Response to Referee #2

Assessing evapotranspiration dynamics across central Europe in the context of land-atmosphere drivers

Comments from Reviewer 2	
Review Report for Manuscript ID: egusphere-2024-3386 Title: Assessing Evapotranspiration Dynamics Across Central Europe in the Context of Land-Atmosphere Drivers General Comments This study provides a comprehensive evaluation of evapotranspiration (ET) products across central Europe using a combination of in-situ, remote sensing, and reanalysis datasets. The authors analyzed the performance of multiple ET datasets under different climatic conditions, particularly focusing on the severe drought. The study effectively addresses a research gap by assessing the agreement and discrepancies between ET products in the context of soil moisture (SM) and vapor pressure deficit (VPD) interactions. The manuscript is relevant for researchers studying land-atmosphere interactions, hydrology, and ecosystem responses to climate extremes. However, several key areas require further clarification and improvement to strengthen the manuscript before publication. The authors should clarify the rationale for dataset selection, improve the discussion on physical interpretability, and provide additional insights into the role of vegetation stress and uncertainty quantification.	Thank you very much for reviewing our study and for outlining the relevance of the manuscript. We tried our best to provide satisfactorily answers to all comments in order to improve and strengthen the manuscript.
Major Comments Justification for Selected ET Products, the authors compare ET estimates from various remote sensing and modeling products (MODIS, SEVIRI, GLEAM, ERA5-land, GLDAS). However, it would be beneficial to explicitly justify the selection of these specific products over other alternatives such as FLUXCOM, ETMonitor, EB-ET. MODIS is kind of ET more relying	Thank you for naming several other ET products and providing the reference for the EB-ET product. The reviewer mentioned products are higher order ET products derivated from either other ET products or satellite products (e.g., Fluxnet ET observations, MODIS), and often provide data outside the time period analyzed in this study. In this study, we compare ET products directly retrieved from the observations and mostly from one single sensor (e.g., ICOS LE observations, MODIS optical data,

on optical data. GLEAM is based on microwave data. Please check other thermal ET product. One could be EB ET. Chen et al. 2021, Remote sensing of global daily evapotranspiration based on a surface energy balance method and reanalysis data. Journal of Geophysical Research: Atmospheres, 126(16): e2020JD032873.	ERA5-land reanalysis, SEVIRI) between 2017-2020. However, we can add the proposed products to our list, when we introduce existing ET products, and further discuss their advantages and disadvantages. We will add sentences to justify in more detail why we choose these ET products (most commonly known and employed ones) and acknowledge the existence of other products, e.g.,: Line 104: 'In this study, we first compare the most common ET products from field measurements, modelling, and remote sensing across 104 central Europe for the period 2017 to 2020. **These selected products are well-known, commonly employed, and freely available.**' Line 488: 'Although there exist other ET products from remote sensing and modelling, e.g., (Jiménez et al., 2011; Mueller et al., 2013; Fisher et al., 2020; De Santis et al., 2022; Yu et al., 2023), the examined ET products in this study are appropriate when addressing global analyses since other products have either a more coarse spatial or temporal resolution (Yu et al., 2023), are limited to clear sky conditions (De Santis et al., 2022), which prohibits continuous time series of ET measurements, **or are higher order derivates from either field measured or merged remote sensing based products (Jung et al., 2019; Chen et al., 2021).'**
Additionally, the manuscript should discuss the potential biases associated with the retrieval algorithms used in each dataset and how these may affect ET estimates under different climatic conditions.	Certainly, we will add a paragraph for discussing on potential biases associated with the retrieval algorithms for every ET product.
Physical Interpretability and Model Dependencies, the study provides robust statistical comparisons but lacks a deeper discussion on the physical implications of the observed differences. For example, why do some products perform better at evergreen needle-leaved sites compared to agricultural sites? How do land cover heterogeneity and seasonal changes influence model uncertainties?	We discuss the performance of the different ET products for varying landcover classes throughout the manuscript, and discuss the influence of land cover heterogeneity on retrieval results in section 4. of the manuscript. More thorough analyses and discussion regarding the physical interpretability and model dependencies with that many ET products would lengthen the paper significantly, which is impractical and potentially more interesting when focusing on two or three ET products.
Since GLEAM incorporates reanalysis and satellite-based observations, its correlation with other datasets like ERA5-land and GLDAS-2 might be inflated. Have the authors accounted for interdependencies between products in their error analysis?	We accounted for the interdependencies between products and calculated the error cross-correlation (ECC) between all products (see sec. 2.3.1.) in order to statistically validate the interdependencies. We discuss this in sec. 3.1. as well as 4.1., and give the ECC results in supplement figure S8.

