EGU24-19266, updated on 27 Jul 2024
https://doi.org/10.5194/egusphere-egu24-19266
EGU General Assembly 2024

Application of novel generative diffusion models to precipitation downscaling

Alex Saoulis[1,2], Sebastian Moraga[1], Natalie Lord[1], Peter Uhe[1], and Nans Addor[1,3]
[1]Fathom, Bristol, United Kingdom
[2]University College London, Physics & Astronomy / Earth Sciences, United Kingdom
[3]University of Exeter, United Kingdom

Machine Learning (ML) is playing an increasingly valuable role in statistical downscaling. Capable of leveraging complex, non-linear relationships latent in the training data, the community has demonstrated significant potential for ML to learn a downscaling mapping. Following the perfect-prognosis (PP) approach, ML models can be trained on historical reanalysis data to learn a relationship between coarse predictors and higher resolution (i.e. downscaled) predictands. Once trained, the models can then be evaluated on general circulation model (GCM) outputs to generate regional downscaled results. Due to the relatively low computational cost of training and utilising these models, they can be used to efficiently downscale large ensembles of climate models over regional to global domains.

This work employs a novel diffusion algorithm to downscale climate data. Diffusion models have proven highly successful in applications such as natural image generation and super-resolution (the natural image analogue to climate downscaling). Diffusion models have been shown to significantly outperform earlier generative ML models such as Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs); they can produce highly diverse samples, emulate fine details with high fidelity, and exhibit much more stable training than alternative ML models.

This work trains and evaluates diffusion models on the Multi-Source Weighted-Ensemble Precipitation (MSWEP) observational dataset over the Colorado River Basin (USA). High resolution (10km x 10km) MSWEP fields are artificially coarsened to generate training data. Once trained, the models are applied to bias-corrected climate model outputs to evaluate their ability to generate realistic downscaled precipitation fields. Performance is compared with several benchmarks, including classical regression techniques as well as alternative ML models.