Referee 2

Review of "Effect of future forest management on carbon and energy budgets in pine forests on mineral soil in southern Finland" by Leskinen et al.

The study by Leskinen et al. examines the climate impact of different forest management scenarios in Finnish pine forests, considering carbon sequestration, surface albedo, and surface turbulent heat fluxes. The main finding is that carbon sequestration is the primary driver of the total climate impact, with lower harvest intensity resulting in the highest carbon sequestration and the most beneficial overall climate change mitigation. While changes in albedo and heat fluxes had minor counteracting effects, they only slightly reduced the climate benefits of the high-sequestration scenarios, confirming that all studied management practices maintained the forests as a net climate benefit.

I have several major criticisms, the first of which is related to the scoping. The research motives are unclear, and research hypotheses are missing. Why was this research carried out or needed? What do the authors seek to discover? How does this build on or compliment existing works? Related to this is the chosen time horizon for the analysis. No argument is made for the short (45-yr) time horizon, which is less that half the age of a typical Finnish pine rotation. If the rationale was to align with some climate policy target (which I suspect wasn't given the odd end year of the time series of 2054) then this is not communicated.
It is rather obvious that a scenario in which forests are allowed to age/grow will build up higher C stores relative to those with higher harvest rates over such a short time horizon.

Another major criticism I have surrounds the methods, specifically, the use of the CO2-equivalent metric for forcings occurring at the Earth's surface, and particularly for the turbulent heat flux changes. CO2 emissions impact global climate by altering the radiative energy balance at the top of Earth's atmosphere, so characterizing the surface forcings in terms of "CO2-equivalent" effects is flawed and misleading. This is especially the case for the turbulent heat flux changes since these are associated with energy re-distributions at the surface only, which do not directly impact Earth's radiative energy balance.

The correct and fair comparison to CO2 would be to quantify the effective radiative forcing (ERF) connected to the surface forcing, which requires looking at how the turbulent heat flux changes affect atmospheric profiles of temperature, humidity, and clouds (i.e., the radiative adjustments), and how the atmosphere attenuates the instantaneous albedo change forcing at the surface. Quantifying ERF is challenging and requires coupled simulations, which I acknowledge is probably not an option available or feasible to the authors. Thus a reasonable compromise I think would be to quantify the top-of-atmosphere (TOA) radiative forcing from the surface albedo change and discuss how the turbulent flux change results might counteract or reinforce this via their impact on the TOA radiative adjustments. Looking at Bright et al. (2025) I find some statistical models (presented in Table S2 of the Supporting Information) that could be useful here.

In general the paper is long and not very well written and organized, and includes a lot of sections/text that are not so pertinent to communicating the core methods and key results.

The Discussion reads more like an extended Results section and lacks important discussion elements (e.g., study caveats/limitations, uncertainties, climate policy relevance, avenues for future research, etc.).

We thank the referee for the detailed general assessment of the submitted manuscript. We value honest feedback, and the in-depth comments and suggestions on how to improve the quality of this manuscript, which we are motivated to implement and carry out the major revisions. The points raised in the general assessment are addressed below, after which detailed answers to the specific comments are given.

First, regarding the scoping. We acknowledge that the motivation behind this study is not very well communicated, which also limits the understanding of how this work compliments existing works. The time horizon is not based on a given climate policy target for a given year. Instead, the choice of time horizon was a compromise of several things. First, the starting point was the harvest levels of the Finnish Climate and Energy Strategy and they covered years up to 2030 only. The time horizon was extended to provide information on the longer run development. Second, in a model with a detailed description of economic sectors, such as the FinFEP model, extending time horizon very far to the future requires assumptions on the future development of technologies, product demands, the prices in the EU emission trading system, policies etc. Long time horizons would be ridden with huge uncertainties and, therefore, not that informative. Third, as the changes in harvest behavior are quickly changing the climate impacts of the forest, there is no need for long time horizons to show the impacts. We admit that a longer time horizon could have been used for illustrating the resulting long-run steady state of the impacts, but given the constantly changing technologies and wood market conditions, such a steady state would not be much more than a modeling artefact.

Secondly, regarding the usage of CO2 metric for the turbulent heat fluxes, and for the surface albedo. We agree with the referee, that the usage of CO2 metric for the turbulent heat fluxes is incorrect, as changes in these fluxes do not directly influence the top of the atmosphere (TOA) radiative forcing. This was a major oversight from the authors, and we will remove the analysis of these fluxes in carbon equivalence for the revised manuscript. The suggestion of using TOA radiative forcing from the surface albedo change, and discussing the turbulent fluxes separately, is a valid alternative approach which we will implement for the revised manuscript. We thank for the reference to *Bright et al. (2025).*

Thirdly, regarding the length, quality of writing and organization. We acknowledge that the manuscript is quite long. Limiting the analysis of the turbulent fluxes will naturally lead to the revised manuscript being shorter. We will rewrite sections of the text to improve its quality and organize some sections of the manuscript to help the flow of the paper. We appreciate the multiple good suggestions made by the referee in the more detailed comments, which can be used to improve the paper.

