Reviewer 1

Peng et al. have devoted considerable efforts to improve their analysis. This revision successfully addresses the major concerns I had.

Response: Thank you very much for your very positive evaluation.

Reviewer 2

Peng et al. have responded to the reviewers' comments, but there are significant concerns that remain inadequately addressed. Therefore, I recommend a MAJOR revision.

Response: Thank you for your constructive feedback.

1. **The novelty of this study lies in the PET computation, as the authors state, "This study proposes to incorporate surface vegetation characteristics, such as vegetation dynamics data, aerodynamic, and physiological parameters, into existing potential evapotranspiration (PET) methods."**

   The authors highlight the inclusion of surface vegetation characteristics—such as vegetation dynamics data, aerodynamic, and physiological parameters—into existing potential evapotranspiration (PET) methods as a novel aspect of their study. However, despite recognizing the importance of these characteristics, the study does not fully utilize the most recent datasets available. Notably, new datasets like the global canopy height dataset released approximately 2 years ago and the global 1k datasets mentioned in prior reviews offer valuable insights. The paper by Sun et al. (2023), while utilizing a different model, aims to leverage the most current global datasets possible. To align with current scientific advancements and fulfill the novelty criteria of the ESSD journal, it is imperative that the authors consider incorporating these more recent datasets into their analysis.

Response: Thanks for emphasizing the importance of utilizing the most recent datasets to enhance the novelty of our study. To advance our usage of vegetation characteristics data, we have incorporated the global canopy height dataset, mentioned in prior reviews and in the paper by Sun et al. (2023), into our analysis.

Regarding the global canopy height dataset, we have carefully reviewed the dataset and assessed its potential to improve our analysis. We first compared the newly developed 10m dataset (Lang et al., 2023) with Simard et al. (2011), which was originally used in our study. We remapped 10-m source points to target cells (0.125 degrees) by calculating their mean, with each target cell containing the mean value from all source points within it. Although the datasets are highly consistent with each other (Figure R1), some discrepancies exist for low-vegetation cases, which is expected due to the fact that Simard et al. focused on the tree height estimates.

