Point-by-point reply to the comments

This document presents comments by reviewers and our point-by-point reply. The reviewer's comments appear in black and our responses appear in blue.

1. Reply to the comments by REVIEWER #1

I read the manuscript "How much water vapour does the Tibetan Plateau release into the atmosphere?" with great interest. The validation of many different ET products over these water towers of Asia has a lot of value. While the manuscript is generally well written and clear, I do have some specific comments and requests for clarification of the presented analyses.

Reply: We thank you for the review and the constructive feedback that helps us to improve our work.

Regarding the validation:

- Provide clear explanation on the temporal scale the analyses were conducted (monthly?), this is not always clear

Reply: We revised accordingly to make it clear. Both the validations based on eddy covariance observations and the basin-scale water balance method were conducted against monthly observations / estimates.

- Provide clear explanation on the period used for the analyses (in some cases the overlap of the *in situ* data (either EC towers or water balance estimates) and products is rather short

Reply: It is true that the overlap period of the *in situ* data and products is short in some instances, and in some case such as the Namco site there is no overlap, since the in-situ measurements started in 2019 while some products did not extend beyond 2019. As regards the site-scale validation, we added a table in the supplementary materials to specify the overlap period for validation, the number of observations, and values of the error metrics. As regards the basin-scale validation, the validation period was also added in the main text.

Our approach was to utilize long time series data (as long as possible) for the inter-comparison and trend analyses. More precisely, for the inter-comparison analysis we used the overlap period of all products (2003~2013). The trend analysis was carried out for the available period for each dataset, being aware that the overlap period of all products was relatively short. We note that many satellite remote sensing ET datasets with high spatial resolution are estimated based on MODIS data, which started from 2000, while there is still a lack of long-term ET datasets with high spatial and temporal resolutions.

- Basins used in the water balance estimation is not always clear, eg figure 1 doesn't show the Heihe basin (is this the Hexi corridor and is the entire basin included in the map/analyses?). In figure 1 what does the stripped area refer to? A table with information would be useful with some additional information on the data used from the studies by Ma and Zhang and Wang et al. Also the basins are referred to as the Yangtze/ Yellow river basin, but as far as I understand these only cover the upper part of the basin. Please provide some additional information on the extent of each of the basins analyses (eg provide name of the gauging station where the basin was delineated). Also in figure 3, there is a reference to TP, which basin/ area does this refer to (the entire TP area shown in figure 1 or the area of all the basins combined, which are two different areas)?

Reply: Sorry for the ambiguities in some of our illustrations and related information. We revised the figures and added more information accordingly. Overall, we used monthly ET_{wb} from five basins from previous studies (Ma and Zhang., 2022; Wang et al. 2021), including the headwaters of Yellow basin (HYE), headwaters of Yangtze basin (HYA), upper Heihe basin (UH), Inner Tibet Plateau (INTP) and Qaidam (QDM) basins. It is true that these only cover the upper part of the basin, and we defined explicitly the extent of these regions and presented this information. A new table (Table 2 in the revised version) was added to provide additional information, i.e., the extent of the basins and the names of the gauging stations.

As regards Figure 3, we intended to use TP to represent the area of all the five basins combined. To avoid the potential misunderstanding, we revised it to *5 basins* (the area of all the basins combined) in the new version.

- Color scheme of figure 3 is not fully intuitive, for example the r2 is deep red for high (=good) values)

Reply: We revised the figure to make it more intuitive.

Figure 5: what do the different colors of the bars mean?

Reply: We added the description in Figure 5. The global satellite remote sensing-based ET dataset are shown in dark blue and model-based ET dataset in light blue, and the regional ET datasets are in red.

Trend analyses (figure 7):

- The calculation of the trends could be affected by an exceptional year with high or low ET at the beginning or end of the time series (since there is quite some yearly variation and the trends are often relatively minor). Could you say something about the significance of these trends as well? Also for the SynthesisET both the first two years and the last two years seem to be outliers and related to the “temporal inconsistencies” of the product. Was this data properly vetted before including in the analyses?

Reply: We fully agree with you that the trend could be affected by the exceptional years at the beginning or end of the time series. This is also why we choose a robust regression method to estimate the trend of ET, rather using simple linear regression, since the robust regression can reduce the impact of outliers. We added the significance level of the trends in the figure and main text.

As regards the temporal inconsistencies of SynthesisET, we carefully checked it for several times and we are pretty sure about the existence of the temporal inconstancies. In fact, this issue was also noticed by the authors of the SynthesisET dataset, and they tried a different synthesis strategy d in a later regional study on the Northern China (Wang et al., 2021). This could also be seen from the temporal variation SynthesisET in Figure RC1-R1. Figure RC1-R1 was used to replace the Figure 7 in the manuscript.

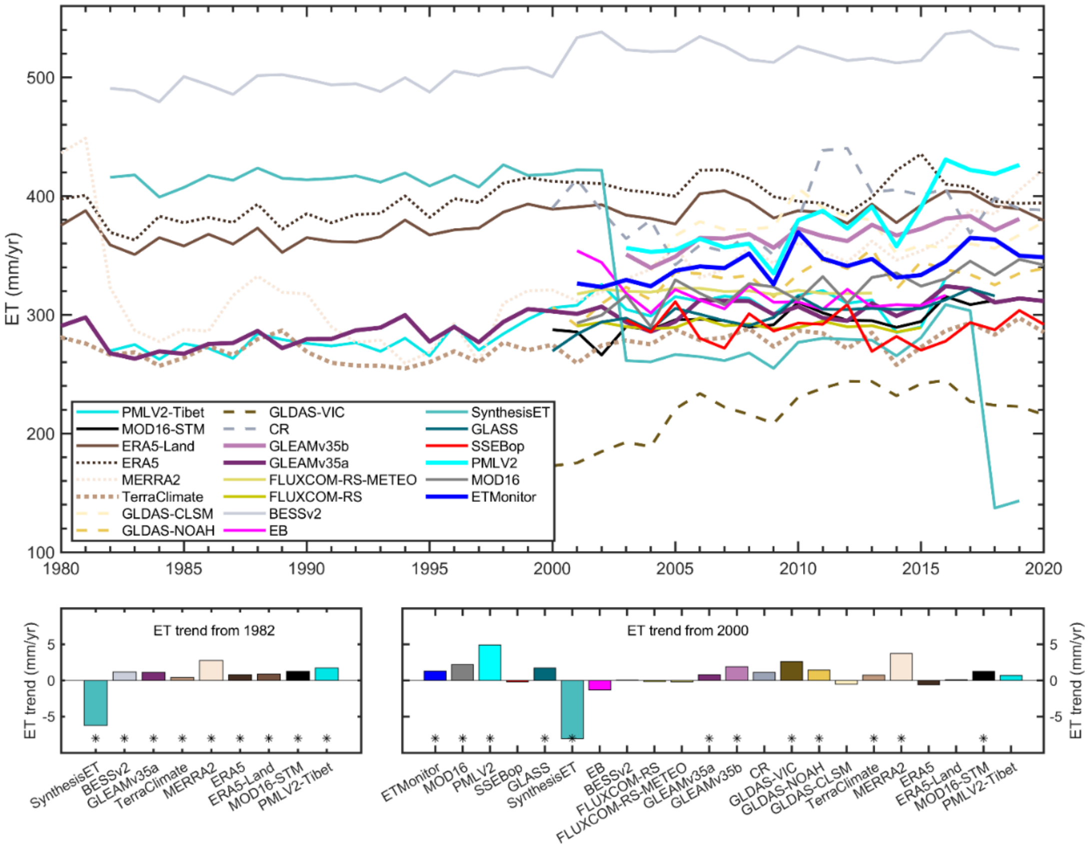
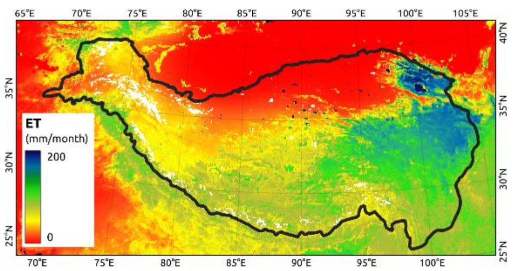
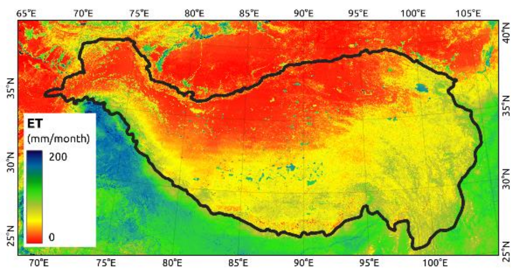


Figure RC1-R1: Yearly variation of ET in the TP by different products. The inset panel shows the annual ET trend by different products. *: trend with significance level ($p<0.05$). In the top panel, the reanalysis data is shown as a dotted line, and the land surface model-based data is shown as a dashed line.

