

EGU24-9172, updated on 13 Jul 2024

<https://doi.org/10.5194/egusphere-egu24-9172>

EGU General Assembly 2024

© Author(s) 2024. This work is distributed under the Creative Commons Attribution 4.0 License.



Forecast Verification in Operational Hydrological Forecasting: A Detailed Benchmark Analysis for EFAS

Maliko Tanguy, Shaun Harrigan, Corentin Carton De Wiart, Michel Wortmann, Thomas Haiden, and Christel Prudhomme

European Centre for Medium-Range Weather Forecasts (ECMWF), Forecast Department, Reading, United Kingdom
(maliko.tanguy@ecmwf.int)

Operational hydrological forecasting systems play a vital role for effective decision-making in disaster management and resource planning. This study addresses the critical need to conduct a fair and realistic assessment of the skill in one such system: the European Flood Awareness System (EFAS). As an integral component of the Copernicus Emergency Management Service (CEMS), EFAS undergoes monthly forecast verification, spanning lead times from 6 hours to 10 days. This verification process provides users with a valuable insight into the system's performance in predicting streamflow for the preceding month.

Motivated by the importance of providing stakeholders with trustworthy information, our research focuses on a thorough examination of benchmark forecasts to evaluate EFAS performance. The choice of benchmark forecasts significantly influences the perceived accuracy of the system, and using benchmarks that are too easy to beat can lead to artificially inflated skill. Therefore, the primary objective of this work is to pinpoint the most suitable benchmark, serving as a robust reference for assessing the true capabilities of EFAS. This will then feed into the development of a 'headline score' which is a unique value of a key metric representative of a geographical domain that enables to track performance evolution.

The study employs various benchmark forecasts, including persistence forecast, climatology, and the previous day's forecast, using the Continuous Ranked Probability Skill Score (CRPSS) for skill assessment. Expanding on previous findings that identified persistence forecast as the most suitable for short lead times and climatology for longer lead times, this work refines and extends these results. We specifically examine the influence of catchment characteristics on the selection of the optimal benchmark at different lead times for operational forecasting evaluation. By uncovering the most robust benchmark, our study contributes to a more accurate understanding of EFAS capabilities, ultimately enhancing the overall performance assessment of EFAS. The nuanced insights gained from this focused examination serve as a step toward refining the methodology and criteria employed to develop new 'headline scores', instrumental in evaluating the evolution of the system's forecasting skill.