

The reviewers' comments are highlighted in blue, the responses are in red.

Response to Review 1

This study proposes a new methodology to separate local and non-local biogeophysical effects of deforestation in model experiments. This new methodology consists of calculating a linear regression between moving 5x5 model grids in the control simulation and the same 5x5 model grids in the deforestation simulation using tree cover, latitude, longitude, and elevation as predictor variables. The slope of the tree cover predictor variable is deemed to represent the local sensitivity of the dependent variable of interest (i.e. surface temperature, albedo, or latent heat flux). Using this new methodology on CMIP6 deforestation model experiments, this study derives emerging relationships between local and total surface temperature changes, albedo changes and latent heat flux changes. It also compares local surface temperature changes, albedo changes, and latent heat flux changes between CMIP6 model experiments and observational data from FLUXNET and MODIS to constrain model results. This study finds that the local surface temperature derived from model experiments tends to be lower than the temperature derived from observations. The albedo response is also stronger in most models compared to observations, particularly in areas with high snow cover.

This is an interesting study that provides an alternative way to calculate local effects from model experiments when it is not possible to use the checkerboard method. This provides us with a deeper understanding of the differences in local effects between models and observations. However, I believe the local effects calculated from this new methodology are contaminated by non-local effects. The extent of which is hard to determine. The magnitude of contamination would depend on the range of forest cover change among pixels included in the regression window, with a greater range likely leading to less contamination and a smaller range likely leading to more contamination. As an extreme example, if the forest cover loss is around 50% for all pixels in the window, and all pixels have a very similar surface temperature change of 1 degree between the control and the deforestation simulation, the resulting slope of the tree cover variable will be around 2. In this case, the bulk of the change in temperature would be attributed to local effects. However, this change in temperature of 1 degree encompasses both local effects and non-local effects coming from pixels outside the window. The contamination of local effects by non-local effects (which are generally a cooling) is likely one of the factors why the local temperature effects calculated by the models are significantly below observations in many cases. This should at the very least be discussed more thoroughly in the limitations and mentioned as an additional plausible cause of differences between observations and models when discussing results.

Response to Reviewer 1 main comment about separation of local and non-local effects, and potential for underestimating the local effect of deforestation on near-surface climate:

We thank the reviewer for his/her comments and appreciate he/she finds our study interesting. We'd like to point out that the moving window regression approach, which our study builds on and develops further, has been applied and evaluated by several previous studies. A very recent study (Jaeger et al., 2025, 10.1029/2024JD042698) proposed an alternative third approach to moving window regression and the chessboard approach and confirms that both are generally valid. However, there has been some evidence that the regression approach underestimates the

local effects to some extent, which, in our original submission, we acknowledge in the Limitations section (Section 5). Previous studies using a similar approach based on linear regression indeed pointed this out, when comparing this method to the factorial experiment approach (Lejeune et al., 2018) and the look-up table approach (Winckler et al., 2019 (Fig. S8)). In order to quantify the discrepancies between the different methods for separating local and non-local effects directly for our simulations, we have performed an additional analysis (see Fig. A6 in revised manuscript) comparing the temperature sensitivities to deforestation in MPI-ESM as calculated in our study based on the *deforest-glob* simulation (via linear regression) and in the study of Winckler et al. (2019), where deforestation is performed in every one or two out of four grid cells (chessboard pattern). While the deforestation patterns in these simulations are different, the magnitude of deforestation is similar and therefore it is possible to approximately compare the two methods – the one based on linear regression and the chessboard pattern method (Fig. A6). This analysis showed that the differences in temperature sensitivities between the two methods vary in geographic space and, importantly, show no systematic bias towards underestimation of the local effect (Fig. A6).

It should be kept in mind that this comparison was done only with one model (due to data availability, this comparison was possible only with the MPI model) and that the chessboard pattern method comes with its own set of uncertainties (Winckler et al. 2019). Importantly, while we cannot exclude contamination of the local effects by non-local cooling, these results suggest no systematic underestimation of the local effect as suggested by previous studies (Lejeune et al., 2018; Winckler et al. 2019). Ultimately, modellers should view our results as a first evaluation and are encouraged to perform offline simulations to further test their model's skill in replicating observationally based temperature (plus albedo- and latent heat-) sensitivities to deforestation.

Further, in this comparison (Fig. A6), we did not apply the scaling of the temperature sensitivities to 100% deforestation (see Section 2.5 in revised manuscript, lines 218-225), in order to keep the magnitude of deforestation similar to the one in the experiments of Winckler et al. (2019). Applying the scaling does lead to amplification of the differences in temperature sensitivities (resulting from the two different methods for extracting the local effects), which are more notable over the boreal region in Eurasia and the central parts of the Amazon and Congo basin (Fig. A6). However, since the same scaling approach has been applied to the observational datasets, a comparison between models and observations is still meaningful and the validity of the emergent constraints is not affected.

