Response to reviewer 4 of the paper entitled
"Different response of surface temperature and air temperature to deforestation in climate models"
Ref.: esd-2018-66

We would thank the reviewer for the time he/she devoted on reviewing the manuscript, and for his/her helpful comments.

Below are the reviewers comments (***bold italic font***) and our responses to each point (normal font). All line numbers that we provide in our responses refer to the revised version of the manuscript in which track changes are not shown.
The original manuscript contained one paper (Winckler et al., 2018) that had not been accepted yet. This manuscript has now been accepted and can be made available to the reviewers.

This paper 'Different response of surface temperature and air temperature to deforestation in climate models' by Johannes Winckler investigated the discrepancy in the temperature response to deforestation between climate model and observations, and how the deforestation impact differs among temperature variables. The question studied here is important to understand the impact of deforestation on temperature. The paper also presents some interesting new findings on this topic. Therefore, I think the paper is suitable for publication in Earth System Dynamics.

We are happy that the reviewer is interested in the topic and finds the paper suitable for publication in Earth System Dynamics.

1. ***Major comments:***
   a) I feel the manuscript needs to be edited to improve the language, especially for the use of preposition like 'the', and some sentences are difficult to understand.
   b) According to results of this study, is it possible to establish a relationship to link the impact on surface temperature and on near-surface air temperature to reconcile their differences (a statistical model or the ratio 0.5 found in the paper)?
   c) When analyzing the discrepancy, model uncertainty should be always kept in mind. How the results of this study would be affected by such uncertainty?

   a) The language was improved, we hope that the sentences are now easier to understand.
   b) We think that it's a good idea by the reviewer to develop a statistical model to derive the T2m response from the Tsurf response. However, this is a non-trivial task as this ratio may vary by location and season (e.g. Figs. S7 and S8; Fig. 3 is only for DJF/JJA in the northern mid-latitudes) and goes beyond the scope of our study. Although the models seem to agree that the ratio of dT2m:dTsurf is around 1:2 over the studied region and the considered seasons, the exact ratio between the response of T2m and Tsurf is still to some extent model specific (range for JJA: 0.35-0.66 excl. HadGEM2-ES, see Table S1).
   c) The reviewer is right that it is important to be aware of model uncertainties. In the last paragraph of section 3.2, we argue that the inter-model differences are large for dTsurf, but smaller for the ratio between dTsurf and dT2m.

2. ***P2 L30: It would be better to also provide the submitted manuscript (Winckler et al., 2018) to reviewers to facilitate the review.***

   We are sorry about the inconvenience in the first phase of the review process. The manuscript

(Winckler et al., 2018) is now published (doi: 10.1029/2018gl080211) and can be made available to the reviewers.

3. ***P2 L31: I think these studies compared nearby locations between forest and non-forest or between locations with and without deforestation.***

   We adjusted the text accordingly.

4. ***P2 L35: Please specify the different mechanisms here.***

   We now specify one sentence later that the local effects act predominantly via changes in turbulent heat fluxes, while the nonlocal effects act predominantly via changes in incoming radiation that reaches the surface (Winckler et al., 2018).

5. ***P3 L6-8 This sentence needs to be revised for clarity.***

   We removed this sentence which was obviously confusing and did not add much value.

6. ***P3 L16: add '. . . from CMIP5.'***

   We adjusted this sentence and included the 'CMIP5'.

7. ***P4: 'The local effects are thus the temperature changes that exceed the nonlocal temperature changes that are obtained by interpolation from nearby non-deforested grid boxes'. I don't understand this sentence.***

   We removed this sentence because it was obviously not clear, and it was anyway only meant to be a summary of what was explained above.

8. ***P4 L2.3 Better specify 'CMIP5' models***

   The title of section 2.3 is now 'Isolation of local effects across CMIP5 models'.

9. ***P5 L9: How about the 2m temperature in other models, is it defined in a similar way and thus have the similar problem?***
   As for 2m temperature from observation, is it the 2m above ground (within canopy), or 2m above canopy?