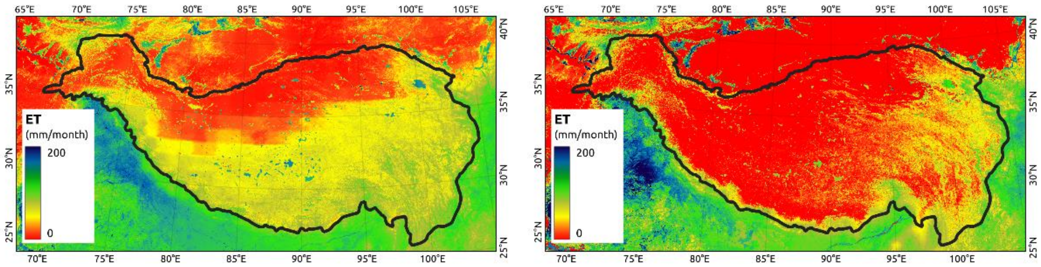
We also checked the spatial variation of ET by SynthesisET (as shown in the following Figures RC1-R2). Before 2000, SynthesisET showed quite high ET values (e.g., in the eastern TP), while after 2019 SynthesisET showed extremely low ET values in the eastern TP.



SynthesisET, July, 2000



SynthesisET, July, 2008



SynthesisET, July, 2013 SynthesisET, July, 2019

Figure RC1-R2: Example of spatial variation of ET by SynthesisET in July of different years.

Reference:

Wang, L.;Wu, B.; Elnashar, A.; Zeng, H.; Zhu,W.; Yan, N. Synthesizing a Regional Territorial Evapotranspiration Dataset for Northern China. Remote Sens. 2021, 13, 1076. https://doi.org/10.3390/rs13061076.

- Why are many of the products with longer time series (eg ERA5Land, SynthesisET, BESS, MERRA2) not presented with their full timeseries?

Reply: The ERA5-Land ET shows a very similar trend as ERA5. As regards SynthesisET, we already noticed its temporal inconsistence, thus we did not include it in the annual trend analysis. To reduce the concerns of reviewers, we included all the long-term ET products in the revised version.

Analyses of "ET components"

- As mentioned by the authors these different sub-components of ET are not validated and with the wide range of values derived from the different products, what conclusions can really be drawn? This is especially a question for the open water ET (maps in figure 9 shows large areas evaporating from water surfaces) and sublimation (which is validated how?)

Reply: It is true that the evaluation of different ET components was still limited due to the scarcity of available data and a comprehensive evaluation based on more observations would help to further evaluate the ET components and improve the algorithm performance. This analysis on the ET components has not been fully investigated in previous studies. We intended to use it to explain the difference among ET products and to answer the question: which processes play a significant role in determining the total ET. We also noticed that previous studies mostly focus on total ET, e.g., magnitude, spatial variation, temporal trend, etc., while the ET components were not fully investigated. Meanwhile, many studies were based on a big-leaf model, and a few studies estimate total ET based on the separate estimation of ET components. These components reflect the different water phase change processes that are regulated by different factors, e.g., transpiration is mainly controlled by the plant physiology through the regulation of stomata behavior, soil evaporation is determined by heat and mass transfer in the top soil with liquid water present at some depth below the surface, the rainfall interception loss is mainly related to the canopy morphology and rainfall intensity and the sublimation is associated with higher enthalpy change than vaporization process and near surface air humidity and temperature. So, we believe this analysis on the ET components is helpful, because at least starts with treating correctly each water phase change.

It is important to note that reliable independent reference measurements on each component of total vapour flux are very scarce. The anonymous Referee #3 (RC3) suggest us to use the ensemble mean of the ET components by different products, which may be close to the truth. We also notice that averaging properly would not provide good estimates, since the it applies only to random errors, not to the use of the wrong algorithm. According to the results in Section 3.2.3, the median values of the ratio of Es, Ec, and Ei to total ET is 50%, 30%, and 5%. A recent study shows the contributions of Es, Ec, and Ei to total ET are 68.21 %, 23.57 %, and 8.21 %, respectively in the Three Rivers Source of the Tibetan Plateau (Zhuang et al., 2024), which is actually quite close to our estimates. After the analysis in our study, we may generally conclude that soil evaporation (Es) contributes most to total ET in the whole TP, and further study should pay more attention to it. We also noticed that the phase change of snowfall is poorly known. These events are short but widespread in the TP, with snow-cover being extensive but short-lived. Both snow melt followed by evaporation and infiltration and sublimation are relevant and will be investigated.

Reference:

Zhuang, J., Li, Y, Bai, P, Chen, L, Guo, X., Xing, Y., Feng, A, Yu, W., Huang, M.: Changed evapotranspiration and its components induced by greening vegetation in the Three Rivers Source of the Tibetan Plateau. J. Hydrol., 633, 130970, https://doi.org/10.1016/j.jhydrol.2024.130970, 2024.

Analyses related to the “response to different environmental factors”

- The purpose of these analyses are not entirely clear to me. First, the analyses are done for the median value of the correlation, whereas it was already very clear that there is a large variance between the different products. Also several products utilize these input data (Rn, LAI, P) for estimating ET, how is this kind of dependency considered in the analyses? Do different types of models have stronger or weaker correlation with these environmental factors? And what does that mean for the interpretation of the analyses?

Reply: Thank you very much for the comments. Analyzing the impact of environmental factors on ET is helpful to reveal the governing factors and the mechanisms determining the variability of ET. It is also helpful to analyze whether and how the ET algorithms/product capture the ET variation caused by the environmental factors. It is true that different models have stronger or weaker correlation with these environmental factors, which indicate the observed response to forcing factors is algorithm dependent. Several products utilize these input data (Rn, LAI, P) for estimating ET, and these products may show higher correlation with these factors. Hence, we think both the algorithm itself and the input data can impact the response of estimated ET to environmental factors.

We also noticed that the current analysis is very limited and a more comprehensive analysis could be done to illustrate this issue better. A proper treatment would require a significant amount of additional materials and we decided to leave it out for the time being. Hence, we removed it from current manuscript and prepare another paper on it for a more robust analysis following the suggestion of Referee #2 (RC2).

- Did any of these factors also influence the partitioning of ET into ETc and ETs?

Reply: We did not mention this issue in the manuscript. But, we think the answer is yes. This is especially true for leaf area index. Higher leaf area index is generally associated with higher plant transpiration and interception loss. For example, a recent study shows that the vegetation greening (judged by increasing LAI by 0.009 $m^2/(m^2\ a)$ with $p < 0.05$) caused different changes in ET and its components, i.e., 1.95 mm/a, −2.41 mm/a, 1.33 mm/a, and 3.03 mm/a for ET, Es, Ec, and Ei, respectively, in the Three Rivers Source of the Tibetan Plateau (Zhuang et al., 2024), which clearly indicates its influence on the ET partitioning.