To respond to the specific example given by the Reviewer, where the forest cover loss for all pixels within the moving window is 50%, we would like to point out that, indeed, this pattern of deforestation has prevented us from applying the classical linear regression approach by Lejeune et al. (2018). To avoid the issue, in our proposed methodology we have taken (instead of change) the values of temperature (T) and tree cover (F), which are different in the *deforest-glob* and *piControl* simulations. And, then we train a linear regression on the pairs of T and F from both simulations.

To reflect on this Reviewer's main comment and our additional analyses, we have added the following text in the revised manuscript in the first paragraph under Limitations (Section 5):

“An additional analysis comparing temperature sensitivities to deforestation based on the linear regression method applied in this study and the chessboard pattern deforestation experiments of Winckler et al. (2019a) did not reveal evidence of systematic underestimation of local effects (Fig. A6e,f). It should be noted that in our study the local effects represent responses due to

incomplete deforestation as defined in the deforest-glob experiment. An additional scaling to 100% was applied in order to make these local effects comparable to observations, which only capture complete deforestation. Thus, it is expected that the modeled effects without the scaling applied (Fig. 9) are attenuated in comparison to the observational datasets. When evaluating non-local effects, scaling to 100% deforestation is not appropriate as non-local effects cannot be directly attributed to the percentage of tree cover, although some authors suggest a linear relationship between non-local effects and the number of deforested grid cells (Winckler et al., 2019a). However, scaling to 100% could lead to overestimation of local effects compared to other methods (Fig. A6g,h) possibly because the relationship between surface temperature and tree cover change is not strictly linear as albedo and evapotranspiration effects vary as a function of initial tree cover and across biomes (Bonan, 2008; Alibakhshi et al., 2020)."

Additionally, although the manuscript is generally well-written, it could benefit from some reorganization to improve the flow, some clarification of concepts in certain areas, and consistency in the terms used to avoid confusion, as indicated in the specific comments. Moreover, the section on non-local effects is not directly relevant to the research objectives (comparing the local effects across a range of ESMs and against observations) and could be moved to the supplementary material. The effectiveness of the figures could also be improved, mainly by making the differences between observations and model experiments stand out more, as per the specific comments.

Regarding the section on non-local effects, we have elaborated in the specific comments on the importance of considering the non-local effects. In this section, we show a way forward how to include in the comparison models that overestimate the albedo-induced cooling.

Specific comments:

14: Typo, northern latitudes

Lines 14: fixed in the revised manuscript

15: Unclear what emergent constraint means without having read the paper, suggest to remove "via an emergent constraint" in the abstract, or elaborate further on the concept of emerging constraints

15: suggested improvement: "[...] via an emergent constraint (i.e. resulting from the linear relationships within the model ensemble)"

55-57: Large-scale deforestation also triggers changes in ocean circulation which would impact non-local effects (Portmann et al., 2022)

55-57: suggested improvement: "But large-scale deforestation in these experiments also triggers strong changes in advection of heat and moisture, as well as in atmospheric and ocean circulation, which influence regions that have not undergone deforestation (Winckler et al., 2017; Portmann et al., 2022)."

75: Typo, cloud cover instead of forest cover

75: to be removed to avoid confusion: "[...] temperature changes ~~resulting from changes in forest cover~~"

91-92: Define what observationally based emergent constraints are. Needs more background. See my comments for line 215-230. It is hard to understand what it means and if it is any different than observational constraints or simply emergent constraints?

91-92: to be removed to avoid confusion: “The objective is to provide ~~observationally based~~ emergent constraints...”; clarification of “emergent constraint” added (see 15)

164-166: Add how the change in forest fraction is represented when using the FLUXNET data to make it consistent with the ESM experiment and the MODIS-based data explanations

164-166: suggested improvement: “The change in forest fraction is 100% and is considered only for pixels, where the respective vegetation cover types are actually present and/or could potentially occur as defined by MODIS-derived land cover maps for 2005 and Köppen-Geiger climate zone maps for the 20th century.”

185: Add comma after the word pair

185: fixed in the revised manuscript

186: You need pixels with greater than 10% forest cover change to calculate the local effects, but pixels with zero or minimal forest cover change can help strip out the non-local effects (as changes in those would be attributed mainly to non-local effects). You could therefore calculate the linear regression *for all pixels*, but *only* in windows where there is a sufficient number of pixels with greater than 10% forest cover change. It would be interesting to see how this changes the magnitude of the local effects.

