# Response to Reviewers' comments

We sincerely thank Prof. John S Kimball for his valuable and constructive comments, which have significantly contributed to improving the quality of our manuscript. We have carefully addressed all comments and will revise the manuscript accordingly. The point-to-point responses to the comments and our plans for revision are listed below.

***Replies to the General Comments:***

*1. This paper describes an update to the established P-LSH ET model, incorporating a new soil moisture constraint and model calibration to improve global performance. Model validation is assessed for global wet and dry climate zones against flux tower measurements and independent watershed level ET estimates to document relative model improvements in relation to the current model (v1) and other global ET records (GLEAM, Penman-Monteith-Leuning). Overall, the results demonstrate clear and meaningful P-LSHv2 ET performance improvement relative to ET observations and the other models. The addition of a model soil moisture constraint more strongly enhances model accuracy in dry climate zones, while providing a more realistic representation of environmental controls on ET trends. The paper is well written and comprehensive, with multiple levels of evidence to support the findings and conclusions, and well illustrated figures and tabular summaries that give the reader a clear understanding of model improvements. I therefore consider the paper to be suitable for publication once the authors address the following the following minor issues.*

**Response:**

We sincerely thank the reviewer for the positive and constructive evaluation. We have carefully addressed all issues raised and provide detailed responses below.

*2. "The authors state that tower measurements, including surface soil moisture, are used to drive and evaluate the ET algorithm (Ln 236-238). However, it's unclear whether the tower level performance assessment (Section 4.2) is based on model ET simulations derived from local tower meteorological measurements or GLDAS inputs. Additional explanation is needed here.*

**Response:**

We appreciate the reviewer's helpful comment. We will clarify the data sources used in Section 4.2. Specifically, in this section related to tower-level performance, the P-LSHv2 model was driven entirely by meteorological measurements from each flux tower, including surface soil moisture, with the only exception being NDVI, which was derived from remote sensing products. This clarification will be explicitly included in the revised manuscript as follows:

"We estimated daily ET at 106 global flux towers using the optimized P-LSHv2 algorithm, driven by tower-based measurements of radiation, meteorology, soil moisture, and remote sensing-based NDVI. The estimated ET was then compared against flux tower measurements for evaluation."

*3. The authors compare model ET performance and soil moisture sensitivity between wet and dry climate zones defined from a global climate aridity index (AI). However, the simple climate AI partitioning groups energy-limited cold land areas, including northern taiga and tundra, into the dry climate category (e.g. Fig. 2) even though these areas have generally wet soils with minimal soil moisture constraints during the short summer growing season. Thus, tundra is grouped with other GRS and OSH dominant land covers even though these other areas may represent much warmer-drier climate zones (e.g. sub-tropical Africa & western CONUS drylands). Failure to distinguish energy limited zones may contribute to the excessive model soil moisture constraint indicated in tundra (Fig. 5) and the corresponding relative ET model underestimation in this region (e.g. Fig. 14). Additional discussion is needed along these lines.*

**Response:**

We sincerely thank the reviewer for the insightful comments on the limitations of the aridity index (AI)-based dry–wet classification, particularly regarding the misclassification of energy-limited ecosystems as "dry" regions, such as the northern taiga and tundra. This concern is well founded. The AI, defined as the ratio of precipitation to potential evapotranspiration (P/PET), is a widely used metric to distinguish between water-limited and energy-limited climates. In high-latitude regions, low temperatures result in low PET, and while annual precipitation may also be low, soils tend to remain moist due to reduced evaporative demand and limited vegetation activity. Consequently, these regions can exhibit low AI values without experiencing true water scarcity. This misclassification may indeed contribute to an overestimation of soil moisture constraints and underestimation of ET in taiga and tundra regions during the growing season, as seen in our results (e.g., Fig. 5 and Fig. 14). During the short but intense growing season, soils are typically saturated or near saturation, and ET is primarily constrained by energy availability rather than soil water.

While this critique is valid, some caveats should be noted. Despite the presence of wet surface soils in summer, the tundra biome is characterized by an extremely short growing season, low radiation inputs, and permafrost-driven hydrological dynamics that distinguish it from both arid and temperate systems. Therefore, annual ET magnitude and variability in these regions are more strongly influenced by energy constraints and vegetation phenology than by seasonal water availability.