Reference:
Zhuang, J., Li, Y., Bai, P., Chen, C., Guo, X., Xing Y., Feng A., Yu W., Huang, M.: Changed evapotranspiration and its components induced by greening vegetation in the Three Rivers Source of the Tibetan Plateau. Journal of Hydrology, 633, https://doi.org/10.1016/j.jhydrol.2024.130970, 2024.

Discussion:

- General reflection of the validation methods employed, doesn't really add much information. The incorporation of seasonal land cover conditions or lack thereof is only explained for 3 products, but then no reflection on how that has affected the results. Or how relevant negative latent heat fluxes are (does this happen often or only occasionally?). The reflection on the water balance estimations are also very general and could have been included in the introduction (there is no reflection based on this specific study). For example, the assumption of not incorporating meltwater could have been explained in the method but is not an outcome of this research.

Reply: We understand the reviewer' concern. We moved some general comments to the introduction and revised the discussion section to focus more on the findings of the current study as follows:

The in-situ observations with an eddy covariance system are recognized as the standard method for monitoring energy and mass fluxes to validate high-resolution ET (Baldocchi, 2020). In addition, the ET products were compared with the basin-scale water balance estimates ET_{wb}. ET_{wb} is obtained at the basin scale (several hundred km^2), which is much larger than the footprint of flux tower observations (approximately km^2, depending on meteorological conditions). Given the relatively sparse distribution and small footprint of the flux-tower-based eddy covariance system observations, the water balance method can serve as a useful complementary reference for ET estimates. This is especially true for the coarse-resolution ET, which has a much larger spatial footprints than eddy covariance observations.

In this study, these two methods gave generally consistent results when evaluating the high-resolution ET. When judged by the KGE of site-scale estimates, the accuracy of the high-resolution ET products can be ranked as follows: PMLV2 > ETMonitor > MOD16-STM > GLASS > MOD16 > SynthesisET > SSEBop. When judged by the KGE of basin-scale validation, the accuracy of the high-resolution ET products can be ranked as: ETMonitor > PMLV2 > MOD16STM > SSEBop >

GLASS > MOD16 > SynthesisET. Although both indicate that ETMonitor, PMLV2, and MOD16STM are the most accurate and the remaining four are less accurate among the high-resolution ET products, some differences in the ranking of the ET products can be observed. This is probably related to the processes captured by the 'ground-truth' data at different scale used in the two evaluation methods. An eddy covariance observation represents the net water vapour flux integrated across different processes at given point (e.g., plant transpiration in the dense vegetation regions, snow sublimation in dry snow cover regions, evaporation of canopy-intercepted water when the canopy is wet due to intercepted rainfall). In addition, the observed vaporization process depends on the land surface conditions at the observation sites during particular times, which may vary seasonally and annually due to factors such as snow/ice, intercepted water, and vegetation. The estimated basin-scale ET by water balance (ET_{wb}) was essentially the residual of the observed water balance terms, which is assumed to be the net liquid water flux loss to the atmosphere at the basin scale. Compared to the site-scale observation, the basin-scale ET_{wb} can capture the effect of land cover dynamic on the ET within the basin. For example, the mean water level of lakes in TP increased by 0.20 m/yr from 2000 to 2009, and the lake water mass increased significantly (Zhang et al., 2013), which caused higher ET in the TP because water evaporation is generally higher than other land cover types. However, most ET products (e.g., MOD16, PMLV2, etc.) assume constant land surface conditions throughout the year or multiple years, which means that they cannot capture the temporal transitions of the vaporization process associated with changes in land cover. In contrast, ETMonitor adjusts the daily land cover based on dynamic land cover conditions, including water bodies cover and snow/ice cover, which allows it to reflect the impact of seasonal and annual open water extent and snow/ice cover on total ET (Zheng et al., 2022). This probably explains in part why ETMonitor performs slightly better than PMLV2 when validated by basin-scale water balance methods, while they are comparable when validated by in-situ observations.

Reference:

Liu, H, Xin, X, Su, Z., Zeng, Y., Lian, T., Li, L., Shanshan S.: Hailong Zhang Intercomparison and evaluation of ten global ET products at site and basin scales. J. Hydrol., 617, 128887, https://doi.org/10.1016/j.jhydrol.2022.128887, 2023.

Zhang, G., Yao, T., Xie, H., Kang, S., and Lei, Y.: Increased mass over the Tibetan Plateau: From lakes or glaciers?, Geophys. Res. Lett., 40, https://doi.org/10.1002/grl.50462, 2013.

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.

Baldocchi, D. D.: How eddy covariance flux measurements have contributed to our understanding of Global Change Biology, https://doi.org/10.1111/gcb.14807, 2020.

Chen, X. Yuan, L., Ma, Y., Chen, D., Su, Z., Cao., D.: A doubled increasing trend of evapotranspiration on the Tibetan Plateau. Sci. Bull., https://doi.org/10.1016/j.scib.2024.03.046, 2024.

- The discussion related to the different types of models comes a bit out of the blue, for example in table 2 the model type is not provided, which makes is difficult to validate a statement such as (first sentence) " PM-type model demonstrated superior accuracy compared to other models". Also ".. models that incorporate soil moisture to detect water stress…" can not be checked, which models do or do not incorporate soil moisture? Also to go in depth into the methodology of each product seems to go beyond the objective of this research, especially since it unclear why some models are singled out and others not (nor a statistical comparison between for example PM vs non-PM models is not done.

Reply: We have double checked and revised the manuscript to make sure all the necessary information is included and the statements can be easily checked. We already stated in the manuscript that "Among the evaluated ET products, there are 14 products that primarily use remote sensing products, including 2 products (SSEBop and EB) based on land surface temperature (LST), 8 products (ETMonitor, MOD16, MOD16-STM, PMLV2, PMLV2-Tibet, GLEAMv35a, GLEAMv35b, BESSv2) based on PM-types models (including Penman-Monteith equation, Priestley-Taylor equation, Shuttleworth-Wallace equation), 4 products (FLUXCOM-RS, FLUXCOM-RS-METEO, GLASS, SynthesisET) based on data-driven methods (machine learning method or ET products ensemble method)." To make the information more intuitive, we moved it to the Section 2.2.2. More information on whether soil moisture is considered in a given data product was added in Table 2 by listing the main forcing data.

Our primary objective is to find out how accurate are the ET products in the TP, which is closely related to the algorithm applied in each product. Since we evaluate 22 products, there are 22 models to be discussed, which is actually too much and will make the manuscript unfocused. Therefore, we discussed the methodology of some representative ET products. Some evidence on the difference between the PM and non-PM model can be found in Section 3.1.1, which shows that the best three products are all PM -type model-based products (ETMonitor, PMLV2, MOD16STM), while the LST-base (SSEBop) and data-driven products (GLASS and SynthesisET) had overall a low accuracy. We revised the manuscript to present this statement more clearly in the revised version.

- The uncertainty of the SynthesisET product was already mentioned in the results section, is this really an important outcome of this research (important enough to single it out in the discussion?)

Reply: Thank you for the comments. In the results section, we evaluated its accuracy and compared with other products to identify a temporal inconsistence. In the discussion section, we try to explain the reason of its relatively poor performance, since we expected the fusion of different datasets should have improved the overall accuracy. We addressed the importance of the ensemble method in the discussion, which might be helpful to guide further studies.

---

2. Reply to the comments by REVIEWER#2

The article "How much water vapor does the Tibetan Plateau (TP) release into the atmosphere?" by Zheng et al. provides a comparative analysis of evapotranspiration (ET) on the Tibetan Plateau, an essential yet uncertain component of the water cycle. This comprehensive review examines various streams of ET data and compares them with in-situ flux measurements, aiming to address a significant research gap: Can ET estimates derived from satellite and land surface models accurately reflect in-situ ET observations?

