We thank the reviewer for their work in reviewing again our paper. We have revised our paper according to their suggestion. Please find our replies below.

The reviewers' comments are shown with regular characters. Our replies are shown in bold. Text citations are shown in *italics*: bold characters highlight new additions, while removed text is ~~crossed out~~. Please, note that all text citations refer to the revised manuscript with tracks.

Reviewer-1

In general, the authors responded very well to my review comments. My only remaining concern is related to the use of E-OBS for comparison of daily precipitation. In addition, I have a few minor comments. Line numbers refer to the manuscript with track changes.

Main point
Regarding the interpolated rain gauge dataset E-OBS, you state in your rebuttal: "We confirm that all the datasets, except for E-OBS, were originally available in UTC time. Indeed, we had overlooked the standard time convention used in E-OBS, given its role as a reference product and the relatively small temporal variation across Europe. However, following the reviewer suggestion, we have re-extracted the data and corrected the interpolation accordingly. As a result, EOBS performance has improved slightly, particularly in the coarse-resolution analysis, where regions with larger temporal discrepancies between standard time and UTC are included."

My point is that E-OBS does effectively not provide an accumulation for that day from 0-24 UTC, because of the different measurement intervals of rain gauge networks, especially from different countries, but often also within countries. The time stamp of data from different networks is mostly in UTC, but the measurement interval often differs. This is at least described in Section 2.2 in Overeem et al. (2023), with, for instance, end times of observation of 06:00 UTC and 18:00 UTC. This is important given the many figures where daily precipitation is compared. I expect that this will not be important for bias computations over longer periods. Timing differences will be negligible then. But I do expect that this will influence the results for other metrics.

In addition, what do you mean by "we have re-extracted the data and corrected the interpolation accordingly"? How to achieve this for E-OBS, since this is already a gridded dataset? And does this imply that you took into account the above mentioned differences in measurement interval? Or did you take this into account by selecting the appropriate satellite data? This is difficult to disentangle, because this would require different selections per country, and measurement intervals even differ within countries.

We thank the reviewer for the clarification. Our initial intervention was related to the fact that E-OBS aggregations are available at local time rather than UTC. Therefore, we interpolated E-OBS data to align the UTC time zone, with slight performance improvement in Eastern Europe areas.

Indeed, the issue arising from E-OBS combining data with different temporal aggregation introduces uncertainty in the evaluation of precipitation products—an inherent limitation that cannot be easily resolved. However, since E-OBS is just one of several datasets used in the assessment, and given its widespread use in the scientific community, we have chosen to retain it as a reference dataset while clearly informing readers about the associated uncertainties (Lines 209–214):

"Note that, in some areas, E-OBS observations are derived by aggregating precipitation stations with time intervals that differ from the standard 00–24 period (Overeem et al., 2023), This can potentially cause uncertainty in the assessment of precipitation products using E-OBS. However, considering that E-OBS is not the only dataset used as reference and the importance of assessing HYPER-P against widely used precipitation products, the uncertainty is deemed acceptable."

I invite the authors to provide an explanation to address this concern. Perhaps I'm overlooking something.

I have a few remaining (very) minor comments:
1) L. 97-98: "Radar measurements are hence often combined with rain gauges, to intercalibrate the measurements and obtain more reliable precipitation estimates.". I suggest to use "adjust" instead of "intercalibrate". In radar meteorology, "calibration" is typically used for hardware calibration.

We thank the reviewer for the suggestion. We have changed the term accordingly.

2) The additions in the text on the suitability of radars for precipitation estimation and the reason to not include their data are much better described by (L. 139-141): "Radar measurements are not included in the merging process due to the lack of a global radar dataset and the limited number of weather radars, particularly in developing countries, but they can be used as a valuable reference.". However, I recommend to be more explicit by adding/rewriting that:
- the main purpose of HYPER-P is to provide a "consistently processed" global dataset, especially geared towards application in (nearly) ungauged regions lacking radar coverage. Hence, the main goal is to improve precipitation data in those regions (although the dataset is also expected to add value for regions in Europe further/far away from radar sites and/or rain gauges).
- that some gauge datasets (e.g. EMO, E-OBS or individual timeseries), as well as radar data

have not been incorporated to keep an independent reference for verification above Europe, which is used as a testbed.

We thank the reviewer for the suggestion. We had already pointed out the potential of HYPER-P to be global in its introduction (line 137-139): "The parent datasets are selected based on criteria such as low latency availability or potential, broad spatial coverage, and high accuracy. As a result, the merged product can be made available globally with relatively short latency—approximately one week**"**
We added two sentences to the text to highlight the above-mentioned topics:

Lines 141-143: "*Local (intended as not-global) datasets from radar and gauge were not included in the merging, but they were used as independent references for assessing the performance of the merged product.*"

Line 147-148: "*Specifically, HYPER-P is expected to be particularly valuable for completely or nearly ungauged areas, which lack stable and high-resolution information from ground networks (gauges and/or radars).*"

3) Note that topography information is used in E-OBS, but not for the variable precipitation (I've checked this with an E-OBS expert).

We thank the reviewer for the suggestion. We have corrected the related sentence in the manuscript. Now it states (Lines 572-574):

"*The absence of gauge stations in this area limits the reliability of both EMO and E-OBS products, which are* ***based on*** ~~*forced to estimate precipitation data from sources other than rain-gauges and radar*~~ ***spatial interpolation techniques*** *(*~~*e.g. topography information*~~*, Cornes et al., 2018).****"***

Bibliography

Overeem, A., van den Besselaar, E., van der Schrier, G., Meirink, J. F., van der Plas, E., and Leijnse, H.: EURADCLIM: the European climatological high-resolution gauge-adjusted radar precipitation dataset, Earth Syst. Sci. Data, 15, 1441–1464, https://doi.org/10.5194/essd-15-1441-2023, 2023.