The role of vegetation stress and physiological controls (e.g., stomatal closure) in driving ET reductions during drought should be better discussed, perhaps using additional to support this point.	Certainly, we will add more discussion on the role of vegetation stress and physiological controls in ET reductions during the drought year 2018.
Evaluation of Uncertainty and Error Cross-Correlation (ECC), The extended triple collocation (ETC) analysis is a valuable approach, but some ECC values are quite high, particularly at agricultural sites. The manuscript should explicitly discuss how ECC influences the reliability of the results and whether certain datasets may be inherently dependent.	We discussed the ECC results and potential interdependencies between ET products in sec. 3.1. as well as 4.1., and showed that ETC results are uncorrelated to ECC results, since at station DE-Rus, which gave high ECC results (potential strong interdependencies), ETC results do not reflect any interdependencies.
Clarity of Figures and Statistical Significance, the scatter plots and time series comparisons are informative, but additional clarity is needed in figures showing product inter-comparisons (e.g., Figures 4, 5, 6). Including a statistical significance test for differences between ET products would enhance the rigor of the results.	We included significance tests in figures 8 and 9 since we agree that some analyses need statistical significance information to discuss the results reliably. However, figure 4 shows the time series of all ET products and for all stations. Meaning, we are showing one single time series of every ET product and for every station individually. We are not sure what kind of statistics would make sense for single time series that would provide any useful information. In figure 5, we show the scatterplots, for which we give the Pearson's correlation coefficient (R^2), root-mean square error (RMSE) and percentage bias (PBIAS) between each product and for each station in supplement figures S5-S7, which we discuss in sec. 3. and 4.
Minor Comments Grammar and Style: Some sentences are long and complex, making them difficult to follow. Consider simplifying and improving readability. For example: The ICOS network has undertaken a large effort to ensure high-quality LE measurements, which are comparable among different ICOS stations.". Suggested revision: "The ICOS network has made significant efforts to ensure consistent high-quality LE measurements across stations.". Line 555 Grammar mistake, This is, products were most consistent with each other at stations with less complex land cover conditions and changes throughout the seasons (the evergreen needle-leaved stations DE-Ruw and FI-Let).	Well taken, we can certainly improve the grammar and style.
Line 49, rephrase the sentence 'Since precipitation (P) 'and evaporation are the two key components of the global water cycle' (Miralles et al., 2011), another important proxy for analyzing water stress and	Done, we can rephrase the sentence: 'Evapotranspiration (ET) is an important proxy for analyzing water stress and its effects on ecosystems since