While I appreciate the insights offered by this article, particularly its thorough incorporation of diverse data sources, there are concerns regarding the clarity and completeness of the methodology. Consequently, the obtained results lack sufficient substantiation. Therefore, before publication, these concerns need to be addressed thoroughly.

Reply: We thank you for the positive and constructive feedback that helps us to improve our work.

Major comments:

Regarding the Methodology:

1. The temporal coverage of the analysis is not clearly defined throughout the article. In line 197, it is written 2003 to 2015, in line 221, 2001 to 2018 while in line 312, it is written 2000 to 2020. These discrepancies need clarification to ensure consistency and accuracy in the reporting of the study period. I suggest keeping the results with consistent temporal coverage in the main section while any other information on supplementary information (SI).

Reply: Thank you very much for the suggestion, and we try to use consistent temporal coverage in the main section. The differences in the overlap period were caused by differences in the temporal coverage of ET products and *in situ* observations. In section 3.1.1 on the validation of ET products against flux tower measurements, the overlap period for *in situ* eddy covariance observations and high-resolution ET products was in most cases from 2001 to 2018, but there are differences for some sites and products. We added Table S1 in the supplementary materials to show the temporal coverage for each site and each product for site-scale validation. In section 3.1.2, when the validation of ET products against basin-scale water balance ET_{wb} is presented the temporal coverage was from 2001 to 2015 with some gaps for some catchments and data products. We also added a table (Table 2) to show the temporal coverage period and integrating the information related to basin-scale validation where necessary. In section 3.2.2, '2000 to 2020' is not the precise temporal coverage of different products. Instead, we determined the median value of ET of all available products for each year between 2000 and 2020. The median value of ET was further used to obtain the overall trend of ET from 2000 to 2020. We revised this part to make it clearer.

Furthermore, to avoid unnecessary confusion caused by differences in the temporal coverage, we try to keep a consistent temporal coverage and add information in each section where necessary. Considering the temporal coverage of all products is from 2003 to 2013, the comparative analysis in Section 2.3.2 and Section 3.2 were conducted by applying the period 2003~2013, unless gaps in data had to be taken into account, leading to a different temporal coverage.

---

2. Although it appears to be conducted at a monthly scale based on the information provided, it is unclear whether all datasets, such as ETMonitor with daily resolution and MOD16 with 8-daily resolution, were aggregated to a monthly scale for comparison or were based on the native resolution of the dataset. Clarity is needed regarding the aggregation process of these datasets to ensure transparency and understanding of the methodology employed.

Reply: It is true that the validation was carried out using monthly data. All the products were temporally aggregated to monthly values from their native temporal resolutions prior to validation and comparison. The data products with a daily resolution were just added up to obtain the monthly ET values. For the data with 8-days resolution, an average ET value was first estimated for the available data in that month, and the monthly ET value was subsequently obtained by multiplying the averaged values by the number of days in the month. We added this description to clarify how the monthly data were obtained.

---

3. In line 135, it is mentioned that months with less than 50% valid daily ET values were excluded from the analysis. However, it remains unclear whether these excluded months were filled to maintain a continuous ET time series or if the comparison was limited to months with more than 50% valid ET values. Clarification on how the missing data was handled and its impact on the analysis is necessary for a comprehensive understanding of the methodology.

Reply: The missing data was not further filled and gaps were excluded to avoid the impact of uncertainty introduced by gap-filling. We stated this in the methodology.

---

4. Providing information on the number of valid observations available for each dataset, either in the supplementary information or elsewhere, would be beneficial for assessing the comparability of sample sizes across datasets, especially if they are not analyzed for same temporal coverage.

Reply: Thank you very much. We added a table in the supplementary materials to include the temporal coverage for site-scale validation and number of valid observations.

Regarding the results:

1. It appears that the regional-based formulations of ET, such as MOD16STM and PLMV2 ET Tibet, demonstrated the highest accuracy when compared to in-situ flux towers. However, it is crucial to ensure that the flux stations utilized in this study for comparison were not already included in the calibration of these datasets. If the same flux stations were used for calibration, the greater accuracy of these products may not be fully substantiated. Therefore, it is imperative to verify whether there is any overlap between the flux stations utilized in this study and those used for calibration to accurately assess the reliability of the results.

Reply: The issue you mention is very important, and we agree with you that the validation results are influenced by the calibration. As a summary, the calibration sites were clearly listed for three high-resolution products (ETMonitor, PMLV2, MOD16STM), and some of these sites were used for validation in this study. Other products did not use flux sites for calibration or this information is not presented in the corresponding studies. Although some coarse-resolution products (e.g., PMLV2-Tibet) were also reported to use flux sites as calibration, they were not validated based on flux site observations in this study, considering the mismatch of spatial representativeness between in-situ observations and coarse-resolution products. In this study, we did not exclude the calibration sites in our validation study for the following reasons:

- The difficulty in maintaining ground-based observations have resulted in a scarcity of flux towers on the TP. If calibration sites were excluded, the validation sites would be scarce, which would raise further concern on the sites' representativeness and relevant uncertainty.

- Different products use different sites for calibration, and some studies did not provide such information. Some products were designed with a clear separation between calibration and validation sites, while others did not. For example, some studies clearly separated calibration and validation samples using data of different years from same sites, while other studies did not provide clear information at all. It seems to be not feasible in practice to apply a well-defined screening of calibration and validation data.
- To achieve high accuracy, model calibration is a valid approach applied for many models before generating datasets. The purpose of this study is to identify how accurate the current ET products are, which might help to achieve an ET product with better accuracy, and efforts on model calibration should be encouraged.

To address this, we included the information on whether the sites were utilized for each ET product calibration in the supplementary materials Tables S1. In addition, we performed basin-scale validation to strengthen our findings. To our best knowledge, there is no product using basin-scale water balance estimates for calibration, i.e. this approach as an independent validation method.

---

2. In Figure 3, it is unclear how the metrics were calculated for the entire Tibetan Plateau (TP). Does the metrics for TP represent averages or medians across the basins or was TP treated as a single basin?

Reply: Sorry for the misunderstanding. We intended to use *TP* to represent the area of all the five basins combined, including headwaters of Yellow basin (HYE), headwaters of Yangtze basin (HYA), upper Heihe basin (UH), Inner Tibet Plateau (INTP) and Qaidam (QDM) basins. To avoid any misinterpretation, we revised it to *5basins* (the area of all the basins combined) in the revised manuscript. A new table was added to provide additional information in the Supplementary., We simply used all samples (each sample represents a valid group of reference data and to-be-validated ET data from one basin) from all 5 basins to estimate the metrics for the *5basins*.

---

3. In Figure 4, the color bar for ET standard deviation (ETsd) differs from the color bars used for other variables. This inconsistency can lead to confusion, particularly since the figures are presented together. Also, if possible, please keep the results in the order of datasets that appears in the Table 1.

Reply: Thank you for the suggestion. We revised the figure accordingly to avoid the confusion. We also moved the information on the spatial variability of ET in each product to the supplementary materials to make the manuscript more concise.

---

4. Regarding Figure 7, it would be beneficial to highlight the trends observed specifically from data with long records to discern the presence of significant trends in ET, because the trend calculated with only some years of data would not add any conclusion to the overall trends in the ET. Additionally, it's essential to clarify how the trends were calculated—whether through linear regression or another method—and whether the significance of these trends was assessed.

Reply: Thank you for the suggestion. We also noticed that the trend could be affected by the temporal coverage of the ET time series, and we also agree that longer records provide more reliable information on trends. We identified the trends estimated with long records in the revised version.

We believe that relatively shorter data records (especially in recent years) remain relevant to document differences across data products, so we kept the results on trends after 2000s.