186: It is not expected that including pixels with less than 10% forest cover change will alter the results substantially, as in most models the forest cover change is much larger than 10% (see revised manuscript, Fig. 1). We have provided an additional analysis comparing the surface temperature sensitivity with and without the 10% threshold for the MPI model (see Appendix at the end of this document, Fig. 1). The MPI model was selected based on the availability of more pixels with less than 10% forest cover change in comparison to other models. The results with using the threshold are more robust, especially for the boreal forest, as evidenced by the number of statistically significant pixels.

192: Consider adding that you then calculate the mean value of the local sensitivity over the period and region of interest for completeness.

192: The information is already (partially) provided in 211-214. Suggested improvement: “In order to evaluate the consistency of seasonal surface temperature responses between the ESMs and the observation-based estimates, we extracted the mean values of local surface temperature responses to deforestation also at monthly timescale for broad latitudinal regions: boreal (from 50° N to 90° N), temperate (from 23° N to 50° N) and tropical (from 23° S to 23° N).”

203: Consider starting the new subsection here “2.5 Comparing local sensitivities between observations and ESMs – emergent constraints”

203: Subsection 2.5. updated accordingly

215-230: Most of this section should go in the Introduction as it provides background info on the emergent constraint approach and how it can be used, and discusses goals of the study.

215-230: Explanation of the emergent constraint concept is moved to Introduction.

223-230: Hard to understand, would benefit from a clearer description of what you are trying to achieve: are you saying that if we know the local effects from observations, and the emerging relationship between local effects and total effects from model experiments, we could use that relationship to constrain the total effects based on the observed local effects? Regarding albedo, if we know the albedo change from observation, and the emerging relationship between albedo and total temperature from model experiments, we could use that relationship to constrain the total temperature change based on the observed albedo change?

223-230: suggested additions/changes: “By knowing the local effects from observations, and the emerging relationship between local and total effects from model experiments, we could use that relationship to constrain the total effects based on the observed local effects. Thus, by using the local sensitivities derived from in-situ and satellite-based data, we are potentially able to provide emergent constraints for surface temperature, albedo and latent heat flux. In addition, we use the relationship between local surface temperature and albedo, and between local surface temperature and latent heat flux, to explore further the emergent constraints.”

227: I don't think we can say that local temperature sensitivities are comparable between models and observations in boreal and temperate regions based on your data (although they appear to be comparable in the tropics). Most models show local temperature effects significantly lower than observations, which, as discussed earlier, could be due to the incorporation of the non-local cooling effects in the local effects calculated from models.

227: The possibility to compare local effects in models with observations has been confirmed by a number of studies (e.g., Bright et al., 2019; Winckler et al., 2019a; Chen&Dirmeyer, 2020). The issue with the possible overestimation of local effects is discussed in the first paragraph. To avoid a misunderstanding about the term “comparable”, we have changed the sentence from “[...] as local surface temperature sensitivities from models and observations are comparable” to “[...] as local surface temperature sensitivities from models and observations can be brought into a meaningful relationship”.

229: Local surface temperature or total surface temperature?

229: Here we refer to the total effects.

249-253: Sentence starting with “It has to be noted...” to sentence ending by Fig A4: could go in the limitations section of the Discussion.

249-253: Moved to Limitations in the revised manuscript.

278: Clarify local or total surface temperature.

278: Here we refer to local surface temperature.

280: The emerging linear relationship between local temperature and total temperature is a key result, I think the figure showing this result (currently A1) should be incorporated in the main text and not the supplementary information.

280: Fig. A1 is moved to the main text in the revised manuscript (Fig. 8).

288: Add “defined by the standard deviation of the *MODIS* observations” for clarity.

288: Changed to “defined by the standard deviation of the *MODIS* observations”

304: Section 3.4 should really be part of section 3.3 and not a separate section.

304: Section 3.4 is merged with Section 3.3

316: This section on the non-local effects doesn't add much information that is directly related to the objectives of the study and can deter from the core message. I would move Figure 7 and corresponding information to the supplementary material.

316: While in this section we report also on non-local effects, which are not the primary focus of the study, we also evaluate the local effects during summer. This evaluation is particularly important because by considering only summer months we are able to exclude the effects of the Siberian cold bias still observed in many CMIP6 models (Portal et al., 2023). It is also relevant in light of the discussion on the contamination of local effects by non-local effects. In the summer, the non-local cooling is weaker (due to the absence of the snow albedo effect), thus giving possibly clearer view on the local effects.

