



How skilful are AI-based forecasts of 2023 Indian summer monsoon precipitation?

Mehzooz Nizar¹, Reinhard Schiemann^{1,2}, Andrew G Turner^{1,2}, Kieran Hunt^{1,2}, and Steffen Tietsche³

¹Department of Meteorology, University of Reading, Reading, United Kingdom of Great Britain – England, Scotland, Wales (m.nizar@pgr.reading.ac.uk)

²National Centre for Atmospheric Science, University of Reading, Reading, United Kingdom of Great Britain – England, Scotland, Wales

³European Centre for Medium-Range Weather Forecasts, Bonn, Germany

India relies on agriculture as one of its main sources of income. Therefore, reliable prediction of Indian summer monsoon rainfall is crucial to the country's policy making and development of crop management strategies. The recent development of global AI Weather Prediction (AIWP) models has revolutionized weather forecasting. Owing to the very recent emergence of AIWP models, their performance in simulating the Indian monsoon system is still insufficiently explored. In this study, we verify the precipitation forecast skill of AIWP models GraphCast and FuXi at a lead time of 1-9 days during Indian summer monsoon 2023 and compare their performance to the physics-based model ECMWF IFS-HRES (IFS). Satellite-derived precipitation dataset IMERG is used as the ground truth to verify precipitation along with ERA5 precipitation. Root mean squared error (RMSE), pattern correlation coefficient (PCC), structure (S)-amplitude (A)-location error (L) and stable equitable error in probability space (SEEPS) were the metrics used to evaluate the models.

A number of case studies, seasonal and intra-seasonal characteristics of precipitation forecast at various lead times were analysed during June-September 2023. The case studies reveal that the AIWP models have lower RMSE and higher PCC than IFS in general, while the AIWP models smoothen (positive S error) precipitation at longer leads. FuXi consistently underestimates precipitation (negative A error) in the case studies. Analysing the daily mean rainfall for the country as a whole and the precipitation bias at a lead time of 5 days, it is confirmed that FuXi shows a

systematic dry bias in forecasting monsoon rainfall. Non-parametric statistical tests were conducted to decide which model performs the best at each metric in forecasting the entire season at various lead times. It is found that FuXi consistently achieved the lowest RMSE, IFS delivered the best S, and GraphCast recorded the smallest SEEPS score at a lead time of 1, 5 and 9 days while no model shows a significant advantage in PCC, A and L. It was also seen that AIWP models outperformed IFS in RMSE and PCC while AIWP models have larger S error than IFS corroborating the findings of case studies. FuXi scored the largest A error across all lead times. The loss functions used to train AIWP models directly penalise point-wise errors, which likely explains their RMSE advantage over IFS.

These results show us that even though AIWP models have good overall accuracy and correlation with observed precipitation, exhibits a lack of realism in capturing the spatial distribution and the intensity of precipitation. Also, model skill is metric dependent and choosing between an AIWP or physics-based model should hinge on the forecaster's priority.