Comment	Response
Reliance on MODIS-based products (GLASS-LAI and FLUXCOM AET) • Although you note the use of GLASS-LAI and FLUXCOM, both high-quality datasets, there is no discussion of the uncertainties associated with remotely sensed products, which I think should be acknowledged.	We thank the reviewer for this important suggestion. This referee comment aligns with the feedback from Referee #2 in [https://doi.org/10.5194/egusphere-2025-3836-RC1]. We therefore respond similar to both referee comments: GLASS-LAI and FLUXCOM-AET are both widely used and are considered high quality. In tropical regions, remote sensing can be influenced due to cloud cover. The cloud influence is particularly prominent in MODIS-derived LAI products. Therefore, we used the more robust GLASS-LAI product. The FLUXCOM product integrated eddy covariance observations, yet these are unevenly distributed across climate zones. As mentioned before, a comparison of GLASS-LAI to modeled LAI with SWAT-T can be found in Merk et al. (2024). We tested the FLUXCOM-AET accordingly before its application in this study, see figure below. Given this promising validation, the products still carry inherent uncertainties. We have added a paragraph discussing the limitations and uncertainties with GLASS-LAI and FLUXCOM-AET. We now explicitly acknowledge that these uncertainties may influence the representation of LAI and AET at the sub-basin scale. Including this discussion clarifies the scope of our validation and improves transparency regarding data limitations. The following paragraph is added to the manuscript after line 502: *The application of satellite-based GLASS-LAI and FLUXCOM-AET data for a tropical catchment like the Bétérou Catchment carries uncertainties. Generally, GLASS-LAI and FLUXCOM-AET are both widely used and are considered high quality. Yet, satellite-based datasets in tropical regions can be subject to cloud contamination and reflectance noise (Viovy et al., 1992; Atkinson et al., 2012) or the lack of observation networks for validation (Weerasinghe et al., 2020). For the Bétérou region, the GLASS-LAI dataset shows promising applicability when compared to monitored data (Merk et al., 2024). FLUXCOM-AET has been extensively validated against eddy covariance (EC) measurements across a wide range of climatic conditions, supporting its general reliability (Jung et al., 2019). This study relies on a pointwise validation of FLUXCOM-AET and GLASS-LAI. While this approach does not fully capture spatial heterogeneity or all sources of uncertainty, it provides a consistent and pragmatic basis for model evaluation in data-scarce tropical regions. This approach ensures that the validation framework remains robust and fit for purpose.*

You comment throughout that you have determined LAI is a good proxy for AET, however I find this a challenging takeaway that is not entirely proven. For example: • Currently, there is no calibration that uses Q + AET only (QA). It would be useful to either include this configuration or explain why it wasn't done and whether it would meaningfully change the results. I imagine it would change the results, as at the moment, the QL and QLA calibrations behave similarly. Could this simply be because LAI is driving the optimisation? Showing the QA calibration (or explaining its omission) would help clarify this. If QA performs similarly to QL, that would support the idea that LAI could be a proxy for AET.	We appreciate the referee's comment on the methodology and calibration approaches. The following explanation briefly summarizes the approach definition and selection. In the region, LAI and AET are strongly linked to each other as shown in literature and observational data. The overarching objective of the study is to evaluate the implications of calibration strategies to estimate AET in a sub-humid and AET-dominant catchment. In addition, the study discusses whether detailed LAI modeling can serve as proxy for predicting AET, even if AET is disregarded in the calibration. For this purpose, three calibration approaches are defined. The Q-only strategy neglects LAI and AET in the optimization and is used as a baseline to understand the AET modeling performance if it is not integrated at all. The Q-only strategy serves as a lower limit in the approach comparison. The Q+LAI strategy includes LAI, but neglects AET in the calibration. With this approach, it can be evaluated how LAI influences the AET estimation. With the application of multi-objective calibration, hydrological models are more robust. The Q+LAI+AET strategy includes all three variables in the optimization. It serves as an upper limit to evaluate best possible optimization with respect to Q, LAI, and AET. The Q+AET strategy would be a potential upper limit for the comparison just like Q+LAI+AET. It can be expected that Q+AET represents well AET rates but so does Q+LAI+AET. From Q+AET calibration, it can be expected that vegetation modeling is neglected to satisfy the water demands of AET. The extent of the study is extensive. As an upper limit in the strategy comparison, we defined a model setup which is representative for the region (Q, LAI, AET). We therefore chose the Q+LAI+AET approach rather than the Q+AET for the comparison of the strategies. Still, the discussion about the upper limit for the LAI-AET evaluation at the catchment scale is crucial. To account for this relevant point, we added a paragraph to the discussion section in the revised manuscript (line 494 and following): *The investigation of a Q+AET calibration could also serve as an upper benchmark and is expected to reproduce AET rates well. It is however not considered in the present study, because it would potentially neglect vegetation dynamics to satisfy the atmospheric water demands. Isolating Q+AET from LAI adds a further layer in understanding the role of vegetation processes in AET-focused calibrations. Looking ahead, future work could explicitly explore Q+AET as a further upper limit together with Q+LAI+AET.*

