EGU24-2897, updated on 08 May 2024
https://doi.org/10.5194/egusphere-egu24-2897
EGU General Assembly 2024

Postprocessing East African rainfall forecasts using a generative machine learning model

Bobby Antonio[1,2], Andrew McRae[1], Dave MacLeod[3], Fenwick Cooper[1], John Marsham[4], Laurence Aitchison[5], Tim Palmer[1], and Peter Watson[2]

[1]Department of Physics, University of Oxford, Oxford, UK (bobby.antonio@physics.ox.ac.uk)
[2]School of Geographical Sciences, University of Bristol, Bristol, UK
[3]School of Earth and Environment Sciences, University of Cardiff, Cardiff, UK
[4]School of Earth and Environment, Univeristy of Leeds, Leeds, UK
[5]Machine Learning and Computational Neuroscience Unit, University of Bristol, UK

Existing weather models are known to have poor skill at forecasting rainfall over East Africa, where there are regular threats of drought and floods. Improved precipitation forecasts could reduce the effects of these extreme weather events and provide significant socioeconomic benefits to the region. We present a novel machine learning based method to improve precipitation forecasts in East Africa, using postprocessing based on a conditional generative adversarial network (cGAN). This addresses the challenge of realistically representing tropical rainfall in this region, where convection dominates and is poorly simulated in conventional global forecast models. We postprocess hourly forecasts made by the European Centre for Medium-Range Weather Forecasts Integrated Forecast System at 6-18h lead times, at 0.1^{o} resolution. We combine the cGAN predictions with a novel neighbourhood version of quantile mapping, to integrate the strengths of both machine learning and conventional postprocessing. Our results indicate that the cGAN substantially improves the diurnal cycle of rainfall, and improves rainfall predictions up to the 99.9th percentile of rainfall. This improvement persists when evaluating against the 2018 March-May season, which had extremely high rainfall, indicating that the approach has some ability to generalise to more extreme conditions. We explore the potential for the cGAN to produce probabilistic forecasts and find that the spread of this ensemble broadly reflects the predictability of the observations, but is also characterised by a mixture of under- and over-dispersion. Overall our results demonstrate how the strengths of machine learning and conventional postprocessing methods can be combined, and illuminate what benefits machine learning approaches can bring to this region.