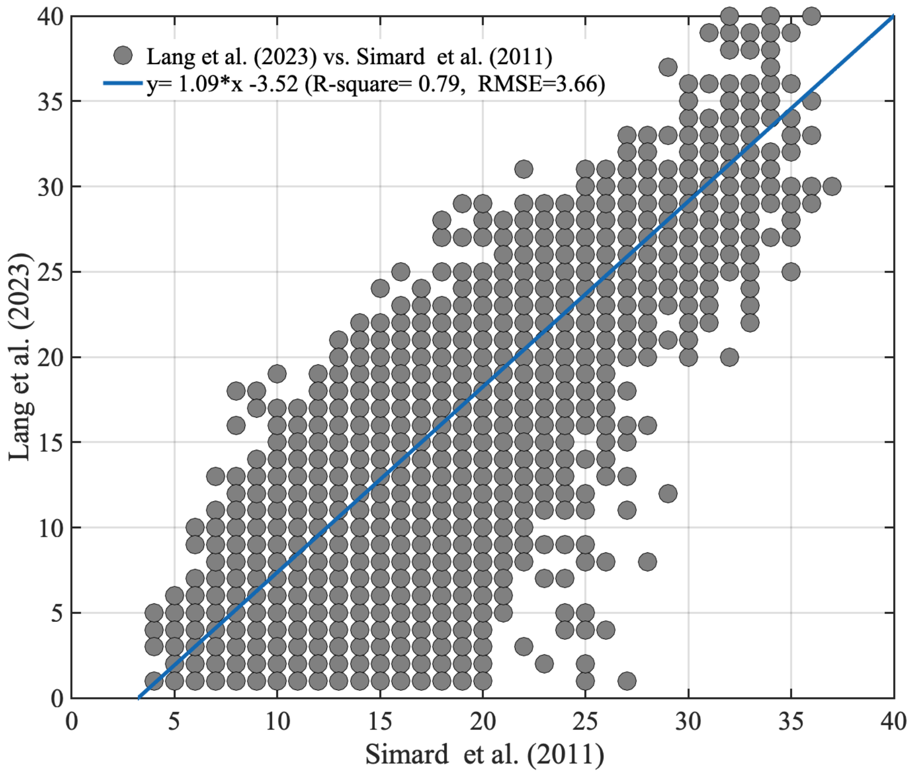

Figure R1. The comparison in canopy height between Lang et al. and Simard et al.

We then further compared the histogram of canopy height in different land cover types between Simard et al. (Figure R2) and Lang et al. (Figure R3). We confirm that the two datasets are highly consistent in the forests, while the Lang et al. provides valuable information in the short vegetation types and indeed could be utilized in our study.

Based on the distribution of canopy height in Lang et al. data for CONUS (0.125 degrees), we update the prior tree height ranges and typical values and expand to other vegetated land cover types (Table R1). We reconstruct the canopy height in each grid cell by comparing the value in Lang et al. with the ranges given the land cover type, if it is out of the range (smaller than h_{min} or greater than h_{max}) then we give the grid cell a typical value of canopy height (h_{typ}). For forests, we continue to follow the definitions from the IGBP land cover classification that the forests are more than 2 m, which we supersede the range in Lang et al. when it gave a range less than 2 m. Typical canopy height is taken from the value of the peak (mode) instead of median for forests. For DBF, Lang et al. only has 3 data points, so we use the distribution of Simard et al. instead, while keeping the lower limit of 6 m from Lang et al.. For grasslands, wetlands, and croplands, the lidar estimates from Lang et al. or Simard et al. are typically more than 3-5 meters, possibly due to the overestimation of the grid cell by the sampling of tall trees. We used conservative estimates (1.5 m for grasslands and 0.5 m for wetlands) from the literature and did not adopt the high canopy height values in these land cover types.

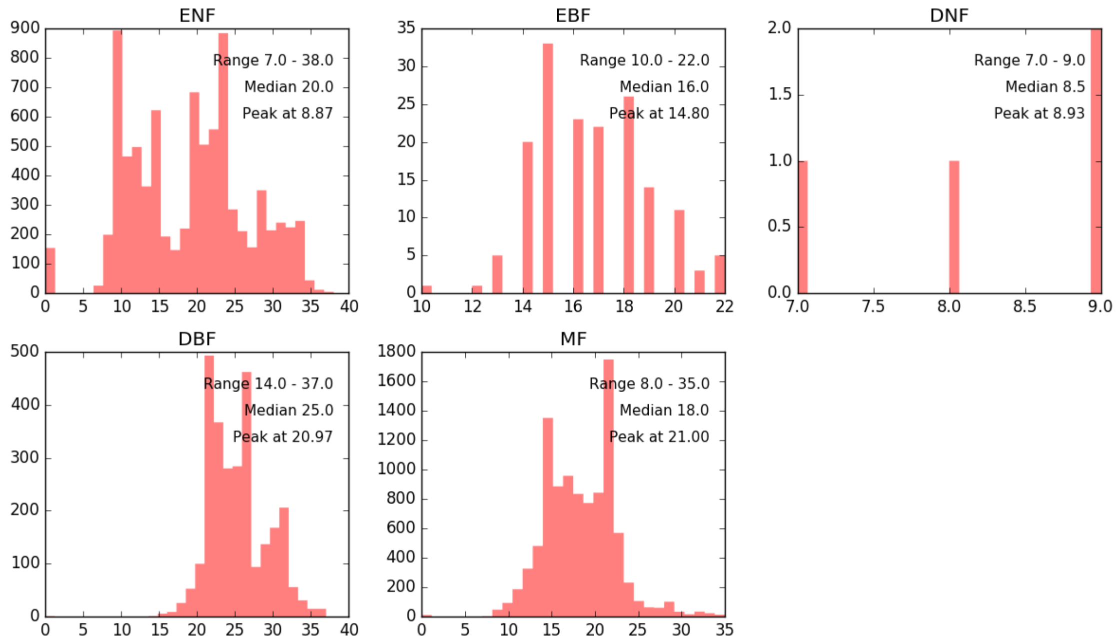

Figure R2. The histogram of tree height of Simard et al. for different forest types over the CONUS.