337: Observational constraints, observational emergent constraints, or emergent constraints? Is there a difference between the three?

337: We refer to observational constraints as a broader term, encompassing emergent constraints. When non-local or total effects are referred to, we use specifically the term (observational) emergent constraints, as these effects can be evaluated only by considering the emergent linear relationship between the different models. "Emergent constraints" and "observational emergent constraints" are synonymous. Clarification is added under Section 2.5.

337: This section would also benefit from a discussion on how we could use the emergent linear relationship between local temperature and total temperature changes to constrain total temperature changes from ESMs based on observed local temperature changes and the limitations/pitfalls of doing so.

337: suggested addition at line 360: "Because of the emergent constraint relationship between the local and total surface temperature effects, this overestimation is valid also for the overall response to deforestation, thus showing that most models exhibit too strong cooling in comparison to observations (Fig. 8). For the tropics, however, approximately half of the ESMs show realistic total surface temperature response, as defined by the emergent constraint based on MODIS data, with fewer models being within the realistic margins defined by the FLUXNET-based dataset (Fig. 8)."

360: Observational constraints, observational emergent constraints, or emergent constraints?

360: see above

378: Observational constraints, observational emergent constraints, or emergent constraints?

378: see above

396: This section could be removed, since it does not add much new information compared to other studies, and if the results regarding the non-local effects are moved to supplementary material as per prior comment.

396: see comment to 316

412-496: This is a major limitation as discussed in my general comments. Indication that non-local effects contaminate local effects should be mentioned along with its repercussions, including overestimation of the local cooling effect.

412-496: see general comments

Fig 3: Replace Bright et al. (2017) by FLUXNET in the legend to make it clearer where this data come from when looking only at the figure. It is already indicated in the text that the FLUXNET data comes from Bright et al. (2017). Y axis: add “local” temperature change to avoid confusion.

Fig. 3: updated in revised manuscript

Fig 4: Use a different color for the MODIS data point (observation) to make it stand out more compared to model data. Y axis: add “local” temperature change to avoid confusion. Replace observational constraints with “observational emergent constraints” in the caption to make it consistent across all figures.

Fig 4: updated in revised manuscript

Fig 5: Same comments as Fig 4. Consider using the same color for the MODIS data point in plots c-d-e and for the MODIS lines in plots a-b, and same color for the CMIP6 data points in plots c-d-e and for the CMIP6 lines in plots a-b for consistency.

Fig 5: updated in revised manuscript

Fig 6: Same comments as Fig 4.

Fig 6: updated in revised manuscript

Fig. A5 and A6: Same comments as Fig 4, as applicable.

Fig. A5 and A6: updated in revised manuscript

References

Portmann, R., Beyerle, U., Davin, E., Fischer, E. M., De Hertog, S., & Schemm, S. (2022). Global forestation and deforestation affect remote climate via adjusted atmosphere and ocean circulation. *Nature Communications*, 13(1), 5569. <https://doi.org/10.1038/s41467-022-33279-9>

References

Alibakhshi, S., Naimi, B., Hovi, A., Crowther, T. W., & Rautiainen, M. (2020). Quantitative analysis of the links between forest structure and land surface albedo on a global scale. *Remote Sensing of Environment*, 246, 111854.

Bonan, G. B. (2008). Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *science*, 320(5882), 1444-1449.

Bright, R. M., Davin, E., O'Halloran, T., Pongratz, J., Zhao, K., & Cescatti, A. (2017). Local temperature response to land cover and management change driven by non-radiative processes. *Nature Climate Change*, 7(4), 296-302.

Jäger, F., Schwaab, J., Bukenberger, M., De Hertog, S. J., & Seneviratne, S. I. (2025). Spectral decomposition and signal separation of climate responses to land cover changes. *Journal of Geophysical Research: Atmospheres*, 130(5), e2024JD042698.

Lejeune, Q., Davin, E. L., Gudmundsson, L., Winckler, J., & Seneviratne, S. I. (2018). Historical deforestation locally increased the intensity of hot days in northern mid-latitudes. *Nature Climate Change*, 8(5), 386-390.

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.

Portal, A., d'Andrea, F., Davini, P., Hamouda, M. E., & Pasquero, C. (2023). Atmospheric response to cold wintertime Tibetan Plateau conditions over eastern Asia in climate models. *Weather and Climate Dynamics*, 4(3), 809-822.