We applied a robust regression method to estimate the trends, rather than using simple linear regression, since the robust regression reduces the impact of outliers. We added the significance level of the trends in the figure and main text.

On results specific to "Response of the ET to main governing factor."

The author's intended message or purpose behind the analysis is not clearly conveyed. It seems to explore the relationship between annual ET and various water, energy, and vegetation variables. I will try to highlight my concerns in points here:

1. In my belief, the analysis of how annual ET responds to different water, energy, and vegetation variables could potentially be a separate study requiring a more comprehensive approach.

Reply: Thank you very much for the suggestion. We agree with you that the response of ET to water, energy, and vegetation variables could be a topic to be addressed by a more comprehensive analysis. We removed it from current manuscript and prepare another paper on it for a more robust analysis.

---

2. For instance, If Leaf Area Index (LAI) correlates well with both/or net radiation (Rn) and precipitation (P), which I believe will be the case, raises doubts about the conclusions drawn regarding the relative influence of these variables on evapotranspiration (ET). This is true especially when conclusion on influence of these variables on ET is drawn simply from correlation of ET with these variables without controlling for the other confounding factors. To check whether this is the case or not, we can simply correlate LAI with Rn and P, as well as by correlating Rn with P.

Reply: It is true that LAI correlates with both Rn and P. LAI is a critical variable that correlates with several climatic and environmental factors as it represents the amount of leaf area per unit ground area and characterizes the canopy structure. LAI influences the interception of radiation and the distribution of light within the canopy, which in turn affects the energy balance of the surface, e.g., net radiation (Rn) and latent heat flux (LE). Also, plant generally grow better in regions with sufficient water supply (high precipitation) and adequate APAR (highly related to Rn).

---

3. Even if one were to accept the current analysis, which I personally disagree with for the reasons outlined in points 2, there remains a crucial need for clarification regarding the rationale behind correlating median ET from all datasets (if I understood it properly) with environmental variables (Figure 10). This need arises primarily from the significant variability observed among different ET datasets in terms of magnitude and hence I believe that the relative importance assessed from the simple correlation of ET with these variables will also vary. Consequently, any conclusions drawn from these correlations may lack robustness.

Reply: We agree that there could be issues with determining the independent effects of these variables on ET if only simple correlations were used. Correlation does not imply causality and more sophisticated statistical methods need to be used, e.g., multiple regression analysis, to control for confounding factors and to determine the relative influence of Rn and P on ET while considering LAI. This would allow to estimate the unique variance explained by each predictor while holding the others constant.

Furthermore, it is important to consider that the relationships between these variables can be complex and non-linear, and they might be influenced by other factors such as soil moisture, air temperature, humidity, wind speed, and atmospheric pressure. To accurately assess the relationships and the potential for misinterpretation, in our next study we will try to employ a multivariate analysis approach to establish the unique contributions of Rn, P, and other factors on ET, while controlling for the potential influence of other relevant factors in another study based on a more robust analysis.

---

4. Again, in regions where Ec and Ei are the dominant modes of evapotranspiration (Figure 8), it would be valuable to investigate their correlation of ET with LAI compared to Rn and P, after removing the confounding effects.

Reply: It is true that the above-mentioned issue for ET is also applicable for Ec and Ei, and the multivariate analysis can be applied to investigate the response of Ec and Ei to environmental factors.

---

5. Nevertheless, I still believe this could be separate research with robust approach.

Reply: Thank you very much for the suggestion again. We removed this part accordingly and will produce another manuscript based on a more robust analysis.

Additional technical comments:

1. Before highlighting the monthly RMSE, it would be helpful to provide information on the magnitude of monthly ET observed at different flux stations based on in-situ observations. This would allow for a comparison of the magnitude of observed ET with the error represented by the RMSE.

Reply: Thank you very much for the suggestion again, we added a table in Supplementary (Table S1) with the mean value of observed ET.

---

2. It's advisable to maintain analysis with consistent spatial and temporal coverage in the main section, while keeping analyses involving datasets with inconsistent coverage to the supplementary section. This will enhance clarity of the manuscript.

Reply: Thank you very much for the suggestion again, and we focus on the products with spatial and temporal continuity for the analysis in the main text to retain the results and analysis with the same spatial and temporal coverages.

---

3. In Figure 8, it is noted that while the total evapotranspiration (ET) may appear similar across different datasets, the partitioning of ET between datasets is not consistent. This observation is indeed a significant finding. However, the substantial explanation provided does not sufficiently clarify why the datasets differ so much, particularly for GLDAS and MERRA2.

Reply: Thank you for appreciating our findings. The partitioning of ET into its components, such as evaporation from the soil (Es) and transpiration from plants (Ec), can vary significantly among different datasets. This discrepancy in the ET partitioning across different datasets cannot be explained by a single factor, and it is difficult to say which one plays a dominant role as they all contribute in some way to the uncertainty in modelling ET, and may even compensate for each other. In general, these differences stem from factors such as differences in the forcing data, model structure and parameterization, spatial and temporal resolution of the products, and the assumptions embedded in each dataset

- Differences in the forcing data. The forcing data could lead to differences in both the total ET and its components. This explains why GLEAMv35a and GLEAMv35b showed different ET partitioning results, although they are based on exactly the same algorithm. ETMonitor uses GLASS-MODIS data (LAI, FVC, and albedo), PMLV2 use the official MODIS dataset (LAI, albedo, and emissivity). A study by Li et al. (2018) has shown that GLASS LAI is more accurate than MODIS LAI, and MODIS LAI is much lower than GLASS LAI in the eastern TP, which partly explains the relatively lower Ec values by PMLV2 than ETMonitor. Moreover, they also use different meteorological datasets. GLDAS-CLSM uses ERA5 data, while GLDAS-Noah and GLDAS-VIC use GLDAS-2.1 meteorological forcing data as input. A recent study shows that GLDAS-2.1 highly overestimates relative humidity during spring and winter time (Xu et al., 2024), which may lead to lower Es.
- Model structure and parameterization. As a most intuitive example, GLDAS-VIC and GLDAS-Noah share the same forcing data, but the estimated ET partitioning differs significantly. GLDAS-VIC gives a much higher Ec/ET and lower Es/ET, consistent with previous studies. This is most likely due to the weaker soil moisture-ET coupling in the applied physical scheme (Feng et al., 2023). The extremely high Ec/ET ratio is mainly due to the "big leaf" vegetation scheme, which assumes that there are no canopy gaps or exposed soil between plants, so that soil evaporation only occurs in unvegetated areas (Bohn and Vivoni, 2016; Sun et al., 2021). It has also been reported that VIC model, with FVC set to 1 as default value, significantly overestimate Ec and suppresses Es in sparse vegetation types with a true FVC between 0.1 and 0.5 (Schaperow et al., 2021). In contrast, GLDAS-CLSM tends to underestimate the Ec/ET ratio and overestimate Es/ET, possibly due to parameterization issues related to the soil or vegetation resistance, or the non-traditional approach of accounting for subgrid heterogeneity in soil moisture (Feng et al., 2023; Sun et al., 2023). CLSM estimates of ET are adjusted by varying the sub-ranges of soil water availability, i.e. the saturation, transpiration and wilting sub-ranges (where transpiration is shut off), which differs from the continuous soil water stress function used in other models. Some other factors, such as the absence of irrigation and the data assimilation procedure, could also affect the ET partitioning in GLDAS models (Li et al., 2022).
- Calibration of model parameter. Some ET algorithms may have been calibrated and evaluated against different observations, which can lead to variations in the model performance and, consequently, the partitioning of ET. The global ET datasets use default parameters assigned according to land surface characteristics, which are inappropriate for TP and certainly contribute to differences in ET partitioning. Many studies have also highlighted the importance of parameter optimization to reflect the local vegetation and soil properties for modelling ET processes (Xu et al., 2019; Zheng et al., 2022).