• Further, the KGE values improve for both AET and streamflow when AET is included (QLA compared to just QL). This raises the question of whether LAI alone adds enough information.	We thank the reviewer for this valuable observation. The study evaluates the impact of different calibration strategies on AET estimation and particularly discusses the role of LAI for AET modeling. AET is an optimization target in the Q+LAI+AET approach, but its optimization is neglected in Q+LAI. Therefore, better AET representation and higher KGE values can be expected for Q+LAI+AET than Q+LAI. For streamflow, the influence of LAI is less significant. The significance of parameters and processes on the streamflow modeling with SWAT-T can be evaluated through the Morris analysis (Figure 3). In Figure 3c, the sensitivity with respect to streamflow is shown where LAI parameters (green circles) are less important than groundwater (orange circles) or AET (purple circles). Therefore, LAI is less sensitive for streamflow, but it is important for AET (Figure 3d). Please note that the caption in Figure 3 will be adjusted in the revised manuscript: panel 3d is "$E_{KGE,AET}$" and 3c is "$E_{KGE,Q}$". We adjusted the manuscript for the caption of Figure 3 and added that LAI alone does not necessarily improve the streamflow modeling with SWAT-T. The manuscript is adjusted in line 389: *The analysis in Figure 5 further indicates that the integration of LAI alone does not necessarily improve the streamflow modeling with SWAT-T. As shown in Figure 3, the influence of LAI parameters on streamflow prediction is less sensitive.*
• Additionally, as noted in your study, none of the calibration strategies reach the benchmark KGE value suggested by Knoben et al. (2020). It may be worth discussing whether LAI limits performance, and whether AET alone could achieve closer to benchmark values.	We highly appreciate the referee's comment on the benchmark limits. We assume that LAI-only and AET-only approaches can increase the KGE values. However, we also expect the LAI-only and AET-only approach to not beat the seasonal benchmarks. GLASS-LAI and FLUXCOM-AET are re-analysis products based on data assimilation and are generated based on seasonal dynamics and statistical means. While they serve as a reliable reference for land-atmosphere modeling, it is challenging for process-based models to outperform those datasets. Although the integration of AET-only and LAI-only is beneficial to assess the model performance against the benchmarks, we focus the analysis on the three defined approaches to keep the manuscript concise and coherent with the climate impact assessment.
Suitability of KGE for LAI • KGE is an appropriate metric for streamflow and for AET given the daily FLUXCOM data. However, LAI changes slowly and has limited intra-seasonal variability, so I question whether KGE is the most informative metric for LAI. I	In total, three performance criteria are used to evaluate Q, LAI, and AET: KGE, PBIAS and R^2. The usage of the performance metrics is explained in Chapter 2.5, line 277. We focus the model optimization on KGE because the KGE metrics considers three statistical means simultaneously: the linear correlation between observation and simulation; the standard deviations as a measure of variability; and the mean values as a measure of bias. The application of KGE is therefore

suggest either using a LAI-specific metric (e.g., RMSE, bias, seasonal amplitude/timing) or including a justification for using KGE in the manuscript.	a comprehensive way to include different statistical means into the performance quantification. The combination of KGE with the percent bias (PBIAS, see Table A2 in appendix) and R^2 (see Table A3 in appendix) further enables a comprehensive evaluation of the LAI dynamics.
Figures 9 and 10 • These are difficult to interpret without a baseline figure showing monthly AET in this format. Consider instead presenting percentage change in monthly AET, which may communicate the intended comparison more clearly, or showing a baseline figure (even in the appendix).	Thank you for pointing out that Figures 9 and 10 are unclear to the reader. The figures are intended to show the baseline monthly AET for each strategy. The baseline for each month and strategy is indicated through the black crosses. We adjusted the caption to make it clearer that Figures 9 and 10 also contain the baseline monthly values.
Methods section • The Methods read a little cluttered in parts, distracting from the overall story. You could consider moving some elements (e.g., the sensitivity analysis description) to the supplementary materials, to tighten this.	We highly appreciate the reviewer's feedback on the phrasing and structure of the Methods section. We moved the equation parts of the description of the sensitivity analysis to the supplementary material and improved the cluttering and phrasing of the section to improve the readability for the reader.