Finally, regarding the discussion. We do acknowledge the lacking elements (study caveats, uncertainties, climate policy relevance and avenues for future research) and will put more emphasis on the addition of study caveats, uncertainties, climate policies and future research.

Other, more detailed comments are provide below.
Focused comments

Abstract: It is way too long. Aim for 200-250 words (it is currently ~430 words).
Response:
We will make the abstract more in line with the expected length. This is an achievable goal, especially since the manuscript will be shortened after reducing the emphasis on the turbulent heat fluxes.

P2, L38-39: Forest albedos are not constant. For which timeframe do these values correspond (I presume growing season)?
Response:
This is a good request for clarification. We refer to "snow free" summer albedo, but it is not emphasized it in the current manuscript version and will be fixed for the revised manuscript.

P2, L38-49: There's a lot of detail here, but it's unclear what the takeaway message is. Given the scoping of the paper I would limit the focus of this section to the influence of forest composition and structure on albedo, since this is what is impacted by forest management.
Response:
We do agree and see that the number of details is large, and some of the details are also irrelevant when it comes to the scope of this article. We are comparing forest management scenarios with each other, not forests with other land cover types.

P2, L50-56: Why the sudden shift in focus on harvest impacts? Up until this point there has been no linking of surface biophysics with forest management. Any why the focus on hydrology only, when harvests also impact the full suite of energy budget mechanisms under scope in the paper?
Response:
We agree that the current manuscript fails to introduce forests from the perspective of forest management before L50, where we abruptly start discussing harvests and their impact on water budget. For the revised manuscript, we will add stronger emphasis to the first chapter of the introduction, that forest management practises are of keen interest in this study.

P2, L57-58: The reader learns here abruptly for the first time that this is a modeling study (it should not be assumed that the reader understands the methods/tools of the study based on information contained in the Abstract). It is better to build the case surrounding the value/need for models as a research tool in the previous paragraphs prior to stating that the present research is supported by them.
Response:
In the abstract, we had the sentence "Three different forest management scenarios with different harvest intensities were modelled with JSBACH_FOM". However, we do acknowledge that the wording makes it somewhat vague, and we will put higher emphasis on this being a modelling study earlier in the text, with clear wording.

"Climate change mitigation potential" should be clearly defined somewhere here or in the succeeding paragraph.
Response:
The term "climate change mitigation potential" will be defined clearly in the revised manuscript.

P3, L64-66: Can the authors justify a 45-yr time horizon, which is relatively short for (even-aged) boreal forestry where rotations are typically 70-120 years?
Response:
We explain this in the general response to this referee comment, and will justify it more clearly for the revised manuscript.

P3-4, Sections 2.1.1 – 2.1.3: I suggest to move many of these details surrounding the JSBACH_FOM model mechanics to the Supporting Information, and consider first introducing the harvest scenarios (i.e, move section 2.4 up to here), since these are currently referred to on multiple occasions throughout sections 2.1 – 2.3.
Response:
The suggestion to move the introduction of harvest scenarios to the beginning is very good, and it will be done in the revised manuscript. This will greatly improve the readability and flow of the text.

P4, Yasso: Missing in the Yasso-related paragraphs is detail surrounding model initialization and spin-up. What efforts were taken to ensure that the model's initial state did not unduly influence the simulation results?
Response:
We will add more in-depth description of Yasso and the model initialization and spin-up for the revised manuscript. During spin-up, soil carbon reached steady state.

P4, Section 2.1.3 "Modeling of Carbon Balance": JSBACH details are presented in terms of fluxes, whereas Yasso and C-cycle results (section 3.1) are presented as stocks/pools. Somewhere these should be reconciled, perhaps with an equation.
Response:
We attempted to explain this with the following in 2.1.3.: "*A fraction of GPP is used for vegetation maintenance respiration Rm and growth respiration Rg . Deducting Rm and Rg from GPP yields net primary productivity (NPP), describing the carbon accumulated in vegetation. Vegetation growth resulting from NPP is allocated to three carbon pools according to corresponding coefficient*". We will add explanations regarding the Yasso soil carbon pools as well. We will consider adding an equation to help the description in the revised manuscript.