Portmann, R., Beyerle, U., Davin, E., Fischer, E. M., De Hertog, S., & Schemm, S. (2022). Global forestation and deforestation affect remote climate via adjusted atmosphere and ocean circulation. *Nature Communications*, 13(1), 5569. <https://doi.org/10.1038/s41467-022-33279-9>

Response to Review 2

The work of Mileva *et al.* explores and quantifies relationships between changes to key surface energy balance variables and forest cover change in several independent observational datasets as well as in ten distinct ESMs. From this, the authors identify robust statistical relationships between surface albedo change and local temperature change, and between local temperature change and total temperature change “emerging” from the ESMs. These findings are important when viewed in the context of similar relationships seen in the observational records since this provides insights critical to the goal of model development/improvement. The paper is well-written with logical organization, and the work has been carried out thoroughly and carefully. The methods are sufficiently described and documented, and the study’s limitations are made clear.

I only have a few comments (or rather, suggestions) for improving the methodological clarity and depth of the discussion. The first is related to the methods. I think it should be made clearer up-front about the limitations of the comparisons being made between the ESM results and the observation-based datasets. “Apples-to-apples” comparisons are never really being made. Perhaps adding a table resembling the following is most efficient at clearly communicating important differences among the datasets being compared in the study.

	Timestamp	Conditional	Aggregation	Physics of Δ LST
ESMs	3-hourly	None	True monthly/seasonal/annual means	Coupled
FLUXNET-based	Monthly	None	True monthly/seasonal/annual means	Uncoupled
MODIS	Local noon (albedo); 13:30 (LST)	Clear skies	Snapshot monthly/seasonal/annual clear-sky means	Coupled

We thank the reviewer for his/her comments and appreciate he/she finds our study interesting.

Response to Reviewer 2 first main comment about improving the methodological clarity:

In order to highlight the differences between the various datasets, Table A1 was updated (see revised manuscript, Table A1), and the following table was added in the revised manuscript:

Table A2. Overview of climate variables datasets. “True” values are those that account for differences in the sun zenith angle.

	Timestamp	Temporal and spatial coverage	Temporal aggregation	Land-atmosphere interactions
ESMs	3-hourly	Complete	True 3-hourly mean	Coupled
FLUXNET-based	Monthly	Complete	True 3-hourly mean	Uncoupled
MODIS	Daily at 13:30 (surface temperature); Daily mean from the acquisitions at 10:30 and 13:30 (albedo); 8-day composite from daily acquisitions at 10:30 (latent heat flux)	Clear sky conditions only	Snapshot monthly/seasonal/annual clear-sky means	Coupled

Additionally, the following text was added to Section 2.1: “All models couple land, atmosphere and ocean in terms of momentum, matter and energy (Lawrance et al., 2016).”

Addition to section 2.3: “The differences between the various datasets in terms of temporal resolution and continuity, consideration of cloud coverage, aggregation methods and consideration of land/ocean/atmosphere interactions are summarized in Table A2.”

Regarding the discussion, I wonder if the paper would be strengthened by expanding on the application benefits of the emergent constraint found in this study as it relates to model development. For example, how, specifically, would it contribute to a “*deeper understanding of how local and non-local biogeophysical effects are represented in ESMs*” as is stated in the Abstract? The authors demonstrate important offsetting biases in a few of the models surrounding the mechanisms governing the local surface temperature response, and thus there is a risk that such an emergent constraint might not lead to any meaningful model improvement.

Response to Reviewer 2 second main comment about improving the discussion on emergent constraints:

Many processes relevant for the climate cannot be explicitly represented by mathematical expressions, which require the parameterization of climate models. The development of parameterizations and the final parameter adjustment (or ‘model tuning’) rely heavily on observations and are two areas where the availability of emergent constraints could lead to new developments and improvements. Suggested addition to the revised manuscript (at lines 496-499): “As models are usually evaluated based on how well they reproduce a subset of past observations, their ability to predict future climate is more uncertain (Flato et al., 2013). By using emergent constraints, modelling centres are potentially able to improve the parameterization and tuning of ESMs, so that they are better adapted to simulate future climate without being overfitted to historical data.”

Additionally, I feel it could also be beneficial to add some discussion surrounding the robustness of the finding of the linear relationship between the local and the total surface temperature change due to deforestation, and whether the authors think this relationship would hold in the case of more realistic (or real world) patterns and scales of deforestation.

Response to Reviewer 2 third main comment about separation of local and non-local effects:

See detailed responses to Reviewer 1 main comment.