its effects on ecosystems is evapotranspiration (ET).' .	precipitation (P) 'and evaporation are the two key components of the global water cycle' (Miralles et al., 2011).'
Line80, optical, thermal, infrared, or microwave observations are used to derive ET based on surface energy balance, physical and empirical models (Bayat et al., 2021, 2024; Rahmati et al., 2020; Zhang et al., 2016). The cited reference does not include thermal observation based ET from surface energy balance method.	Thank you for pointing this out. We will include two additional references, explicitly addressing thermal based ET and the surface energy balance method: 'Although it is not directly measurable from remote sensing acquisitions, optical, thermal, infrared, or microwave observations are used to derive ET based on surface energy balance, physical and empirical models (Zhang et al., 2016; Rahmati et al., 2020; **Singh et al., 2020;** Bayat et al., 2021; **Bhattacharya et al., 2022;** Bayat et al., 2024).' These references are: Singh, R.P., Paramanik, S., Bhattacharya, B.K. *et al.* Modelling of evapotranspiration using land surface energy balance and thermal infrared remote sensing. *Trop Ecol* **61**, 42–50 (2020). https://doi.org/10.1007/s42965-020-00076-8 Bhattacharya, B. K., Mallick, K., Desai, D., Bhat, G. S., Morrison, R., Clevery, J. R., Woodgate, W., Beringer, J., Cawse-Nicholson, K., Ma, S., Verfaillie, J., and Baldocchi, D.: A coupled ground heat flux–surface energy balance model of evaporation using thermal remote sensing observations, Biogeosciences, 19, 5521–5551, https://doi.org/10.5194/bg-19-5521-2022, 2022.
Terminology Consistency: The terms "ET estimation," "ET retrieval," and "ET modeling" are used interchangeably. It would be beneficial to define them more precisely and use consistent terminology throughout the manuscript.	Done. We will use one terminology (ET estimation) throughout the manuscript. The term 'modelling' is only used for modelled ET products (e.g., GLDAS-2). The term 'retrieval' is only used when talking about the retrieval algorithm/method of the ET products.
Temporal Aggregation Effects: Some ET product has a lower temporal resolution than other datasets. Have the authors checked whether this affects the observed discrepancies, if upscaled to 15 days, even monthly temporal resolution?	Yes, we have checked the effect of temporal aggregation. In the manuscript, we included some of these analyses in figure 9 and supplement figure S10 by comparing the Kernel density estimates of ET anomalies on daily and 8-daily time scales. Here, we clearly see a difference in the density curves and hence, included it in the publication as these are the main temporal scales included in the paper. As most products provide high temporal resolution (< daily) except for MODIS, which provides 8-daily ET data, analyzing other temporal scales (15-daily, monthly) is out of the scope of the current manuscript.
Line 558, The authors wrote that: The remote sensing products, SEVIRI, MODIS, and GLEAM, performed equivalently well or even better than the in-situ	Thank you for pointing this out. We are not saying that remote sensing products are better than in-situ observations. What we mean is, that for our specific study design

measured (ICOS), I don`t understand why the remote sensing products can be better than in-situ measured data? This is confusing readers. How can a satellite ET product be better than measurement?	(3 km footprint, daily analyses), the performance of the remote sensing products is overall more comparable and consistent among all investigated datasets, including ICOS. We discussed the reasons for this, e.g., in lines 465-478, also providing the drawbacks of in-situ measured ET observations at ICOS stations. We can rephrase the sentence in line 558 in order to clarify the meaning. 'The remote sensing products, SEVIRI, MODIS, and GLEAM, performed equivalently well or even better than the in-situ measured (ICOS), modelled (GLDAS-2) or reanalysis (ERA5-land) products **for this specific study concept (3 km footprint, daily analyses)**.'
Line 514, ET is more controlled by atmospheric demand rather than atmospheric supply, I can understand when the atmosphere is warming, it will need more vapor evaporated from ground. This could be a kind of atmospheric demand, but do not understand what is atmospheric supply? What kind of supply from atmosphere can influence ET? Are you saying precipitation? Please rephrase this sentence to make it clear.	Thank you for pointing this out. We actually meant soil water supply (soil moisture) as we see no dependency of ET on SM but variations of ET with VPD during wet years. We can rephrase the sentence accordingly: 'Here, our results indicate that ET is more controlled by atmospheric demand rather than water supply from atmosphere **(precipitation) and soil (soil moisture)** as reported also by Zhou et al., (2019).'
Line 523, Further, results show that VPD and SM are negatively coupled during extreme events as reported also by (Zhou et al., 2019)-à by Zhou et al. 2019. Same as reported by (De Santis et al., 2022).	Thank you, we will include De Santis et al., 2022 as additional reference.
Section 4, there are many other global ET product, which are not discussed. Please check and compare them.	As mentioned in our second answer above, we will mention other ET products and include additional information why we choose to compare these ET products.