P5, Section 2.2: Is the described calibration procedure documented somewhere else? If so, please state that clearly here, and if not, please elaborate on what PFT parameters have been updated or modified. Was any spin-up of JSBACH_FOM executed? A spin-up period (for soil moisture) is important and 6-12 months seems pretty standard in offline studies.
Response:
The aim was to describe and document the calibration procedure fully in sections 2.2.1. – 2.2.6 with detail. Soil parameters are described in 2.2.2 and exact properties are found from the supplementary. In 2.2.3, we state that the visible light albedo was changed from 0.04 to 0.03, from the default PFT value. Other parameters were not changed from those present in the default PFT. In 2.2.6, it is explained how the results of the calibration model run was compared with the SMEAR III observations. This calibration is referred to in 2.2.5: "Numerical values for the constants in the equations were derived based on the work of Marklund (1988) with some adjustments during model calibration". We do acknowledge that the spin-up information is limited in the current manuscript, and we will provide more details for the revised manuscript. Model spin-up period was long enough to allow carbon, temperature and moisture of the soil to reach their respective steady states.

P5, L132: Why the coarser EUR-44 and not EUR-11?
We will add a analysis of the agreement between the CORDEX data and meteorological observations for air temperature to the supplement. The distribution of 25 years of 2 meter temperature showed a good agreement between the nearest grid-point in the EUR-44 domain and the observations.

P6, L133-135: Which RCMs were used for the downscaling? How was the bias-correction Implemented?
We will add further detail about the downscaling and bias-correction. The regional climate model RCA4 was used as a downscaling model for all three global climate driver models. A distribution-based bias-correction method (SMHI-DBS45) had been applied to all datasets that we used. The reference data-set was daily E-OBS12, and the reference period was 1981-2010.

P6, L134-136: Is this somehow a rationale or justification for why the three aforementioned driving GCMs were chosen? Given your use of JSBACH as your land model, some argument for why none of the MPI-driven runs were utilized, since this is the only GCM/ESM in CMIP5 using the same land model.
We wanted to use datasets with identical downscaling and bias-correction. The three models that we used were the only ones available with the same downscaling and bias-correction methods, including both RCP4.5 and RCP8.5 climate change scenarios.

P6, Sections 2.2.2-2.2.6: Suggest to move to Supporting Info.
Response:
The suggestion has been noted and will be considered together what was said with the referee comment before this one, about: "*P5, Section 2.2: ...*" and decide if the calibration process should be entirely found from the supporting information.

P7, L171-172: At what temporal resolution were the outputs benchmarked? Was surface albedo not benchmarked?
Response:
The original data had a temporal resolution of 1 hour. Surface albedo was not benchmarked with observations. Instead, albedo for pine PFT was calibrated based on literature. We will add more details to the revised manuscript.

P8, L208-209: "RF_d_abl" should be the radiative forcing at TOA not surface.
Response:
We will implement TOA radiative forcing in the revised manuscript, instead of surface RF.

P8, L210-212: I don't follow what has been done here and why it has been done. What is the rationale of NOT using the temporally-explicit annual mean RF values here, especially since these were available from JSBACH? A very good argument for why the 45-yr average is used here instead (assumed to be repeated for all 45 time steps?), since the whole benefit of the TDEE metric is its accommodation of a temporally explicit albedo (and d_albedo RF) Scenario.
Response:
Not using temporally-explicit annual RF was firstly, because the main focus is the total climate impact across the 45-year time period. Secondly, in the climate models and scenarios, precipitation varies between the models, so the differences between for example CanESM2 and MIROC5 during exact year of 2037 would not provide meaningful information without adding more complexity to the manuscript via additional analysis, as in one model it could be a snow-heavy year, and snow-free in the other. The paper is already quite long. For the revised manuscript, we will consider adding temporally-explicit RF profiles of the different scenarios (and the differences between them) to the supplementary material, as we understand the interest towards this topic.

P11, L262-264: I had a difficult time understanding these two sentences and only did so through inference from the subsequently presented results. Consider re-writing for clarity.
Response:
We agree with the referee and the sentences will be rewritten with better clarity.

P12, Figure 5 caption: Use of the word "change" is confusing here as it often implies a difference between scenarios. I suggest something like "temporal evolution in…".
Response:
We acknowledge that the word selection here has been poor, temporal evolution would be a much better alternative.