Further, the magnitude of deforestation in the *deforest-glob* simulation amounts to 20 million km² and is therefore similar to historical deforestation from 800 to 2015 (~22 million km²) (Boysen et al., 2020). The overestimation of albedo sensitivities over snow and the difficulties of models to represent turbulent heat fluxes, as found in our study, have been documented also in the study of Luo et al. (2023), which is based on historical land use and land cover changes and therefore represents more realistic patterns of deforestation. Taking into account also the similar local surface temperature response that can be observed in our study and the study of Luo et al. (2023) and the similar magnitude of deforestation, one can postulate that the surface temperature emergent constraint would hold true also under more realistic conditions. Testing the emergent constraint in different experiments and multi-model ensembles is an important next step towards confirming its robustness. To reflect on this comment, the following text was added to the revised manuscript: “Lastly, testing the emergent constraints in different experiments and multi-model ensembles is an important next step towards confirming its robustness. The overestimation of albedo sensitivities over snow and the difficulties of models to represent turbulent heat fluxes, as found in our study, have been documented also in the study of Luo et al. (2023), which is based on historical land use and land cover changes and thus represents more realistic patterns of deforestation. Therefore, one can expect that the surface temperature emergent constraint would hold true also under more realistic conditions, however, more studies applying the emergent constraint concept in land use and land cover change scenarios are needed.”

Detailed comments

P5, L142-143: Since this sentence is lumped into the paragraph describing the remote sensing datasets, clarify that “all” here includes the FLUXNET-based dataset presented in the previous paragraph.

P5, L142-143: This sentence refers only to the MODIS products. Suggested change: “All MODIS-based datasets were reprojected to the WGS84 coordinate system and resampled to 0.05°.”

P6, L147: “using an averaging filter”. What is this (or how is it different than averaging)?

P6, L147: To avoid confusion, “filter” is removed: “[...] the pixels at the original resolution of 30 m were resampled to 0.05° using averaging”

Section 3.1 and Figure 2: The authors acknowledge the clear-sky bias limitations of observations based on satellite remote sensing (P3, L73), so I don’t understand why the models’ all-sky responses are being compared to the MODIS-based response? Wouldn’t it make more sense to compare to the FLUXNET-based dataset here which the authors acknowledge (P3, L75) overcomes this important limitation?

Section 3.1 and Figure 2: The FLUXNET-based dataset only provides changes in surface temperature but not in albedo and latent heat flux, which are available from MODIS observations

and are essential for explaining the temperature response. Therefore, it was important to include the MODIS datasets from the beginning of the analysis. Here, we also wanted to present an updated version of local surface temperature effects as shown in the study of Alkama&Cescatti (2016), which is also based on MODIS data, by using the newest generation of these products (v061), which have undergone various calibration improvements. Detailed maps of the FLUXNET-based temperature responses are available in Bright et al. (2017).

Figure 2: I don't see any stippling on any of the figure panels.

Figure 2: All figures are to be provided at 330 dpi resolution, so that the stippling is visible (minimum required is 300 dpi).

P17, L388: I'm confused by this sentence, as these two models seem to be among those with the weakest latent heat flux sensitivity?

P17, L388: Here we refer mainly to Fig. 7a in the revised manuscript, where MPI and IPSL are among the models with the highest latent heat flux sensitivity (together with UKESM and GISS) and thus further away from the observational constraint for latent heat flux, depicted on the x-axis.

Figure A3: Consider adding the observations here as in Fig. 3 so it becomes easier for the reader to more easily benchmark the individual models.

Figure A3: updated in revised manuscript (now Fig. A2)

References

- Alibakhshi, S., Naimi, B., Hovi, A., Crowther, T. W., & Rautiainen, M. (2020). Quantitative analysis of the links between forest structure and land surface albedo on a global scale. *Remote Sensing of Environment*, 246, 111854.
- Alkama, R., & Cescatti, A. (2016). Biophysical climate impacts of recent changes in global forest cover. *Science*, 351(6273), 600-604.
- Bonan, G. B. (2008). Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *science*, 320(5882), 1444-1449.
- Boysen, L., Brovkin, V., Pongratz, J., Lawrence, D., Lawrence, P., Vuichard, N., ... & Lo, M. H. (2020). Global climate response to idealized deforestation in CMIP6 models. *Biogeosciences Discussions*, 2020, 1-35.
- Bright, R. M., Davin, E., O'Halloran, T., Pongratz, J., Zhao, K., & Cescatti, A. (2017). Local temperature response to land cover and management change driven by non-radiative processes. *Nature Climate Change*, 7(4), 296-302.
- Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., ... & Rummukainen, M. (2014). Evaluation of climate models. In *Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 741-866). Cambridge University Press.
- Luo, X., Ge, J., Guo, W., Cao, Y., Liu, Y., Chen, C., & Yang, L. (2023). An evaluation of CMIP6 models in representing the biophysical effects of deforestation with satellite-based observations. *Journal of Geophysical Research: Atmospheres*, 128(12), e2022JD038198.

