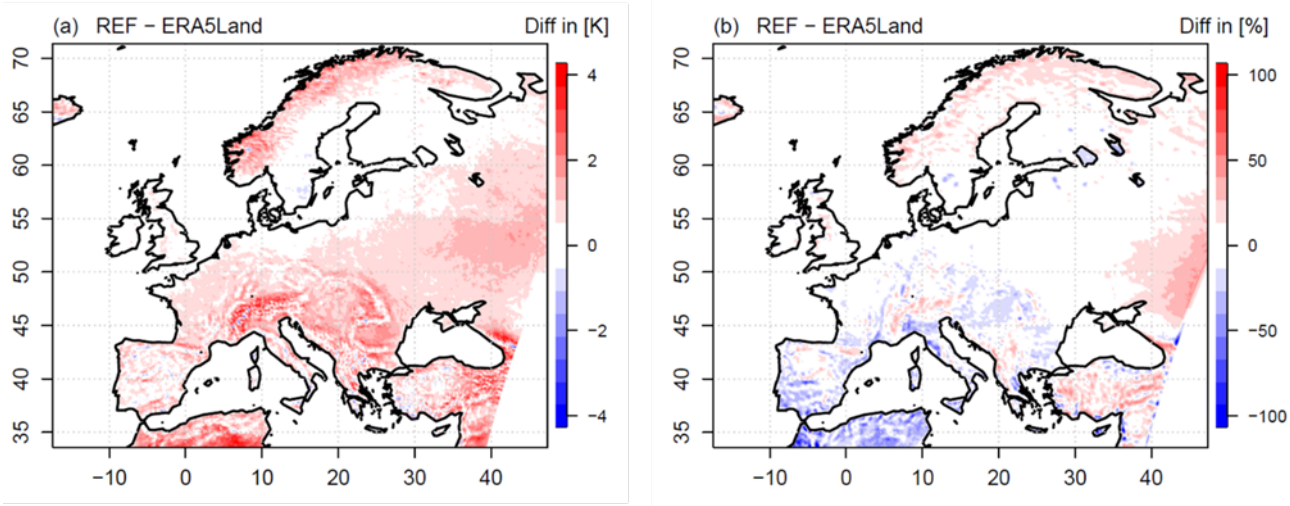

Figure 2: Differences in the (a) yearly mean 2 m temperature and (b) the percentage deviation in the yearly mean total precipitation sums between REF and the ERA5-Land reanalysis for the period 1986-2015.

Could you indicate or mention in Figures 3 and 4 which values were significant?

Reply: We thank the reviewer this comment. In general, the visualization of local effects is very challenging, since afforestation took place only on small spatial scales and isolated locations during the period 1986-2015 (see Figure 1b). Highlighting significant local effects in individual grid cells would consequently be very difficult at this high spatial resolution, and the local changes would not be visible. Therefore, we have decided for visualization purposes to summarize the local effects of afforestation for three characteristic sub-regions (NE, CE, SE). However, this means that both significant and non-significant changes in individual grid cells are included in Figures 3 and 4 and summarized in one bar. In order to be able to visualize the fraction of significant local changes, we decided to divide the bars in Figure 4 in two fractions, a significant one, which is highlighted with dashed lines and a non-significant one, which is filled blue, thus indicating the relationship between significant and non-significant results for each region and season.

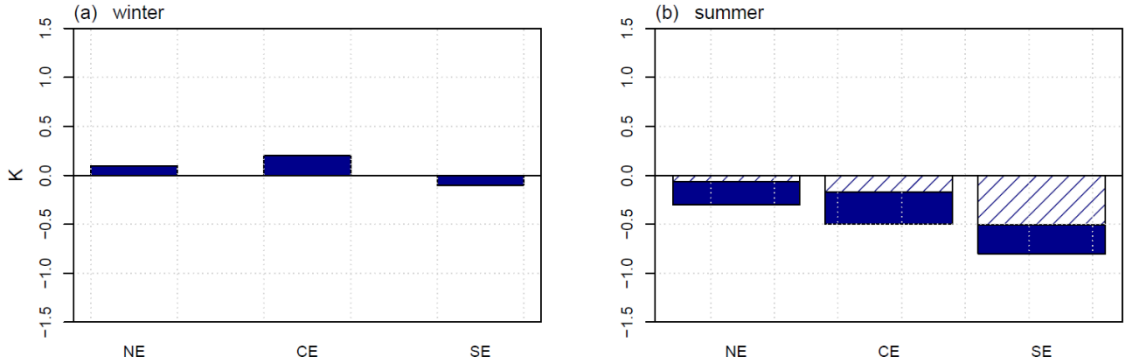

Figure 4: Local effects of afforestation (AFF-REF) on the mean surface temperature in (a) winter, and (b) summer for the three subregions NE, CE and SE. The fraction of significant local effects in the respective subregions (calculated with a Wilcoxon-Rank-Sum-Test at a 95 % level) are indicated by dashed lines.

Maybe 2 stylistic notes:

The authors repeat the mechanisms that lead to the efects very often. I think it is good to repeat things but maybe not so often.

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.

[Disclaimer: not a native speaker] The English word "whereby" means "by, through", it doesn't match the German "wobei".
Reply: Thanks for the comment. We rephrased the corresponding text and removed the word "whereby" in the respective sentences.

Other minor notes on English: 294 "visible" -> "can be seen"
Reply: is replaced

320 "stronger pronounced" -> "pronounced more strongly" (occurs several times)
Reply: is replaced throughout the whole manuscript.

414 is "buoyancy" the correct word here (not saying it's wrong and not entirely sure what you wanna say, but I only know it from liquids, do you mean something like convection?)
Reply: We really mean buoyancy in the hydrostatic sense. According to Archimedes' principle, the buoyancy of a body in a medium (fluid or gas) is equal to the weight of the medium displaced by the body. Transferred to the atmosphere, this means that an air mass with a high density (cold air mass) sinks, and an air mass with a low density (warm air mass) is lifted until the density of the air mass is equal to the surrounding air. Thus, buoyancy is the driving force behind convection.
The strength of turbulent heat fluxes from the surface to the atmosphere depends on two factors: buoyancy and wind shear. In southern Europe, buoyancy is generally higher than in northern Europe due to the generally higher near surface temperatures and the resulting larger density difference to the overlaying atmosphere. In combination with the higher surface roughness of forests and the associated increased wind shear, afforested areas in southern Europe can transform solar radiation into turbulent heat very efficiently.

In order to prevent ambiguities, we rephrased the corresponding text in the following way:
"In southern Europe, where insolation is higher, snow cover plays a minor role for the surface energy balance. Surface temperatures are typically higher than for central and northern Europe, and therefore, buoyancy is generally stronger in this region. In combination with the higher surface roughness of forests and the associated increased wind shear, afforested areas in southern Europe are consequently able to transform this increased energy input from solar radiation efficiently into turbulent heat and release the energy into the atmosphere (e.g. Breil et al., 2020), counteracting the increased solar radiation." (Lines 428-434)

474 "certain" -> "some"
Reply: is replaced

474 "reached as high" -> "reached values up to"

Reply: is replaced.

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.