Validation of precipitation reanalysis products for rainfall runoff modelling in Slovenia

The authors present a study evaluating two reanalysis products against a gridded precipitation dataset that was constructed based on ground observational data from errors in input sources and errors propagated into hydrologic simulation across 16 catchments in Slovenia. The paper is within the scope of EGUsphere, however, major revisions must be made before considering publication. Below lists my major concerns regarding the scope of this study, methodology, and results presentation.

Major comments:

1. The scope or motivation of this study is not well articulated. Apparently, reanalysis datasets are not used to solve precipitation, but rather other meteorological variables. Precipitation variable is a by-product. Most of the reanalysis datasets do not directly assimilate precipitation into model chain (Although ERA5 and MERRA2 did assimilate precipitation). For hydrologic utilities - like floods the authors mentioned, a reanalysis dataset is not qualified and should not be a supplement for real observations unless no observations available. In summary, the authors should provide a strong argument prior to their analysis - why use reanalysis for hydrologic modelling since observational datasets exist.

2. In the introduction, I am not convinced by the facts why the authors chose to evaluate reanalysis data. In line 61-63, the authors stated that " mostly data gathered by means of remote sensing technology. ". First, I would like the authors to state why they are not considering other satellite derived products which have been proven to perform better than reanalysis. I am concerned the scope of this study. See my first comment.

3 The authors site a range of literature on evaluation of reanalysis in Lines 52- 93. However, I don't see a summary of what the reviewed studies did not do, that is compensated by the current study and why reanalysis for Slovenian catchments. The authors should make this clear in the introduction.

3. Why is the ARSO-d data considered as the benchmark? There are other observational precipitation datasets available such as IMERG, MSWEP and GSMaP. The authors should provide more information to claim that ARSO-d data is more accurate than reanalysis products. More details of the on this data should be provided, such as data quality, and the data source, validation etc.

4. How is calibration applied to different forcing?

5. Include the location of the catchments in Table 1 (Lat and Lon), catchment characteristics relevant in the formation and dominant hydrological processes

6. In line 250 – 254, the authors state "The simulation period is split-sampled into….."but do not justify the use of Split Sample method in model calibration. Refer to Arsenault et. al., 2018. The hazards of split-sample validation in hydrological model calibration. *Journal of hydrology*, *566*, pp.346-362.

7. Line 268 – 273 Please check the font and line spacing

8. There are many recent references focusing on evaluation studies and application, you may find it helpful.
Wanzala, M. A., Ficchi, A., Cloke, H. L., Stephens, E. M., Badjana, H. M., & Lavers, D. A. (2022). Assessment of global reanalysis precipitation for hydrological modelling in data-scarce regions: A case study of Kenya. *Journal of Hydrology: Regional Studies*, *41*, 101105.

Tang, G., Clark, M. P., Papalexiou, S. M., Ma, Z., & Hong, Y. (2020). Have satellite precipitation products improved over last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets. *Remote sensing of environment*, *240*, 111697.