

## **Referee comments on “Daily ensemble river discharge reforecasts and real-time forecasts from the operational Global Flood Awareness System” by Harrigan et al., 2020 in HESSD.**

This paper shows recent progress in global river forecasts from the Glofas modelling system. Such data is indeed very useful and appreciated by many users at the global scale, especially by low- and middle-income countries who might not have access to their own river-forecast system. Accordingly, it is very important to evaluate such systems scientifically before launching them operationally.

The paper gives a very good overview of a river-discharge forecast system, which is indeed valuable for the scientific community to learn more about, as such systems are dedicated to national/international institutes with advanced IT infrastructure and operational production.

My main concern with this paper is that I miss a scientific question and the story of what kind of new scientific knowledge we have learnt from using the forecasting system and evaluation method described.

The Glofas model and forecasting system has been described before in the scientific literature and the focus of this paper seem to be that the results are now part of the climate service C3S, but this is hardly a scientific finding. New datasets should rather be published in ESSD, in which Glofas results have already been published. Likewise, the methods used for forecast evaluation are standard and has been published before. For publications in HESS I expect a more scientific analysis of the results and conclusions about new knowledge from the identified scientific achievements with impact on our understanding of Hydrology or Earth Systems. Right now, I have difficulties to find a clear take-home message in the current version of this paper. It is very descriptive and less analytic.

I therefore suggest to find a scientific angle from current discussions in the research community and tell the story of the results from that perspective.

Interesting scientific questions could for instance be:

- On the method side: How should we evaluate forecasts – what metrics are there, how do they compare and what does different metrics contribute in understanding/reliability for the user community and research community, respectively?
- Could the metrics presented (and argued for?) in this paper be compared with other metrics, to show their excellence and benefits to users/scientists? (is there a take-home message or guide-lines to the scientific community from using a specific metric/evaluation method compared to another?) What are the options?
- On the understanding of hydrology: what are the attributes for catchments/regions with high or low skills in forecasting? i.e. which processes do we need to learn more about to improve the quality of river-discharge forecasts?
- How does different global river-discharge forecast systems compare to each other? Can we learn from different model setups and elaborations on procedures, process descriptions or geophysical representation?

Please, find some detailed comments on current manuscript below. Apologies for mentioning my own work, but I am very eager to start comparing model results at the global scale soon. ☺

### **Introduction**

Line 31: Reference Blöschl, et al. 2019 does not evaluate risks or hazards.

Line 37: also note the global and continental scale forecasting based on sharing the world-wide HYPE model:

*Arheimer, B., Pimentel, R., Isberg, K., Crochemore, L., Andersson, J. C. M., Hasan, A., and Pineda, L., 2020. Global catchment modelling using World-Wide HYPE (WWH), open data and stepwise parameter estimation, Hydrol. Earth Syst. Sci. 24, 535–559, <https://doi.org/10.5194/hess-24-535-2020>*

Line 60-70: In fact, global river forecasts and reforecasts are also available at <https://hypeweb.smhi.se/> where the user can subscribe to seasonal forecasts with monthly data. In addition, 1-10 days forecasts with thresholds based on return periods of high flows can be ordered at <https://hypeweb.smhi.se/water-services/data-delivery-services/>

Section 2. even though Glofas has been evaluated against observed river flow in previous publications, it would be helpful to include such information about model performance vs absolute values also. For instance, Fig 4 could also include colors of KGE performance (modelled values vs observed values) in the circles showing location of gauges. This would make this figure much more informative and help the reader a lot to judge model performance. Please, check the color coding in Arheimer et al., so the overall pattern of model performance could be compared. Please, also mention median KGE at global scale (no you only say that it was skillful, which is very vague).

Section 3: please start with some sentences summarizing the evaluation concept – e.g. that you use scores with met. model vs observed met. model (“a perfect weather model”) and correlation with observations. It would also be interesting for many users to actually see some scores to absolute values as well – or at least to discuss the difficulties here.

Section 4: the Glofas results could be compared with results from another model, using the same metrics across Europe, presented by:

*Pechlivanidis, I. G., Crochemore, L., Rosberg, J., & Bosshard, T. (2020). What are the key drivers controlling the quality of seasonal streamflow forecasts? Water Resources Research, 56, e2019WR026987. <https://doi.org/10.1029/2019WR026987>*

To further explore and evaluate the added value of the Glofas system, it could also be compared to warning issued by National forecast services for specific regions or countries, or to soft information from new items reporting floods, to check if the alerts actually captured something real.

Line 265: Attribution is also discussed in the above-mentioned paper. It is another interesting scientific analysis, which deserves much more attention – also in this global study of model performance. Such an analysis would make this paper much more scientifically interesting.

I am looking forward to read a new more elaborated version of this paper, with a scientific discussion linked to the methodological description.