   We think that it is reasonable for climate models to use semi-empirical formulas based on Monin-Obukhov similarity theory (see last paragraph of section 3.2), and thus we expect that also in other models temperature is defined 2m above $d + z_0$ rather than 2m above the surface or canopy.
   Concerning the observations, we state in the discussions section (around p. 13, l. 19) that weather stations (i.e. in forest clearings) record temperatures at a height of between 1.2m and 2.0m above ground level while temperature at Fluxnet sites is typically recorded 2-15m above forest canopies.

10. ***P5 L31: Why only 30 years for the non-local effect? I realized that this is explained later. Maybe some rearrangements can be done for this.***

    Section 2.3 is now re-arranged such that first the '30 years' are introduced.

11. ***Figure 1: Since the transition latitude from warming to cooling is discussed in the paper, it would be useful to have a latitudinal averaged temperature response for different temperature variables (or in a separate figure).***

    We now provide zonal land averages of the responses of T2m and Tsurf in Fig. S2.

12. ***P8 L30-35. If the discrepancy is explained this way by Richardson number, it sounds like such discrepancy is a model-dependent artifact instead of actual phenomenon. The discrepancy can be seen in observations (e.g., Baldocchi 2013), suggesting it is not just the Richardson number reason. I guess that the differences in the magnitude of Tmin/Tmax and seasonal responses could play a role because they cancel out each other at the annual mean scale.***

    The reviewer is right that the magnitude of the Tmin/Tmax and seasonal responses are important to explain the annual mean response of Tsurf and T2m, we added this as the last sentence in section 3.1. However, we think that differences between Tmin/Tmax and between seasonal responses are not the only reason why Tsurf and T2m can respond differently; e.g., even for Tmax in JJA, in some regions T2m and Tsurf can show a different response (Fig. S9.)
    We re-wrote the last two paragraphs of section 3.1 (for Tmax and analogously for Tmin) to clarify that there is a plausible mechanism why T2m and Tsurf could respond differently in reality, and how this mechanism is implicitly accounted for in the calculation of T2m in the MPI-ESM.

13. ***P10: L11-13: I don't understand this sentence.***
    P10 L15: 'all but one model show a surface warming locally' this sentence may cause confusion.

    Both sentences were rewritten, we hope they are now more clear.

14. ***P11 L13-14. The 0.5 ratio is an interesting number. Is it applicable to section 3.1?***

    As can be seen in Fig. S2, this ratio varies with latitude, even when focusing on annual means. It seems plausible that this ratio may vary also with the considered season.

15. ***P12 L23: With the scale of deforestation in reality much smaller than the model simulation, the non-local effect is negligible and the local effect is dominant, this makes the climate model and observation more comparable.***

    The reviewer is right that the nonlocal effects in reality are much smaller than in our simulation '3/4'. We now clarify in the caption of Fig. 1 that the shown results refer to this simulation. Furthermore, we added in the text that the nonlocal effects are expected to be large especially in simulations of large-scale deforestation. This does not alter our statement that including the nonlocal effects causes an inconsistency in comparing the models and observations.

16. ***P13 L9-10: There is a possibility that this is due to in climate model uncertainty, we don't know if the model is able to perfectly simulate Tmax response. Model uncertainty needs to be taken into account when making this statement.***

    We replaced this sentence by the following:
    'Our results for the MPI-ESM suggest that the difference between T_{2m} and T_{surf} is particularly strong for mean daily maximum temperature (see Fig. 2). Further studies may investigate whether this is also true for other climate models and observation-based data-sets.'

17. ***A recent paper by Melo-Aguilar (2018) might be helpful.***

    Thanks! The two suggested references are now included in the introduction.

    Reference:
    Baldocchi D, Ma S. How will land use affect air temperature in the surface boundary layer? Lessons learned from a comparative study on the energy balance of an oak savanna and annual grassland. Tellus B. 2013
    Melo-Aguilar C, Gonzalez-Rouco JF, Garcia-Bustamante E, Navarro-Montesinos J, Steinert N. Influence of radiative forcing factors on ground – air temperature coupling during the last millennium: implications for borehole climatology. Clim Past. 2018;1583–606.

References

Winckler, J., Reick, C. H., Lejeune, Q., and Pongratz, J. (2018). Nonlocal effects dominate the global mean surface temperature response to the biogeophysical effects of deforestation. *Geophysical Research Letters*.