Technical Corrections

Line 72 – "if or if not" is awkward. Suggest "whether" or "whether or not."	We highly appreciate the reviewer's valuable feedback on language, wording, and phrasing to improve the readability and overall quality of the manuscript. We adjusted the manuscript for the suggested technical corrections.
Line 140 – "We use the Penman–Monteith method..." appears to be repeated.	
Line 189 – Extra space before the bracket in "(5 km resolution)."	
Figure 3 – Consider adding detail to panel lettering (e.g., "a) all variables", "b) LAI..."). Increase font size of u* on the axes and briefly define it in the caption (e.g., "higher u* = more sensitive").	
Figure 5 – Consider relabelling the y-axis to something clearer (e.g., "Cumulative probability") or define F(x) in the caption.	
Line 398 – missing "to" after according.	
Figure 8 – Consider adding the projection period to each panel	

(e.g., "a) 2031–2050" or "a) near-future").	
Figure 11 – Axis font sizes are too small; consider increasing.	
Lines 489–490 – The sentence structure is unclear, and "with particularly for AET" is incorrect grammar. A clearer option might be: "Similar to previous comparative studies, we investigate simple to comprehensive calibration strategies, with a particular focus on AET."	

Supplemental Figures

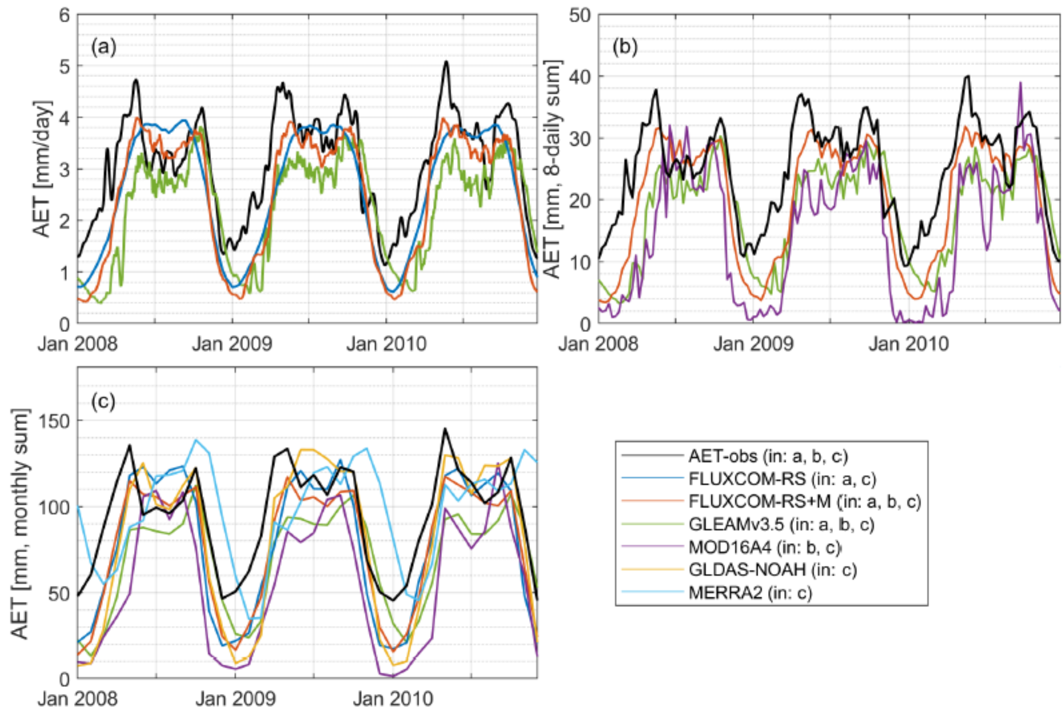

Figure 1 Comparison of monitored AET (AET-obs, black line) to multiple satellite-based and reanalysis AET products at the forested footprint for different temporal scales a) daily, b) 8-daily sum, and c) monthly sum. Please note the varying y-axis scales for readability.

Literature

Please note that literature already cited in the manuscript is not added to the following list.

Atkinson, P.M., Jeganathan, C., Dash, J., Atzberger, C., 2012. Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology. Remote Sensing of Environment 123, 400–417. doi:10.1016/j.rse.2012.04.001

Viovy, N., Arino, O., Belward, A.S., 1992. The best index slope extraction (bise): A method for reducing noise in ndvi time-series. International Journal of Remote Sensing 13, 1585–1590. doi:10.1080/01431169208904212.

Weerasinghe, I., Bastiaanssen, W., Mul, M., Jia, L., and van Griensven, A.: Can we trust remote sensing evapotranspiration products over Africa?, Hydrol. Earth Syst. Sci., 24, 1565–1586, https://doi.org/10.5194/hess-24-1565-2020, 2020