

Author's Response for HESS Preprint EGUSPHERE-2025-4978

Dear Micha Werner,

we thank you for the review process. Please find attached with this document our new revised manuscript and a manuscript with track changes, both incorporating the suggested changes from the review process. Below here, we summarize the major changes made in the new manuscript:

1. Introduction

- We reformulated the sentences according to the suggestions of the referees

2. Data and Methods

- We reformulated the sentences according to the suggestions of the referees
- We justify using the Forecast Baseline experiment as lower benchmark following Seibert et al. (2018)
- the Cross Domain Evaluation Experiment is explained in more detail, with Table 1 showing the differences in the forcing data sets
- We introduce the new chapter “2.3 Quantifying Domain Shift and Its Physiographic Controls”, expanding the analysis by looking into spatial differences of forecast improvements

3. Results and Discussion

- We have revised the phrasing throughout the section in response to the referees' suggestions and added clarifying explanations where needed to improve the accessibility and precision of the presentation.
- The new chapter “3.1.1 Linking Meteorological Domain Shift to Topographical Features” builds on the Cross Domain Evaluation Experiment and quantifies the distributional differences of the input domains with regard to catchment characteristics.
- In Former Figure 4 (now Figure 5), the Delta NSE value in the figure was corrected from 0.3 to 0.24
- In response to the referees' comments on basin-scale variability and the role of catchment attributes in modulating model skill, we add a new Section “3.6 Physiographic Controls on Model Skill” to the revised manuscript. In this section, we compute the per-basin Δ NSE between each experimental configuration and the Baseline Forecast, visualized as a heatmap with basins sorted by their mean Δ NSE in ascending order, providing a clear picture of which catchments consistently benefit from the proposed strategies and to what extent. We further compute the Spearman correlation between the per-basin Δ NSE and 33 static catchment attributes across all experimental configurations to identify which physiographic characteristics are associated with the largest forecast skill improvements, also shown in a heatmap. We further added a map, showing the 1-Wasserstein distance between Forecast and Reanalysis inputs and the NSE values of the Baseline Forecast.

- The Limitations and Future Directions section has been generously expanded to include referee comments
4. Conclusion and Outlook
 - We have revised the phrasing throughout the section in response to the referees' suggestions and added clarifying explanations where needed to improve the accessibility and precision of the presentation.
 - Paragraphs drawing conclusions of the new chapter 3.6 were added
 5. Appendix A
 - Table A1 caption was extended
 6. Appendix B
 - We found an error in the y-axis label of the Solar Radiation violin plot. The plot was changed.
 - A description of the violin plots is now below the plot
 7. Appendix E
 - Included Figure E2. Sequential Forecast LSTM
 8. Appendix F
 - new appendix, containing a table with Summary Statistics of Model Performance Across All Experimental Configurations
 9. References
 - Included the missing url for the ECMWF data source

Despite the substantial additions and revisions described above, the core findings and conclusions of the paper remain unchanged. We are confident that the modifications comprehensively address all points raised by both referees and have resulted in a significantly strengthened manuscript. We are grateful to both referees for their thorough and constructive engagement with our work, which has been invaluable in improving the quality and clarity of the paper.

Best regards,

Oliver Konold