- Effects of spatial heterogeneity and resolution. Higher spatial resolution data may more accurately capture details of the local variability in land surface characteristics and associated vapour fluxes in heterogeneous areas (Chen et al., 2019), leading to differences in ET estimates compared to coarser resolution datasets.

Reference:

Bohn, T.J., Vivoni, E.R.: Process-based characterization of evapotranspiration sources over the north american monsoon region. Water Resour. Res., 52 (1), 358–384, https://doi.org/10.1002/2015WR017934. 2016.

Chen, Q., Jia, L., Menenti, M., Hutjes, R., Hu, G., Zheng, C., and Wang, K.: A numerical analysis of aggregation error in evapotranspiration estimates due to heterogeneity of soil moisture and leaf area index, Agric. For. Meteorol., 269–270, 335–350, https://doi.org/10.1016/j.agrformet.2019.02.017, 2019.

Feng, H., Wu, Z., Dong, J., Zhou, J., Brocca, L., He, H.: Transpiration – Soil evaporation partitioning determines inter-model differences in soil moisture and evapotranspiration coupling. Remote Sensing of Environment, 298, https://doi.org/10.1016/j.rse.2023.113841, 2023.

Li X, Lu H, Yu L, Yang K.: Comparison of the Spatial Characteristics of Four Remotely Sensed Leaf Area Index Products over China: Direct Validation and Relative Uncertainties. Remote Sensing. 10(1),148, https://doi.org/10.3390/rs10010148, 2018.

Li, C., Liu, Z., Tu, Z., Shen, J., He, Y., Yang., H.: Assessment of global gridded transpiration products using the extended instrumental variable technique (EIVD). J. Hydrol., 623, https://doi.org/10.1016/j.jhydrol.2023.129880, 2023.

Li, C., Yang, H., Yang, W., Liu, Z., Jia, Y., Li, S., Yang, D.: Error characterization of global land evapotranspiration products: collocation-based approach. J. Hydrol. 612, 128102 https://doi.org/10.1016/j.jhydrol.2022.128102. 2022.

Schaperow, J.R., Li, D., Margulis, S.A., Lettenmaier, D.P.: A near-global, high resolution land surface parameter dataset for the variable infiltration capacity model. Sci. Data, 8 (1), 216. https://doi.org/10.1038/s41597-021-00999-4. 2021.

Sun, R., Duan Q., Wang, J.: Understanding the spatial patterns of evapotranspiration estimates from land surface models over China. J. Hydrol., 595, 126021, https://doi.org/10.1016/j.jhydrol.2021.126021, 2021.

Xu, C., Wang, W., Hu, Y., Liu. Y.: Evaluation of ERA5, ERA5-Land, GLDAS-2.1, and GLEAM potential evapotranspiration data over mainland China. Journal of Hydrology: Regional Studies, 51, https://doi.org/10.1016/j.ejrh.2023.101651, 2024.

Xu, T., Guo, Z., Xia, Y., Ferreira, V.G., Liu, S., Wang, K., Yao, Y., Zhang, X., Zhao, C.: Evaluation of twelve evapotranspiration products from machine learning, remote sensing and land surface models over conterminous United States. J. Hydrol., 578, 124105, https://doi.org/10.1016/j.jhydrol.2019.124105. 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.

Other comments:

Overall, there are numerous instances in the text which exhibits repetition and with typos, with numerous lines conveying similar information and occasionally out of context. Therefore, significant restructuring of the article's text is necessary.

Reply: We apologize for the repetition and typos. We went through the manuscript again to improve it.

For instances:

1. The passage from lines 60-65 highlights the significant uncertainty surrounding evapotranspiration (ET) estimation on the Tibetan Plateau (TP). However, the paragraph falls short in effectively conveying how the present research differs from existing literature. It is evident that this study introduces novelty to the field, particularly through its comprehensive comparison of various ET products with in-situ observations in TP. This contribution warrants greater emphasis in the introduction section.

Reply: Thank you. We emphasized the novelty in the introduction section in the revised manuscript. Previous validations were generally based on either in-situ measurement by the eddy covariance system or the basin-scale ET estimated by water balance method, which represent the surface net water flux at different scales, while these ET products mainly focus on the upward water vapour flux. Recently, Chen et al. (2024) evaluated several ET products with spatial resolutions ranging from 1km to 50km against site-scale eddy covariance observations. It is important to note that the observations from tower-based eddy covariance systems have a very small footprint (approximately several hundred metres depending on weather conditions), and direct comparison of site-scale observations with the coarse-resolution ET products (e.g., 25km) is problematic due to the severe problem of spatial mismatch. In order to increase the credibility of currently available ET products, this study will undertake a more comprehensive evaluation, taking into account both in-situ observations and basin-scale measurements.

Reference:

Chen, X. Yuan, L., Ma, Y., Chen, D., Su, Z., Cao., D.: A doubled increasing trend of evapotranspiration on the Tibetan Plateau. Sci. Bull., https://doi.org/10.1016/j.scib.2024.03.046, 2024.

Liu, H, Xin, X, Su, Z., Zeng, Y., Lian, T., Li, L., Shanshan S.: Hailong Zhang Intercomparison and evaluation of ten global ET products at site and basin scales. J. Hydrol., 617, 128887, https://doi.org/10.1016/j.jhydrol.2022.128887, 2023.

---

2. The final three paragraphs in the introduction section needs to be rephrased for coherence, eliminating redundancy to convey the message clearly. For example, lines 81-86 present the research questions effectively. However, the same information is reiterated in the following paragraph (lines 88-91) within the main objectives, which essentially duplicates the content. This and other redundancy should be streamlined for clarity.

Reply: The questions presented in lines 81-86 serve to identify the scientific problems after the literature review. The last paragraph is intended to clarify the research objectives. To avoid the duplicates, we revised the introduction to ensure that these issues are addressed in a clear and concise manner.

---

3. The classifications in the discussion sections (ET based on PM model, LST-based model, data-driven, and LSM type) seem abrupt as they haven't been introduced earlier in the text. Section 4.1 should be emphasized when analyzing the results, as much of the content there appears redundant in the manuscript, despite its scientific validity. This caveat should be highlighted without unnecessary repetition

Reply: We moved the classifications in the discussion sections to the Section 2.2.2 ET Products when they are first introduced. We also revised the results section to emphasize the discussions relevant to Section 4.1.

On the introduction section, it appears:

These validations were generally based on either in-situ measurement by the eddy covariance system or the basin-scale ET estimated by water balance method, which represent the surface net water flux that integrates different processes (e.g., plant transpiration for the dense vegetation regions, snow sublimation for the dry snow cover periods for the eddy covariance system observations, even condensation when negative latent heat flux occurs),while these ET products mainly focus on the ET (positive upward latent heat flux), which attributes to the validation uncertainty.

While in the section 4.1.1, it appears:

The eddy covariance system observation represents the net water flux integrated across different processes (e.g., plant transpiration in the dense vegetation regions, snow sublimation during the dry snow cover periods, evaporation of canopy-intercepted water when the canopy is wet due to intercepted rainfall). The vaporization process observed by the eddy covariance system depends on the land surface condition, which may vary seasonally and yearly due to factors such as snow/ice, intercepted water, and vegetation. Meanwhile, eddy covariance system observation includes condensation when negative latent heat flux occurs. Remote sensing-based ET products mainly focus on positive ET (positive upward latent heat flux) and omit processes such as condensation.

These two instances basically convey same information. I do agree this is important point to make reader aware about the validation. However, I think the author could be concise about it and avoid unnecessary repetitions.

Reply: In the introduction, we intended to introduce generically the uncertainty caused by the validation method, while in the discussion we focused on the processes captured by tower-based observations, as documented by our findings. To avoid the repetitions, we revised the introduction and discussions accordingly.