P12, L274-278: It is unclear how the LAI results relate to or explain the C-cycle results presented in Figure 5. I'm wondering if showing the breakdown into the various pools (i.e., living biomass, deadwood, litter, soil, etc) is more helpful in interpreting the aggregate trends (of Fig. 5).
Response:
We agree with the LAI not being relevant in terms of it being a major result and will move it to supporting information. Breakdown of carbon into the various pools is something we can do and agree that it would add value in understanding the trend of total carbon pool in Fig. 5.

P13, L281-283: Again, why are all 45 years of the results time series averaged? Myself and probably many readers would find the interannual albedo dynamics relevant and interesting. You could consider aggregating months into "winter" vs. "summer", or "snow-covered" vs. "snow-free" to reduce the number of plots.
We will consider adding a time series plot, and examining winter and summer months separately for the revised manuscript.

P14, Figure 7: I feel the breakdown of albedo into its broadband constituents is unnecessary and distracting. Only the entire shortwave broadband albedo is of relevance/interest in this study (i.e, "NIR+VIS"), and most readers would probably rather see this plotted here instead for the three harvest/management scenarios, which would align with the presentation style of the previous results figures.
Response:
This was a decision we originally made because JSBACH outputs the albedo bands separately, but after reconsideration it is indeed a better option to display just the broadband albedo. We agree with the referee, that the presentation style should be more in line with the previous result figures, as it would improve the readability of the results.

P16, Figures 9-10: The snow cover results help explain the albedo results and could be moved to the Supporting Info in effort to shorten the manuscript.P17, Figure 11: Same comment as previous regarding the choice not to show the interannual dynamics. Also, consider harmonizing all difference plots so that differences are consistent across plots. In Figure 8 differences are defined as "delta = XX - EC", so doing the same for Figure 11 (instead of of "delta = EC – XX" as here) would make it easier for the reader to connect this result to the delta albedo result.
Response:
The manuscript will be shortened by decreasing the analysis of the turbulent heat fluxes. We will consider moving these to either the appendix or the supplementary, if the length of the manuscript remains long. The comment about harmonizing all the difference plots is valid, and this will be done for the revised manuscript.

P21, Section 3.3 "Water Balance": These results are not a major result and can be moved to Supporting Info.
Response:
We agree, since the water budget has received limited attention in general in the manuscript, so displaying these results here is not needed and will be moved for the revised manuscript.

Sections 3.2 – 3.4 and Figures 13-15, 17, Table 3: This content should be updated if choosing to the carry out the major revision I argue for above, with climate impacts quantified in terms of TOA ERF (albedo IRF + Adjustments). I have therefore refrained from commenting on the remaining results of these sections.
Response:
The content here will go through major revision based on the valuable feedback received.

Discussion: I also refrain from making detailed comments given the likelihood that the content will change should the authors choose to implement the suggested major revisions.
Response:
The authors will implement the suggested major revisions, so the content of discussion will have major changes in the revised manuscript.

I will though make a general comment, however, that the Discussion is missing important elements such as i) the study's limitations and/or uncertainties ii) benchmarking of main findings to those of other, similar studies (there have been a lot of similar studies published for Fennoscandian region in the past 10 years); iii) recommendations for future research.
Response:
We appreciate this general comment and agree that higher emphasis on the i) study's limitations and uncertainties ii) similar studies and iii) future research would improve the manuscript. There will be higher focus on these in the revised manuscript.

P28, Conclusions: That the EC scenario is "the most beneficial in terms of climate change mitigation" is not surprising given that differences in biogeophysical effects were negligible and since forests were allowed to grow – particularly since the analytical time horizon was constrained to 45 years.
Response:
We agree that this is not surprising, in terms of carbon sequestration. However, it was not clear how significant the lowest albedo in EC scenario is in relation to carbon sequestration in terms of the total climate effect. We will add discussion elements covering the topic that it is not surprising that the carbon sequestration was highest in the harvest scenario with lowest harvest.

And while this conclusion may be valid at the ecosystem level (i.e., when the system boundaries are limited to the forest), it may not actually hold when one takes an expanded system perspective that includes the fate of C in the wood that is harvested and exported from the forest as well as the fossil fuel C substitution occurring when the harvested wood displaces more emission-intensive non-wood products. I strongly encourage the authors to bring these elements into the discussion and recommend that they clearly state that their main conclusion is only valid at the ecosystem level.
Response:
We agree with the referee, that it is important to discuss the topic of what happens to the carbon stored in the wood after it is harvested. We will add these elements to the discussion in the revised manuscript and put strong emphasis on the main conclusion being only valid at the ecosystem level.

Bright et al. (2025), Biogeophysical radiative forcings of large-scale afforestation in Europe are highly localized and dominated by surface albedo change, GRL 52(1): [*https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2024GL112739*](https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2024GL112739)