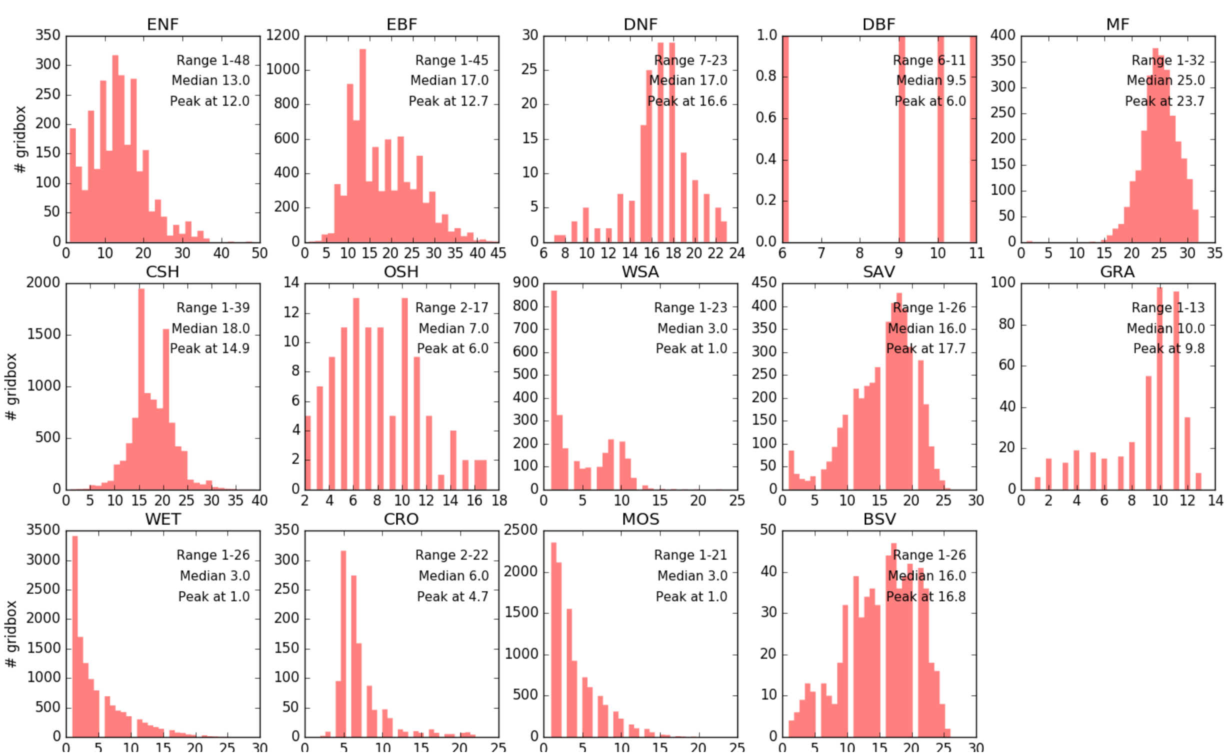

Figure R3. The histogram of canopy height of Lang et al. for different vegetated land cover types over the CONUS.

Table R1. Canopy height parameters by IGBP land cover* under previous and revised versions.

