Referee 2:

General comment:

This study presents a 23-year long-term benchmark ET dataset (2000–2022) based on global FLUXNET2015 observations. The dataset effectively addresses critical gaps in existing ET records at both hourly and daily scales, while also extending the time span. This makes it highly valuable for validating ET models and satellite-derived ET products. Therefore, this work has significant importance for the ET research community and is suitable for ESSD. Below are some minor suggestions to further enhance the quality of the manuscript.

Reply: Thank you for taking your precious time to review our article and also recognizing our work. We have made a detailed reply and they are as follows. We have also carefully examined the entire manuscript and corrected some ambiguous or incorrect expressions therein. All the changes are marked in red.

1: Line 14: "terrestrial"

Reply: Thank you for pointing out the mistake. We have modified "the terrestrial" to "terrestrial".

In revised paper Line 14:

"Evapotranspiration (ET) is a crucial component of terrestrial hydrological cycle."

2: Line 19: "This hinders their application." the sentence is too short

Reply: Thanks for your comment. We have rewritten this sentence to make it more specific.

In revised paper Line 17-19:

"However, existing LE_{EC} datasets, like FLUXNET2015, face significant challenges due to limited observation periods and extensive data gaps, which hinders their application in ET modelling and global change analysis"

3: Line 45: "With the abundance of data and the development of models" the sentence is not very clear.

Reply: Thanks for your comment. We have rewritten this sentence to make it clarity.

In revised paper Line 47-49:

"With the abundance of remotely-sensed and reanalysis data and the development of ET models, more and more ET products based on remote sensing or earth system model simulation are produced and shared."

4: Line 50: "as" is missing in "Since LEEC data are considered..."

Reply: Thank you for pointing out the mistake. We have added "as" after "considered".

In revised paper Line 53:

"Since LE_{EC} data are considered as the ground truth, researchers are eager to find evidence from ground observations to support their hypotheses."

5: Line 62: "hopes" seems not suitable

Reply: Thanks for your suggestion. We have replaced the word to "aspires" and rewritten this

sentence.

In revised paper Line 64-65:

"The research community aspires to leverage the most recent, long-term LE_{EC} data; however, there is a lack of up-to-date datasets that are readily accessible for their use."

6: Eq. 2-3: Ta and Td is better than ta and td. And also, ata is easily mistaken for a single symbol. **Reply:** Thanks for your suggestion. Here, we have changed the symbol of air temperature from 'ta' to 'T_a' and dewpoint temperature from 'td' to 'T_d'. The use of 'ata' can indeed lead to confusion. Therefore, we add a multiplication symbol between 'a' and 'T_a'.

In revised paper Eq. 2-3:

$$e_s = 6.1078 \times exp\left(\frac{a \times T_a}{T_a + 273.15 - b}\right) \begin{cases} a = 17.27, \ b = 35.86, T_a > 0\\ a = 21.87, \ b = 7.66, T_a \le 0 \end{cases},$$
(2)

$$e = 6.1078 \times exp\left(\frac{a \times T_d}{T_d + 273.15 - b}\right) \begin{cases} a = 17.27, \ b = 35.86, T_a > 0\\ a = 21.87, \ b = 7.66, T_a \le 0 \end{cases}$$
(3)

where e_s is the saturated vapour pressure (kPa), e is the actual vapour pressure (kPa), T_a is the air temperature, and T_d is the dewpoint temperature (°C).

7: Line 109: The sentence of "Its spatial resolution..." needs to be refined and polished. **Reply:** Thanks for your suggestion. We have rewritten this sentence and it is terse now.

In revised paper Line 111: "Its spatial and temporal resolutions are 250 m and 16 days, respectively."

8: Line 112: replace "contained" by "includes"

Reply: Thanks for your suggestion. We have replaced "contained" by "includes".

In revised paper Line 114:

"The gap-filling and prolongation framework for LE_{EC} data mainly includes 3 parts:..."

9: Line 123: "if"-> "when"

Reply: Thanks for your suggestion. We have replaced "if" by "when".

In revised paper Line 124-125:

"We calculated the daily energy balance ratio (EBR) when there were \geq 36 (18 for hourly data) valid observations in a day."

10: Line 129: What does the sentence of "thus we chose 2 more sites with relatively good data quality" mean?

Reply: Thanks for your question. Following our site selection criteria, no African sites initially qualified for inclusion. To enhance regional representation, we incorporated two additional African sites (CG-Tch and SD-Dem) that partially met the requirements. While these sites satisfied Criteria 1 and 3, they exhibited higher data gaps (57% and 64% missing values respectively, see Appendix A: Site information) compared to our 50% threshold. Nevertheless, these sites represent the best available data quality within the region. For clarity, we have reformulated this explanation in the revised manuscript.

In revised paper Line 130-131:

"Notably, no sites in Africa fully met the specified criteria. Consequently, we selected two additional sites that substantially met the essential requirements."

