Response to RC2

Review for " Saudi Rainfall (SaRa): Hourly 0.1° Gridded Rainfall (1979–Present) for Saudi Arabia via Machine Learning Fusion of Satellite and Model Data" by Wang et al. submitted to EGUsphere (MS No.: egusphere-2025-254).

General comments:

The authors introduce *Saudi Rainfall (SaRa)*, a gridded precipitation product for the Arabian Peninsula developed using Machine Learning (ML) techniques. They clearly present the motivation behind the development of such a dataset, describe the procedures used to generate the SaRa product, and evaluate its performance. By leveraging a large amount of available gauge-based and gridded datasets, the authors produce a new dataset that shows improved performance compared to existing products—particularly in areas with sparse station observations and in the dominantly arid regions of the Arabian Peninsula.

This work makes a valuable contribution to the data community and enhances scientific understanding of precipitation patterns in data-scarce, arid environments. The overall quality of the manuscript is good, with well-cited references and generally clear writing.

We thank the reviewer for their useful feedback.

However, there is still room for further improvement. In particular, I would like to raise two main concerns:

1. Limitations of Machine Learning: What are the potential limitations, challenges and sources of error introduced by using Machine Learning techniques in generating this dataset? A discussion on uncertainties and biases associated with ML itself would strengthen the paper.

This is a great comment. We agree that ML is subject to some limitations, but as our independent validation shows, our ML model outperforms all other gridded datasets for all metrics. Nevertheless, to address this comment, we will add the following text to section 3.1:

“A key limitation of ML-based P estimation is poor generalizability; models often fail in regions lacking training data (Xu et al., 2024). To assess whether this applies to our models, we analyzed KGE values of the evaluation stations as a function of distance to the nearest training station (Figure 4). The results show no clear decline in KGE with increasing distance, indicating satisfactory spatial generalizability. Another common criticism is the “black-box" nature of ML models, which limits interpretability.

To improve transparency, we computed predictor importances for all four submodels of model_01 (Table 5). IMERG-L V07 consistently ranked higher than GSMaP-MVK V8 in importance, indicating a model preference for IMERG, which aligns with its superior validation performance (Table 4). ERA5 was the most important predictor for the daily submodel, whereas IMERG dominated in the 3-hourly and hourly submodels. This likely reflects the advantage of observational datasets like IMERG in accurately capturing event timing. Static predictors were overall much less important than dynamic ones. Among the static predictors, abs_lat and lat had the highest importances, likely reflecting the latitudinal dependence of P product performance observed in global evaluations (e.g., Beck et al., 2017)."

A third challenge is that ML methods inherently tend to underestimate extreme precipitation events due to regression toward the mean. This is discussed in sections 2.5 and 3.1, and was addressed in SaRa by the inclusion of submodel 3, which corrects the CDF.

2. Broader Impact and Global Appeal: What is the relevance of this work beyond the Arabian Peninsula? Discussing the broader applicability of the methodology and insights would enhance the global significance of the study.

This is a great point! We believe the algorithm we used in this paper could also be applied globally. We are actually also working on the new version of MSWEP (Multi-Source Weighted-Ensemble Precipitation) by using the SaRa algorithm as the base. To highlight the broader applicability, we have the following text:

"The SaRa dataset, available for use and distribution at www.gloh2o.org/sara, equips researchers, professionals, and policymakers with the tools needed to tackle pressing environmental and socio-economic challenges in Saudi Arabia, and serves as a potential framework for filling this data gap in other arid and dryland regions."

We will not mention MSWEP in the current context as it is still in development.

In addition, I suggest the authors consider the following points:

Include a study area map: Add a map of the Arabian Peninsula showing the region's topography and its location in a global context. This would help orient readers unfamiliar with the area.

We will include a new figure to indicate the study area

Describe ML Challenges: Provide a more detailed discussion of the challenges and limitations in implementing ML for P data generation.

We will add a detailed paragraph discussing the main limitations of ML for precipitation estimation. Please see our response to the reviewer's first point.

Discuss Practical Applications: Expand the discussion to highlight potential applications of the dataset, such as its use in flash flood risk mitigation, water resource management, or climate-related decision-making in arid regions.

Thank you. We have the following text in the paper which we believe addresses your comment:

"The SaRa dataset, available for use and distribution at www.gloh2o.org/sara, equips researchers, professionals, and policymakers with the tools needed to tackle pressing environmental and socio-economic challenges in Saudi Arabia, and serves as a potential framework for filling this data gap in other arid and dryland regions. The product delivers a high-resolution, near real-time resource designed to support a diverse range of applications, including water resource management, hydrological modeling, agricultural planning, disaster risk reduction, and climate studies."

References

Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., Van Dijk, A. I., Weedon, G. P., ... & Wood, E. F. (2017). Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. Hydrology and Earth System Sciences, 21(12), 6201-6217.

Xu, Y., Tang, G., Li, L., & Wan, W. (2024). Multi-source precipitation estimation using machine learning: Clarification and benchmarking. Journal of Hydrology, 635, 131195.