			Previous			Revised		
ID	Code	Name	h_{min} (m)	h_{max} (m)	h_{typ} (m)	h_{min} (m)	h_{max} (m)	h_{typ} (m)
0	WB	Water body	0.001	0.02	0.01	0.001	0.02	0.01
1	ENF	Evergreen needleleaf	2	50	18	2 [a]	48 [b]	13 [b]
2	EBF	Evergreen broadleaf	2	50	30	2 [a]	45 [b]	17 [b]
3	DNF	Deciduous needleleaf	2	50	15	7 [b]	23 [b]	17 [b]
4	DBF	Deciduous broadleaf	2	50	17	6 [b]	37 [c]	9.5 [c]
5	MF	Mixed forest	2	50	19	2 [a]	32 [b]	25 [b]
6	CSH	Closed shrublands	0.1	5	4	1 [b]	39 [b]	14.9 [b]
7	OSH	Open shrublands	0.1	5	2	2 [b]	17 [b]	6 [b]
8	WSA	Woody savannas	2	30	14	1 [b]	23 [b]	1 [b]
9	SAV	Savannas	2	30	8	1 [b]	26 [b]	17.7 [b]
10	GRA	Grasslands	0.1	3	0.5	0.1	3	1.5
11	WET	Permanent wetlands	0.1	5	0.5	0.1	5	0.5
12	CRO	Croplands	0.1	5	1	0.1	5	1
13	URB	Urban and built up	2	50	13	2	50	13
14	MOS	Cropland/vegetation	0.1	30	12	0.1	21	12
15	SNO	Snow/ice	0.001	0.02	0.01	0.001	0.02	0.01
16	BSV	Barren	0.01	0.1	0.05	0.01	0.1	0.05

*The above estimates are collected from [a]IGBP classification, [b]Lang et al. (2023), [c]Simard et al. (2023), otherwise based on authors' best estimates.

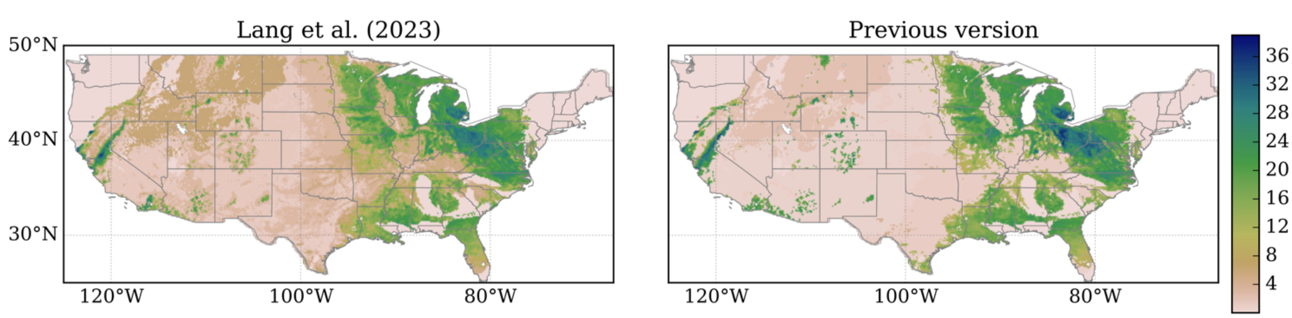

Figure R4. Left: The reconstructed canopy height based on Lang et al. and different data sources including Simard et al.; Right: The canopy height based on land cover types and Simard et al. for forests over the CONUS.

Previously we used the spatial map from Simard et al. derived from lidar data only for the forests (ENF, EBF, DNF, DBF, MF), and only used literature values for short vegetation. The reconstructed canopy height (Fig. R4 left panel) has much more spatial details, especially for non-forest grid cells, than the previous approach mostly based on land cover (Fig. R4 right panel). We admitted that this is indeed one of the limitations in our previous data source and did not clarify our processing and selection. We use the new data from Lang et al. and constrain the range using the distribution within the CONUS.