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.

Response to Review 3

The study uses simulations from eleven Earth System Models (ESMs) participating in CMIP6 to investigate the sensitivities of local and non-local biogeophysical effects to idealised global deforestation. Specifically, the authors analyse how surface temperature, albedo, and latent heat flux respond locally and non-locally to forest loss. The study introduces a modified regression-based method tailored to the linear deforestation experiment setup to disentangle these effects. Furthermore, the emergent constraint framework is applied to compare the simulated biogeophysical effects with observational evidence from satellite and flux tower data. The study finds that while ESMs agree on the general spatial pattern of deforestation-induced temperature changes (cooling at high latitudes and warming in the tropics), they tend to overestimate the cooling effect.

The manuscript is well written and uses adapted methods in a novel way to provide new insights into ESM-simulated biogeophysical responses to deforestation.

General comment:

A more thorough explanation or demonstration of the approach used to separate local and non-local effects would strengthen the manuscript. It remains unclear to me how the approach achieves the full separation of local effects (within the window) and non-local effects of deforestation outside the window. How can it be ensured that deforestation outside the moving window does not influence the climate variables within it, thereby affecting the linear fit that is interpreted as purely "local"? If non-local effects are not fully excluded, couldn't it affect the validity of the emergent constraint formed between local and total temperature changes? Since the separation is a key aspect of the study, a more detailed discussion of the validity of this method would enhance the robustness of the results and better establish the developed approach.

We thank the reviewer for his/her comments and appreciate he/she finds our study interesting.

Response to Reviewer 3 main comment about separation of local and non-local effects, and potential for underestimating the local effect of deforestation on near-surface climate:

See detailed responses to Reviewer 1 main comment.

In addition, it should be noted that the key idea of the separation method(s) is that the nonlocal effects are part of the background climate that is similar in close-by grid cells. Deforestation outside the moving window undoubtedly affects the climate variable within, but not in a systematic way (e.g. as a function of tree cover change within the moving window) and therefore related effects can be excluded.

Specific comments:

Lines 130 – 133: The assumption that a grassland to DBF transition is adequate to represent all deforestation in the temperate regions seems questionable. E.g. in Europe, coniferous forests are at least if not more prevalent (Naudts et al. 2016, <https://doi.org/10.1126/science.aad7270>).

Response to specific comments of Reviewer 3:

Lines 130 – 133: The local surface temperature response of the two transitions – grassland to evergreen needleleaved forest (ENF) and grassland to deciduous broadleaved forest (DBF), is almost identical (Bright et al., 2017, Fig. 3 d,e). Choosing the ENF instead of the DBF transition will, therefore, not influence the validity of our results for the temperate region. To reflect on this comment, the following text will be added to the revised manuscript: “While broadleaved forests in Europe were replaced by coniferous forests throughout the last centuries and may not be the dominant forest type in Europe anymore (Naudts et al., 2016), the local surface temperature response of the two types of forest conversion is almost identical (Bright et al., 2017) and therefore the choice of a specific forest transition does not influence the observational constraint.”. We have also provided in the Appendix of the revised manuscript an additional plot (similar to Fig. 3b in the original manuscript), showing both transitions (Fig. A4):

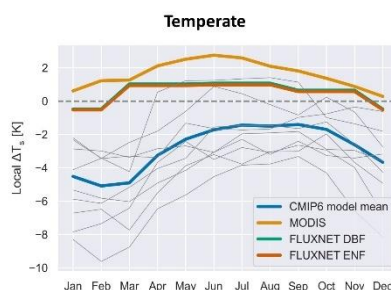


Figure 1: How is the green area defined?

Figure 1: The green area is defined by the slope of the multiple linear regression and is represented as a plane in the multidimensional space, defined by the predictor variables (tree cover, latitude, longitude and elevation).

Lines 277 – 279. I urge the authors to be more explicit about what they precisely mean by consistency. E.g. consistency in shape/behaviour or magnitude? I feel the broadly stated consistency over all hardly holds if looked at more individually. E.g. there are considerable differences in magnitude between models and observations for temperature and albedo. Further, there are quite considerable differences in shape in latent heat fluxes between ESMs.

Lines 277 – 279: The emergent constraint relies mostly on the consistent behaviour of the ESMs (e.g., decrease of temperature with increasing albedo), but also on the presence of variation/errors in the models (Williamson et al., 2021). For example, if all ESMs exhibit the same error or the spread between models is removed (e.g., due to model improvements), the emergent constraint could disappear (Hall et al., 2019). Suggested improvement: “The largely consistent behaviour...”