11: Line 212: Why did the author prolong the daily ET? not the hourly ET?

Reply: Thanks for your question. The decision to focus exclusively on daily-scale ET prolongation was mainly based on the consideration that **current mainstream ET products are predominantly available at daily and monthly scales** (Zhang et al., 2019; Zheng et al., 2022; Miralles et al., 2025). Therefore, prolonging daily-scale ET data aligns best with current practical application scenarios. Of course, if required in the future, we can also extend hourly-scale data to meet future research needs. We have added this explanation in our revised manuscript.

Reference:

Zhang, Y., Kong, D., Gan, R., Chiew, F. H., McVicar, T. R., Zhang, Q., and Yang, Y.: Coupled estimation of 500 m and 8-day resolution global evapotranspiration and gross primary production in 2002–2017, Remote Sens. Environ., 222, 165-182, https://doi.org/10.1016/j.rse.2018.12.031, 2019.

Zheng, C., Jia, L., and Hu, G.: Global land surface evapotranspiration monitoring by etmonitor model driven by multi-source satellite earth observations, J. Hydrol., 613, 128444, https://doi.org/10.1016/j.jhydrol.2022.128444, 2022.

Miralles, D. G., Bonte, O., Koppa, A., Baez-Villanueva, O. M., Tronquo, E., Zhong, F., Beck, H. E., Hulsman, P., Dorigo, W., Verhoest, N. E. C., and Haghdoost, S.: Gleam4: Global land evaporation and soil moisture dataset at 0.1° resolution from 1980 to near present, Sci. Data, 12, 10.1038/s41597-025-04610-y, 2025.

In revised paper Line 130-131:

"Current mainstream ET products are predominantly available at daily and monthly scales (Zhang et al., 2019; Zheng et al., 2022; Miralles et al., 2025). Therefore, prolonging dailyscale ET data aligns best with current practical application scenarios."

12: Line 304: What's the difference between the forward and the backward prolongation?

Reply: Thanks for your question.

The difference between the forward and the backward prolongation is two opposite time direction. For example, one site has the time cover from 2007-2014. Therefore, prolongation of 2000-2006 is the backward direction, and prolongation of 2015-2022 is the forward direction. A schematic diagram is shown below. We expect that the prolongation performance will be consistent in both directions. Our analysis confirms that the extension algorithm achieves consistent performance in both directions, meeting our expectations (Figure 7 in the manuscript). We have added the explanation in our revised manuscript.

In revised paper Line 232-234:

"The prolongation at the daily scale was conducted into two time directions: forward and backward. For example, one site has the time cover from 2007-2014. Therefore, prolongation of 2000-2006 is the backward direction, and prolongation of 2015-2022 is the forward direction. We expect that the prolongation performance will be consistent in both directions."



A schematic diagram for the forward and the backward prolongation.

13: For the datasets, what is the difference between the aggregated_daily and the prolonged_daily_200_2022?

Reply: Thanks for your question.

"Aggregated_daily" data was aggregated from the gap-filled half-hourly data to a daily scale. The start and the end times match the observation period at each site. "Prolonged_daily_2000-2022" data provides the prolonged daily data using the novel bias-corrected RF algorithm. The seamless data spans 2000-02-18 to 2022-12-31. For the prolonged part, the quality flag is set to 2. The rest is consistent with the aggregated daily data. In general, "Aggregated_daily" data is subset of the "Prolonged_daily_2000-2022" data. We have modified the "Data availability" part in our revised paper.

In revised paper Line 130:

"6 Data availability

Our released dataset mainly contains four types of data:

(1) Half-hourly or hourly gap-filled data: The data were well gap-filled data using the novel bias-corrected RF algorithm. Filenames include "HH" for half-hourly or "HR" for hourly data. The time format follows FLUXNET2015 standards, with paired timestamps recorded in local time. The start and end times align with the observation period at each site. For QC flags, a value of 0 indicates observed data, while 1 indicates gap-filled data.

(2) Aggregated daily data. This daily dataset was aggregated from the gap-filled half-hourly data to a daily scale. The start and the end times match the observation period at each site. *QC flags represent the percentage of valid hourly observations for each day.*

(3) Prolonged daily data: This dataset provides the prolonged daily data using the novel biascorrected RF algorithm. The seamless data spans 2000-02-18 to 2022-12-31. For the prolonged part, the quality flag is set to 2. The rest is consistent with the aggregated daily data.
(4) Aggregated monthly and yearly data: These datasets were aggregated from the prolonged daily data. QC flags indicate the proportion of days with >90% valid hourly observations per month or year. No distinction is made between prolonged data and completely missing daily

data. The time span for the monthly data is 2000-03 to 2022-12, and that for the yearly data is 2001-2022.

All files are formatted as csv files. NDVI and debiased reference variables from ERA5-Land are also provided in our released data. The product has been deposited at https://doi.org/10.5281/zenodo.13853409 (Li et al., 2024) and can be downloaded publicly."