In addition to this new dataset, we add a new Canopy Height based PET approach (CH) to incorporate the dynamic canopy height dataset, rerun our models, and update our findings accordingly. We clarify the updated approach in the Data and Methods sections as below:

1) We updated the data description in L135: "This study uses the newly developed 10-m global canopy height dataset that merges the Global Ecosystem Dynamics Investigation (GEDI) space-borne LiDAR height data with Sentinel-2 satellite data (Lang et al., 2023). The original 10-m resolution was remapped to 0.125o using the average. Additionally, this study uses a global tree height dataset at 1-km for 2005 using spaceborne lidar (Simard et al., 2011) for complementary analysis in the forests (Appendix B)."
2) We added the CH approach for aerodynamic conductance in L225-240.
3) We documented in detail how we use the canopy height information in the two approaches, LC and CH, in L275: "The first method uses literature values and is adopted in the Land Cover approach (LC, Eq.9). For most of the land cover types (ID 6-16), we applied the values from the look up table except for the forests, where we determined canopy height by calculating the median height within each land cover from the tree height lidar data (Simard et al., 2011).

   The second more comprehensive method is adopted in the Canopy Height approach (CH, Eq. 11) and the SW two source model (Appendix A, Eq. A9-10). It takes into account three factors: land cover type, measured canopy height, and dynamic LAI. We overlayed the land cover map (Fig. 1) and the canopy/tree height data (Lang et al., 2023; Simard et al., 2011) to obtain the distribution in each land cover type (Appendix B). Based on the distribution of the two datasets, land cover definition, and literature ranges, we estimated the minimum canopy height (h_min) and maximum canopy height (h_max) by land cover type (Table 2). As for quality control, we set the outlier (smaller than h_min or greater than h_max) to a typical value of canopy height given land cover type (h_typ, obtained through the mode of the distribution). The actual canopy height is then determined by assuming a linear relationship with dynamic LAI following Zhou et al. (2006)."

4) We added a Table 2 for the canopy height parameters for each land cover type.
5) We added a Table 4 for the detailed formula and parameterization for each PET method.
6) We added an Appendix B to describe the evaluation and processing of canopy height.

With the updated CH approach, we re-ran our models, updated the Figures 4-7 by adding the new CH approach, and rewrote the results and findings in sections 6.1-6.3 surrounding these figures.

Figure R5 (Fig. 4 in the manuscript) shows a significant improvement in the CH method relative to the LC method in non-forested areas (Fig. R5c), highlighting the uncertainty of these parameters in the sparse vegetation and the importance of incorporating actual canopy height. We believe the newly implemented dataset will significantly strengthen the findings of our study. Thank you for your suggestion.