---

4. In line 100, it might be more appropriate to adhere to existing climatic regime classifications, such as those based on AI or other established frameworks. Because the term rather "monsoon" is kept here in between arid and humid climate types. So, how "different" is "monsoon" from the humid in these classifications? Or what does that monsoon mean when compared with "arid" and "humid"?

Reply: We agree with you that it is more appropriate to use the existing climatic regime classification and monsoon is not a standard climate type. According to the Köppen classification, there are dry, subtropical, temperate, subpolar and polar climate types in the TP. These climate types are influenced by both westerlies and the Asian monsoon, which is also enhanced by the thermal forcing of the TP (Zhou et al., 2009; Wu et al., 2012; Yang et al., 2014). The aridity index (P/PET) or Budyko dryness ratio (PET/P) are also widely utilized to characterize the aridity level. A recent study has shown that the dryness ratio has a large spatial variability in the TP, from humid climate with dryness ratio less than 0.3 to hyper-arid climate with dryness ratio larger than 3 (Feng et al., 2024). We revised the description accordingly to avoid any ambiguity.

Reference:

Feng, Y.,·Du, S.,·Fraedrich, K., ·Zhang, X., ·Du, M., ·Cheng, W.: Local climate regionalization of the Tibetan Plateau: A data-driven scale-dependent analysis. Theor. Appl. Climatol., https://doi.org/10.1007/s00704-024-04916-8, 2024.
Wu, G., Liu, Y., He, B., Bao, Q., Duan, A., Jin, F.F.: Thermal controls on the Asian summer monsoon. Sci. Rep., 2. http://dx.doi.org/10.1038/srep00404, 2012.
Yang, K., Wu, H., Qin, J., Lin, C., Tang, W., and Chen, Y.: Recent climate changes over the Tibetan Plateau and their impacts on energy and water cycle: A review, Glob. Planet. Change, 112, https://doi.org/10.1016/j.gloplacha.2013.12.001, 2014.
Zhou, X., Zhao, P., Chen, J., Chen, L., Li, W.: Impacts of thermodynamic processes over the Tibetan Plateau on the Northern Hemispheric climate. Sci. China Ser. D Earth Sci. 52, 1679–169,. https://doi.org/10.1007/s11430-009-0194-9, 2009.

---

5. Figure 1: It's not clear what does hashing represent. And for some "red" labels, they are not clear like names around "XG".

Reply: We revised Figure1 to improve clarity.

---

6. Equation (1), please write equations of all metrics or skip even KGE. Please make it coherent.

Reply: We added equations of all metrics.

---

3. Reply to the comments by REVIEWER#3

When I read this paper, I found that the author may not comprehensively review the following papers: Chen, X. et al., 2024. A doubled increasing trend of evapotranspiration on the Tibetan Plateau. Science Bulletin.

Yuan, L. et al., 2024. Long-term monthly 0.05° terrestrial evapotranspiration dataset (1982–2018) for the Tibetan Plateau. Earth Syst. Sci. Data, 16(2): 775-801.

Wang, B.*, Y. Ma*, Z. Su, Y. Wang and W. Ma. Quantifying the evaporation amounts of 75 high-elevation large dimictic lakes on the Tibetan Plateau. Science Advances, 2020, 6, eaay8558.

I agree to the author that they have collected more ET products in this study, but the generally conclusions are not really new compared with previous ET studies on the TP. Hereby, I suggest to focus more on ET components verification and their trends. This part has not been fully investigated by previous publications. The ET trends and annual ET estimation does not deserve more energy on it. This means that the title should be also changed. There are also some water balance ET studies. Hereby, this analysis is also not new. Introduction should really have a in depth review of previous work.

Reply: Thank you for providing the latest publications and constructive suggestions. It is true that ET components are important and not well studied, however we think clarify the total ET and ET trends is also helpful, especially considering that differences in ET components can surely lead to different total ET. Although the previous studies by Chen et al. (2024) and Yuan et al. (2024) have demonstrated the difference of area-averaged ET in the TP, they did not investigate the spatial variability of this difference which actually is very large. Furthermore, previous studies on ET mostly applied the old TP boundary, which only includes the region inside China. Recent studies emphasized the geographic integrity of the TP and a new boundary of TP was applied (Zhang et al., 2013; Zhang et al., 2021), which is larger than the area of the old boundary by 20%. This boundary is more reliable as it is based on geomorphology and formation processes that considers factors such as elevation and watershed boundaries. Hence, the comparison of ET amount and trend is still necessary. We strengthened the materials on ET components in the revised version following the suggestion.

Reference

Chen, X. Yuan, L., Ma, Y., Chen, D., Su, Z., Cao., D.: A doubled increasing trend of evapotranspiration on the Tibetan Plateau. Sci. Bull., https://doi.org/10.1016/j.scib.2024.03.046, 2024.

Wang, B., Y. Ma, Z. Su, Y. Wang and W. Ma. Quantifying the evaporation amounts of 75 high-elevation large dimictic lakes on the Tibetan Plateau. Sci. Adv., 6, https://doi.org/10.1126/sciadv.aay8558, 2020.

Yuan, L., Chen, X., Ma, Y., Han, C., Wang, B., and Ma, W.: Long-term monthly 0.05° terrestrial evapotranspiration dataset (1982–2018) for the Tibetan Plateau, Earth Syst. Sci. Data, 16, 775–801. https://doi.org/10.5194/essd-16-775-2024, 2024.

Zhang, Y., Li, B., Liu, L., Zheng, D: Redetermine the region and boundaries of Tibetan Plateau, Geogr. Res., 40, https://doi.org/10.11821/dlyj020210138, 2021.

Zhang, G., Yao, T., Xie, H., Kang, S., and Lei, Y.: Increased mass over the Tibetan Plateau: From lakes or glaciers?, Geophys. Res. Lett., 40, https://doi.org/10.1002/grl.50462, 2013.

The large uncertainty of ET products over the TP has been reported by Chen et al. 2024 and Yuan et al. 2024. The abstract should more focus on the new scientific questions. Please revise the sentence: there is still significant uncertainty regarding the amount of water vapour released by the TP into the atmosphere, otherwise remove it. The abstract should emphasize the innovative results, not repeated information.

Reply: We agree with you that Chen et al. (2024) and Yuan et al. (2024) have reported the large uncertainty of ET data products. However, their studies did not mention the spatial variability of the uncertainty and the ET components. We revised that sentence to 'there is still considerable uncertainty in the magnitude and spatial variability in the water vapour released from the TP into the atmosphere'.

The response of annual ET to total precipitation, net radiation and leaf area index was explored to present their governing effect on ET, and the results indicated that precipitation effect mostly in the middle and northern TP and net radiation play significant role in the eastern TP. There are many other factors which also influence ET. But they are not included in this paper. In addition, this conclusion is normal as other study. I suggest to remove this weak point from this paper.

Reply: Thank you very much for the suggestion. We agree with you that the response of ET to water, energy, and vegetation variables could be done better by a more comprehensive analysis. And we removed it from current manuscript and prepare another paper on it for a more robust analysis.

TP has been indicated before line 60, hereby please replace "Tibetan Plateau" with "TP".

Reply: We revised it accordingly.

Line 61, Chen et al. 2024 and Yuan et al. 2024 have listed the big differences of annual ET estimation for the TP. It is better to cite their results directly, since they have compared most ET product for the TP region.

Reply: We revised it accordingly.

Line 80, these specificities, are you talking about negative latent heat? If yes, please use negative latent heat directly.

Reply: We revised it accordingly.