Furthermore, as shown in Fig. 13, both GLEAM and our earlier P-LSHv1 tended to overestimate ET in northern basins The revised P-LSHv2 model provides more reasonable estimates, despite showing slight underestimation in Fig. 14—an underestimation that is relative to our previous version, not to ground-truth observations.

To ensure global applicability and operational feasibility, the current version of our model uses a parameterization scheme based on vegetation type and climate zone derived from widely available static datasets. We did not incorporate phonological dynamics or vegetation growth in this version. While this simplification may introduce uncertainty in regions like the taiga and tundra, it represents a necessary trade-off between model complexity and global generalizability. Moreover, evaluating soil moisture constraints at the annual scale, rather than exclusively during the growing season, can still yield meaningful insights for ecosystems with short active periods.

The AI remains an internationally accepted and widely applied metric for global climate classification. Despite its weaker physical interpretation in cold regions, its statistical properties and global consistency make it a practical basis for large-scale comparisons. We recognize that more refined classification systems—potentially incorporating growing season length, vegetation phenology, and permafrost dynamics—could better distinguish energy-limited from water-limited systems. Nonetheless, even hybrid approaches (e.g., combining AI with temperature) may still fall short of fully capturing these complexities.

Given these challenges, we have opted not to modify the AI-based climate zonation in the current version. However, we fully acknowledge the value of more sophisticated climate classifications—particularly in cold regions—and suggest that future efforts explore integrated ecohydrological zoning approaches to better represent energy-limited ecosystems in global ET modeling. The relevant discussion will be added to the revised Discussion section to better acknowledge this limitation and guide future work. We appreciate the reviewer for highlighting this important issue, which has helped us clarify model limitations and identify future directions for improvement.

*4. Ln 42: Text should be modified similar to: MODIS data do not cover the pre-2000 period and are of insufficient length to represent longer-term interannual variability and trends, and attribution analysis in ET. The revised statement more correctly acknowledges the longer MOD16 ET record available from the NASA Terra satellite. Moreover, while the MODIS record is too short to capture climate "normals" that would require a minimum 30-year span, the data record does represent a comprehensive (500m, 8-day) multi-decadal global operational satellite ET record, which has been used to evaluate more recent interannual variability and trends (e.g. Hall et al. 2023, Roman et al. 2024).*

**Response:**

Thank you for the helpful suggestion. We agree that our previous statement did not adequately assess the time span and utility of the MOD16 ET record. We will revise the sentence to better reflect the MOD16 dataset's temporal coverage and value for assessing recent interannual variability, while also acknowledging its limitations in capturing long-term climate normal. Relevant citations (Hall et al., 2023; Roman et al., 2024) will also be added in our revised manuscript.

***Replies to the Specific Comments***

*1. Ln 14: "curcial" should be "crucial".*

**Response:**

Thank you for pointing this out. We will correct the spelling from "curcial" to “crucial”.

*2. Ln 236: Please define what is meant by “surface” soil moisture here; e.g., 0-5cm depth?*

**Response:**

In the case of GLDAS, we used the 0–10 cm soil moisture product. For the flux tower measurements, we selected the shallowest available data from each tower, typically represented by variables such as “SWC_F_MDS_1”, “SWC_PI_1”, or “SWC_1_1_1”, where the “_1” suffix denotes the top soil layer. The measurement depth varies slightly across various towers, but these variables generally represent soil moisture in the uppermost 0–5 cm or 0–10 cm layer.

*3. Ln 497: Include supporting citation on the noted net radiation decline since 2016.*

**Response:**

Thank you for the suggestion. The statement regarding the decline in net radiation since 2016 was based on a preliminary analysis of our input data. However, since we could not identify a well-established peer-reviewed reference to support this trend, we have decided to remove the description to maintain the scientific rigor of the manuscript. The sentence will be revised as follows:

“The trend of P-LSHv2 is comparable to PML (0.68 mm $yr^{-2}$), and higher than GLEAM (0.38 mm $yr^{-2}$). P-LSHv2 ET increased by 0.46 mm $yr^{-2}$ ($p < 0.001$) from 1982 to 2023, although the rate of increase appears to have slowed in recent years.”