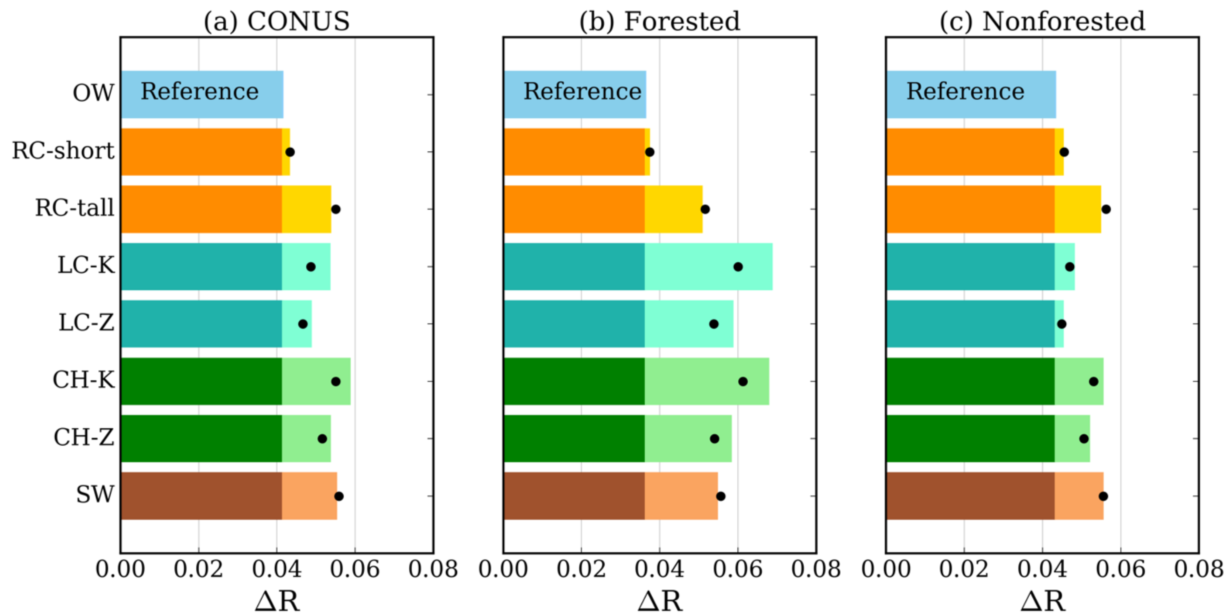

Figure R5. Differences in correlations (**Δ*R***) for selected PET methods versus the control scenario (PET = 0). Correlations were computed between the 1-month SPEI and SMsurf series across: (a) CONUS, (b) forested grids, and (c) nonforested grids.

2. **Regarding the modification, "A recent study by Sun et al. (2023) highlighted the importance of incorporating surface properties, especially vegetation control, in PET and used a two- source model designed for sparse vegetation surfaces. However, its applicability beyond sparse vegetation remains unclear, raising questions about data requirements and potential uncertainties." It's unclear why the S-W model used by Sun et al. (2023) is deemed unsuitable for areas beyond sparse vegetation, without further explanation or references.**

Response: Thanks for pointing out the lack of references that discuss the application of the S-W model across different vegetation types in this paragraph. We added the supporting literature and modified the original sentence in L75 as: "A recent study by Sun et al. (2023) highlighted the importance of incorporating surface properties especially vegetation control in PET and used a two source Shuttleworth-Wallace (SW) model designed and validated for sparse and fragmented vegetation surfaces. However, without further calibration and parameterization, the SW model's broader applicability beyond sparse vegetation is uncertain, and additionally it may increase data requirements and associated uncertainties (Gao et al., 2021; Abeysiriwardana et al., 2022)".

3. **A main focus here is on the PET products, and there are numerous available PET datasets over CONUS. However, the comparison between your products and other reference datasets is lacking. As noted in previous comments, the authors only provide a visual comparison, "The LC method not only yields modest absolute PET values (Fig.5a) but also demonstrates better performance across many regions (Fig.6). Specifically, LC estimates an annual PET of approximately 1200 mm, aligning with PET estimates for the same region and temperate zones reported in a recent study (Fig.8 in Sun et al., 2023)." A more comprehensive comparison with other reference PET datasets seems necessary. It appears that the response from the authors does not attempt to resolve the issues but rather tries to avoid directly addressing the comments.**

Response: Thanks for suggesting a more comprehensive comparison between our PET products and other reference datasets. To strengthen our analysis, we have downloaded their datasets, processed to the same spatial resolution, and compared our LC and CH methods with the proposed PET dataset by Sun et al. (2023). The comparison shows that significant spatial pattern differences can be found (Fig. R6), although the time series variations are highly consistent (Fig. R7). The magnitude of Sun et al. is closer to the PT and RC-short methods. The LC and CH methods have medium magnitudes among all the PET methods.