Line 81, How accurate are these improved ET products, I understand that this question is already answered at least partly in Chen and Yuan`s publication. The snow/ice sublimation is new in this study. I suggest to revise the second question to: which processes play a significant role to the ET components trend. The third question, I did not find the author provide answers to which factor dominant different ET products. Hereby, the introduction should be rewritten and new scientific questions to be raised. Current formation is quite weak and not comprehensive.

Reply: It is true that previous studies did some evaluation already of new data products. However, it should be noticed that this validation was conducted only at site scale by comparison of eddy covariance observations and ET products. The tower-based eddy covariance observations have a very small footprint (roughly several hundred square meters depending on the weather conditions), and direct comparison of site-scale observations with the coarse-resolution ET product (e.g., 25km), suffers severe problem of spatial mismatch. Hence, we only used site-scale observations to validate the high-resolution ET products (~ 1 km^2). We used basin-scale ET_{wb} to validate both high-resolution and low-resolution ET product. In this sense, our comparison is more robust and comprehensive. To address this aspect, we included this point in our revised version.

For the second question, we intend to figure out how much water is vaporized in TP and which processes (e.g., plant transpiration, soil evaporation, snow/ice sublimation) dominant the total ET. The different components here correspond to different bio-geophysical processes, and this is addressed in Section 4.3 in the revised version. The plan transpiration from plant leaves is mainly controlled by the stomata behaviour in response to environmental conditions, soil evaporation is controlled by soil structure and soil water content, the rainfall interception is determined by canopy morphology and rainfall intensity, and vapour transport after sublimation is determined by near surface boundary layer conditions and the higher latent heat of sublimation. As regards the third question, we removed the section on the response analysis following the suggestion of RC#2.

The first aim of this paper is already investigated by Chen et al. Please change this point or further deep this aim. Actually, there are many attribution studies of TP ET trend. Please review their studies, then make a revision for the num 3 aim.

Reply: The first point is addressed in our reply to the previous comment above.

Line 94, I don`t really agree that pearson correlation analysis can provide us the response of ET to precipitation, Rn and LAI. Indeed, I don`t suggest to include this correlation analysis in this paper. These analysis weaken this paper, it does not benefit to this work.

Reply: We agree with you and removed it in the revised version.

Lines 122, These the sites, please correct this error.

Reply: We revised it accordingly.

Table 2, EB is a daily ET product, not monthly.

Reply: We revised it accordingly.

Figure 4, SEBS should be EB?

Reply: We revised it accordingly.

Please revise ‘in Tibetan Plateau’ to be ‘in the Tibetan Plateau’ or ‘in the TP’.

Reply: We revised it accordingly.

Figure 5, the figure caption should explain what is meaning for different colored bars.

Reply: The global satellite remote sensing-based ET datasets are in dark blue, and the land surface model-based and analysis global ET dataset are in light blue, while the regional ET datasets are in red. We added this explanation in the revised version.

Figure 7, it is quite difficulty to recognize which bar represent which product. Add the product name corresponding each would be more useful. All the trends are ended in 2020? Their curves in figure 7 do not exhibit the same end year.

Reply: We revised Figure 7 to include the products’ name. It is true that different products end in different years, and the end year for the trend analysis depend on the end year of the products. This information was added in the revised version.

Line 322, Among these products, there are nine that provide the main components of ET (Ec, Es, and Ei), it is better to directly say that 'Nine products provide …'.
Reply: We revised it accordingly.

It is important to note that there is no independent reference available for the ET components. I suggest to use the ensemble mean of ET components to check their differences with the ensemble mean. Nine products have provided the ET components. It's a lot. Their ensemble may be close to the truth.
Reply: Thank you for the suggestion. We also notice that averaging properly would not provide good estimates, since the it applies only to random errors, not to the use of the wrong algorithm. According to the results in Section 3.2.3, the median values of the ratio of Ec, Es, and Ei to total ET was 50%, 30%, and 5%. The ET partitioning ratios, i.e. 52%, 43%, and 5% by ETMonitor are the closest ones to the median values. A recent study shows the contributions of Es, Ec, and Ei to total ET are 68.21 %, 23.57 %, and 8.21 %, respectively in the Three Rivers Source of the Tibetan Plateau (Zhuang et al., 2024), which is actually quite close to our estimates.

Furthermore, we discussed the likely reasons causing such differences and the reliability of the partitioning results among different products in the discussion. For example, there are already reports that the overestimation of the Ec/ET ratio by GLDAS-VIC and GLEAM is due to the "big leaf" vegetation scheme assumption that there are no canopy gaps or exposed soil between plants, so soil evaporation only occurs in unvegetated areas (Bohn and Vivoni 2016; Sun et al., 2021; Miralles, et al., 2016). In contrast, GLDAS-CLSM tends to underestimate the Ec/ET ratio and to overestimate Es/ET, possibly due to the parameter problems related to the soil evaporation resistance or vegetation related resistance or the non-traditional approach to consider the subgrid heterogeneity of soil moisture (Feng et al., 2023; Sun et al., 2021). Therefore, to avoid the bias due to these already-known uncertainties, we removed these products in the calculation of the ensemble mean values.

Reference:

Bohn, T.J., Vivoni, E.R.: Process-based characterization of evapotranspiration sources over the north American monsoon region. Water Resour. Res., 52 (1), 358–384, https://doi.org/10.1002/2015WR017934. 2016.
Feng, H., Wu, Z., Dong, J., Zhou, J., Brocca, L., He, H.: Transpiration – Soil evaporation partitioning determines inter-model differences in soil moisture and evapotranspiration coupling. Remote Sensing of Environment, 298, https://doi.org/10.1016/j.rse.2023.113841, 2023.
Sun, R., Duan Q., Wang, J.: Understanding the spatial patterns of evapotranspiration estimates from land surface models over China. J. Hydrol., 595, 126021, https://doi.org/10.1016/j.jhydrol.2021.126021, 2021.
Miralles, D. G., C Jiménez, Jung, M., Michel, D., & D Fernández-Prieto.: The WACMOS-ET project – Part 2: evaluation of global terrestrial evaporation data sets. Hydrology and Earth System Sciences, 20(2), 823-842, https://doi.org/10.5194/hess-20-823-2016. 2016.

Figure 8, the blue color around TP lakes may not reflect the truth. Please check if this is caused by a wrong lake mask.

Reply: We do not use a lake mask here. Please notice that here *Ew* represents the open water evaporation, which actually comes from either lakes or other water bodies, e.g., rivers, snow/ice melt water, flooded pixels, etc.

Figure 9, there are some reports about the annual ET amount for the TP lakes. Please cite these papers to verify Ew shown in the figure. I understand that Wang et al. Science Advance should also provide the Ew estimation for the TP. This study could benefit to verify the result in the figure.

Reply: We added some comparison with these results in the revised version. According to Wang et al. (2020), the total water evaporation is about 29.4 ± 1.2 km^3/yr (≈1111.5 mm/yr) from the 75 lakes in the TP with total area of 26,450 km^2 (accounting for approximately 56.9% of the total lake area in the whole TP), and the total lake evaporation (51.7 ± 2.1 km^3/yr) for all plateau lakes. The total open water evaporation amount from ETMonitor gives a value of 945.3mm/yr for the permanent water surface over the TP. The total water area is 1.29×10^6 km^2 in the TP when seasonal water bodies are taken into account, which is much larger than the permanent water surface. ETMonitor takes into account the seasonality of water surface areas when estimate ET, and the multi-year mean total annual water evaporation in the TP estimated by the ETMonitor is about at 44.4 km^3/yr, which is lower than that given by Wang et al. (2020).

Section 3.3, this part is not really persuasive. A simple correlation is not meaningful, in addition, other factors were not fully considered, such as air temperature, soil moisture, wind speed etc. In addition, the correlation of abnormal should be analyzed, not the original signal. I suggest to remove this section.

Reply: Thank you again, and we removed it accordingly.

"the daily land cover inputted" please revise this.

Reply: We revised it accordingly.
