EGU22-11063
https://doi.org/10.5194/egusphere-egu22-11063
EGU General Assembly 2022

ML-based Probabilistic Prediction of 2m Temperature and Total Precipitation

Mohamed Akram Zaytar[1], Bianca Zadrozny[2], Campbell Watson[3], Daniel Salles Civitarese[2], Etienne Eben Vos[4], Thabang Michael Mathonsi[4], and Thabang Lukhetho Mashinini[4]
[1]IBM Research, Kenya (mohamed.akram.zaytar@ibm.com)
[2]IBM Research, Brazil
[3]IBM Research, US
[4]IBM Research, South Africa

The need to build reliable weather forecasting systems for subseasonal to seasonal (S2S) timescales has never been greater as the world continues to experience increased numbers of extreme weather events. This study addresses the skill gap between numerical weather prediction (NWP) and seasonal forecasting by proposing a daily probabilistic forecast model that predicts 2-meter temperature and total precipitation on a global scale. It combines multimodal data (e.g., physics-based ensembles, climate modes, recent climatology) into feature vectors given as inputs to three ML models: Extreme Gradient Boosting, U-Net, and Natural Gradient Boosting. We use Bayesian hyperparameter search and leave-one-year-out RPSS cross-validated scores to accelerate learning and ensure generalizability. Our method consistently outperforms both ECMWF 46-day forecasts and climatology. We find that augmenting physics-based issued forecasts with other sources of predictability greatly improves the performance of the underlying dynamical models. We hope that by improving the physics-based probabilistic forecasts, we will unlock skill in predicting climate extremes-oriented indices. Subsequent representation learning models may be trained to efficiently navigate the ensembles' uncertainty space and estimate the likelihood of extreme events.