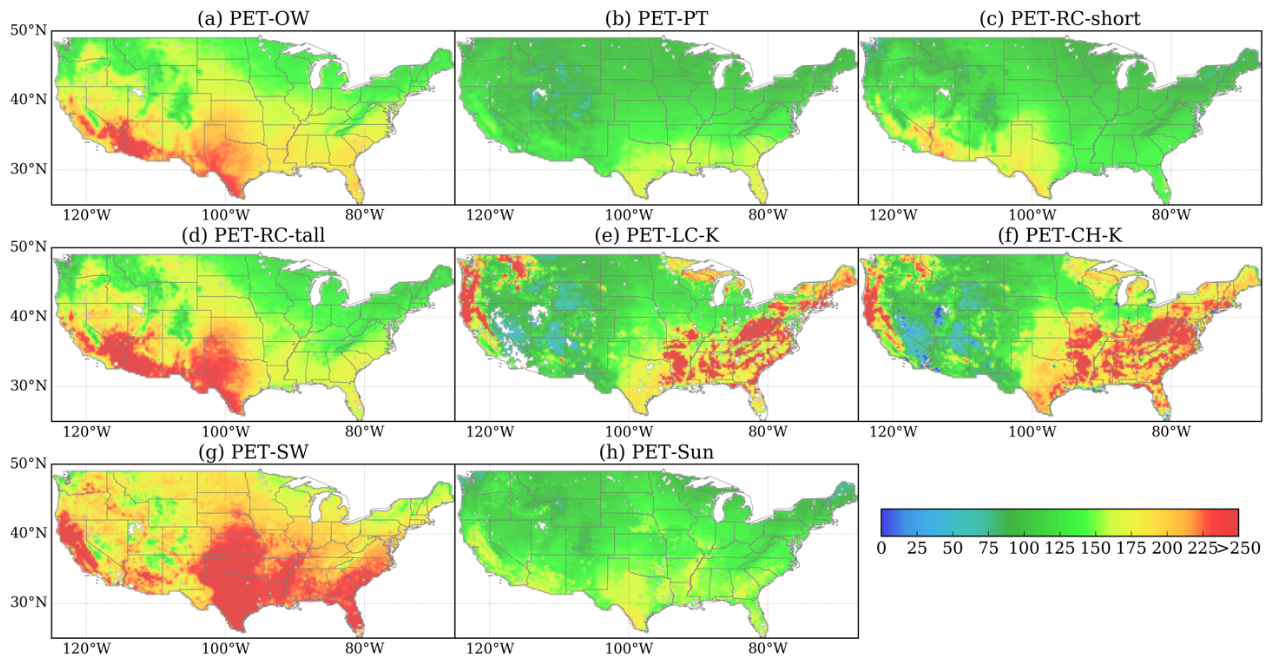

Figure R6. Growing season averages of the PET methods in this study (Table 4) and the PET dataset by Sun et al. (2023) over the CONUS.