Lines 318 – 320. Does an approach working with two standard deviations make sense for a study of 11 models? Assuming a normal distribution, wouldn't that only lead to about half a model outside 2sd on average? It would be helpful if the authors could state explicitly how many models were excluded.

Lines 318 – 320: The main idea here was to provide a plausible ensemble mean by excluding outlier models. The excluded models are MIROC, CanESM, CNRM and UKESM (Fig. 9 in revised manuscript). Suggested improvement: “In this analysis, we concentrate on models having plausible local albedo and latent heat flux sensitivities, defined here as being within two standard deviations of the model ensemble mean, which leads to the exclusion of four out of the eleven ESMs...”

References

Alibakhshi, S., Naimi, B., Hovi, A., Crowther, T. W., & Rautiainen, M. (2020). Quantitative analysis of the links between forest structure and land surface albedo on a global scale. *Remote Sensing of Environment*, 246, 111854.

Bonan, G. B. (2008). Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *science*, 320(5882), 1444-1449.

Bright, R. M., Davin, E., O'Halloran, T., Pongratz, J., Zhao, K., & Cescatti, A. (2017). Local temperature response to land cover and management change driven by non-radiative processes. *Nature Climate Change*, 7(4), 296-302.

Hall, A., Cox, P., Huntingford, C., & Klein, S. (2019). Progressing emergent constraints on future climate change. *Nature Climate Change*, 9(4), 269-278.

Ma, H., Crowther, T. W., Mo, L., Maynard, D. S., Renner, S. S., Van den Hoogen, J., ... & Parada-Gutierrez, A. (2023). The global biogeography of tree leaf form and habit. *Nature Plants*, 9(11), 1795-1809.

Naudts, K., Chen, Y., McGrath, M. J., Ryder, J., Valade, A., Otto, J., & Luyssaert, S. (2016). Europe's forest management did not mitigate climate warming. *Science*, 351(6273), 597-600.

Williamson, M. S., Thackeray, C. W., Cox, P. M., Hall, A., Huntingford, C., & Nijssse, F. J. (2021). Emergent constraints on climate sensitivities. *Reviews of Modern Physics*, 93(2), 025004.

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.

Appendix

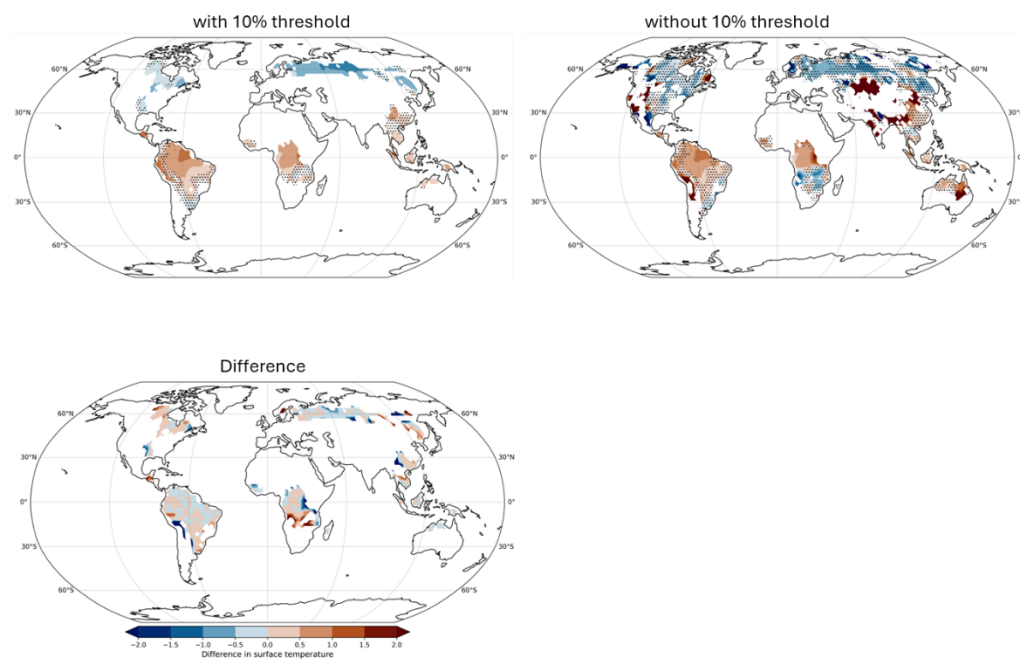


Fig. 1 Comparison between the local effects in the *deforest-glob* simulation with and without the 10% threshold being applied. All simulations are performed with MPI-ESM.