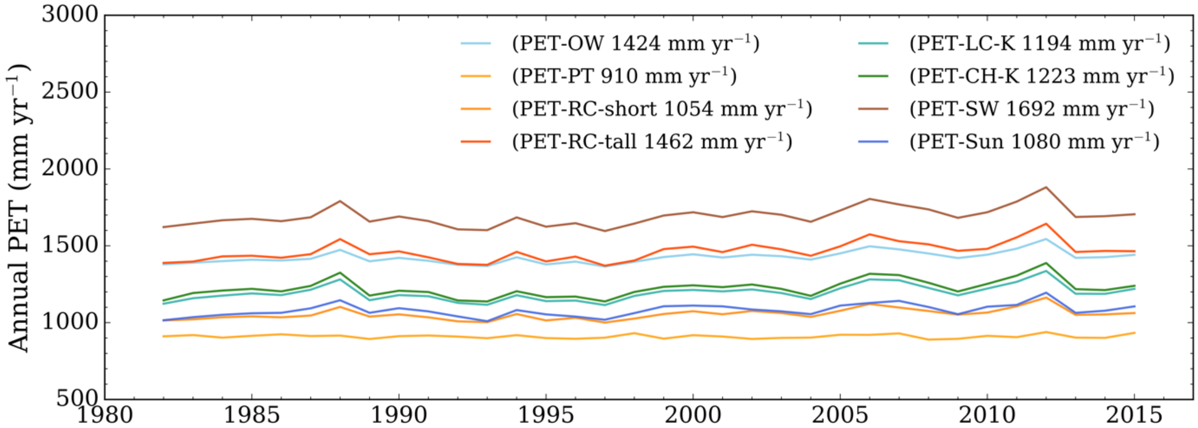

Figure R7. Annual times series of the PET methods in this study (Table 4) and the PET dataset by Sun et al. (2023) over the CONUS.

In the revised manuscript, we added Appendix C with the two figures. We also discuss them in section 5.3 Spatial patterns analysis as well as section 6.3: "Both the CH and LC methods not only provide modest absolute PET values (Fig. 5a, C2) but also display better performance across many areas (Fig. 6). Specifically, LC and CH estimate an annual PET of roughly 1200 mm, which is within the range of the higher OW value (1424 mm) and the lower values around 1100 mm from RC-short as well as Sun et al. PET dataset (Fig. C2). As Ershadi et al. (2015) pointed out, no single model consistently outperformed any other when considered across all land cover types. the selection of PET for model simulation varies depending on the region (Pimentel et al., 2023)". We believe that incorporating the most recent and relevant global datasets can strengthen the novelty and scientific contribution of our study.

4. **Attempting to access the data, I found it currently unavailable. Does ESSD typically require data accessibility for reviewers? Additionally, statements like "The data generated in our study are published in this public repository: https://doi.org/10.6084/m9.figshare.12132696.v1 (active after acceptance)" are too vague. Specific details regarding data accessibility (e.g., which specifical datasets) should be provided.**

Response: Thanks for pointing out the issues with the need for clearer information regarding the availability of our datasets. The previous provided link was intended to be activated post-acceptance. We will ensure it active and meet the standards of the journal. Here is the private link for reviewers to access the data: https://figshare.com/s/6e1eb7c9f2b6fd13edba

We provide the inputs including key parameters like canopy height and roughness, and outputs and add the following in the data statement so that it is clearly described in the manuscript: "The data provided along with this study include the key surface parameters, PET annual data from the main methods, precipitation, and SPEI dataset, available in this public repository: https://doi.org/10.6084/m9.figshare.12132696.v1".

To further enhance our transparency, we also add a code statement in the manuscript: "The code used to process data and perform analysis for this study is available in the public repository at https://github.com/pitcheverlasting/spei-pet-evaluation/"

In addition to all the above updates, we also proofread the manuscript, updated the figures and tables (like Table 3 and Fig. 2), corrected any remaining typos, resolved any conflicts/confusion in the previous draft.

References:

Abeysiriwardana, H. D., Muttil, N., & Rathnayake, U. (2022). A comparative study of potential evapotranspiration estimation by three methods with FAO Penman–Monteith method across Sri Lanka. *Hydrology*, *9*(11), 206.

Ershadi, A., McCabe, M. F., Evans, J. P., & Wood, E. F. (2015). Impact of model structure and parameterization on Penman–Monteith type evaporation models. *Journal of Hydrology*, *525*, 521-535.

Gao, G., Feng, Q., Liu, X., & Zhao, Y. (2021). Measuring and modeling evapotranspiration of a Populus euphratica forest in northwestern China. *Journal of Forestry Research*, *32*(5), 1963-1977.

Pimentel, R., Arheimer, B., Crochemore, L., Andersson, J. C. M., Pechlivanidis, I. G., & Gustafsson, D. (2023). Which potential evapotranspiration formula to use in hydrological modeling world-wide? *Water Resources Research*, 59, e2022WR033447. https://doi.org/10.1029/2022WR033447