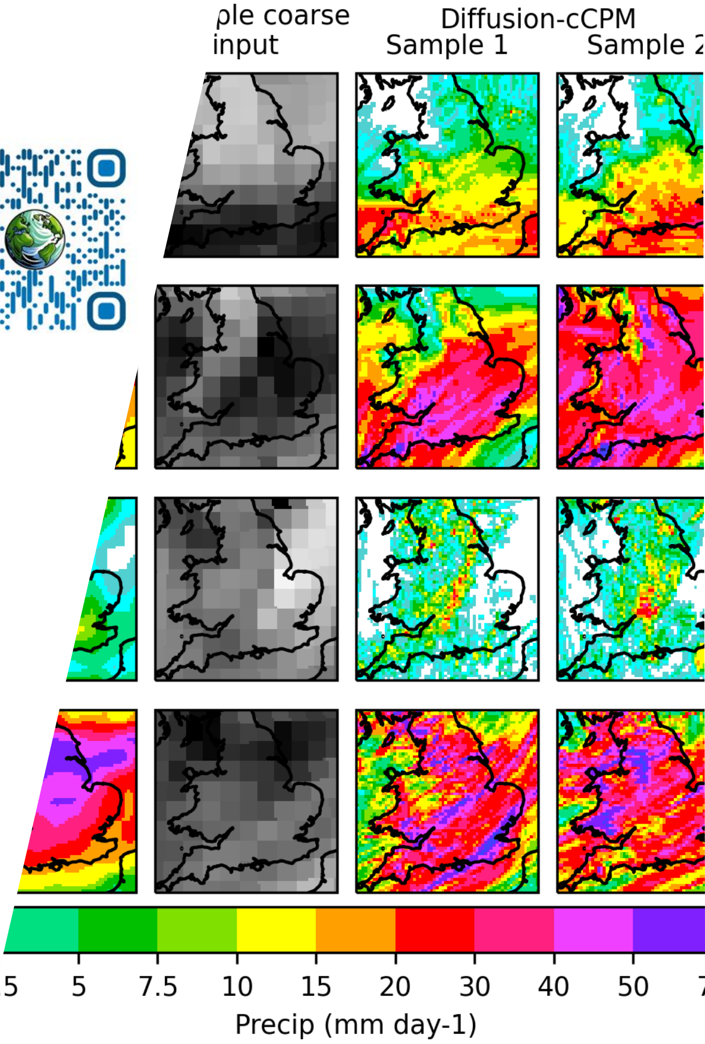
Machine learning emulation of precipitation from km-scale regional climate simulations using a diffusion model

Henry Addison

Supervisors: Peter Watson and Laurence Aitchison

with Elizabeth Kendon and Suman Ravuri



University of BRISTOL

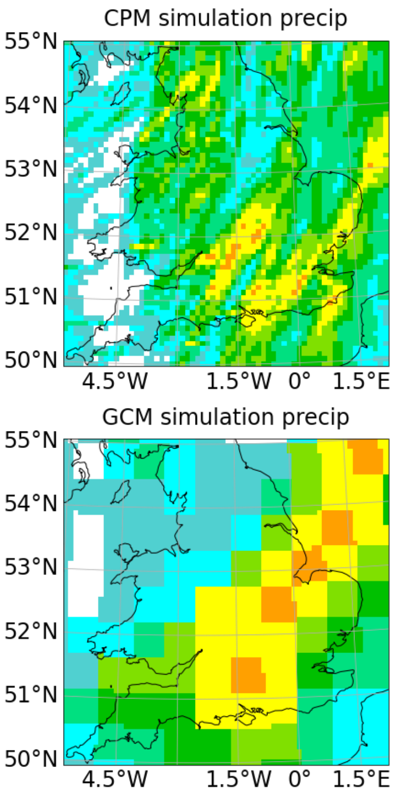
INTERACTIVEAI
Centre for Doctoral Training

UK Research and Innovation

bristol.ac.uk

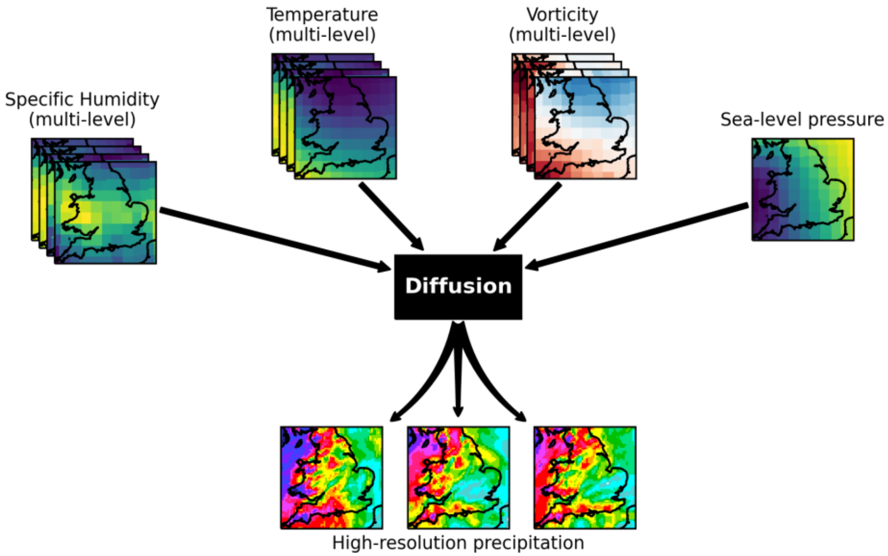
The problem

- High-resolution climate simulations are expensive
- **Can we use probabilistic ML methods to emulate a hi-res, CPM simulator using coarse (GCM) climate variables?**
- Met Office UKCP Local and UKCP18
  - **Low resolution:** Global Climate Model (**GCM**) @ **60km**
  - **High resolution:** Convection-Permitting Model (**CPM**) @ 2.2km (using **8.8km**)
  - Daily frequency



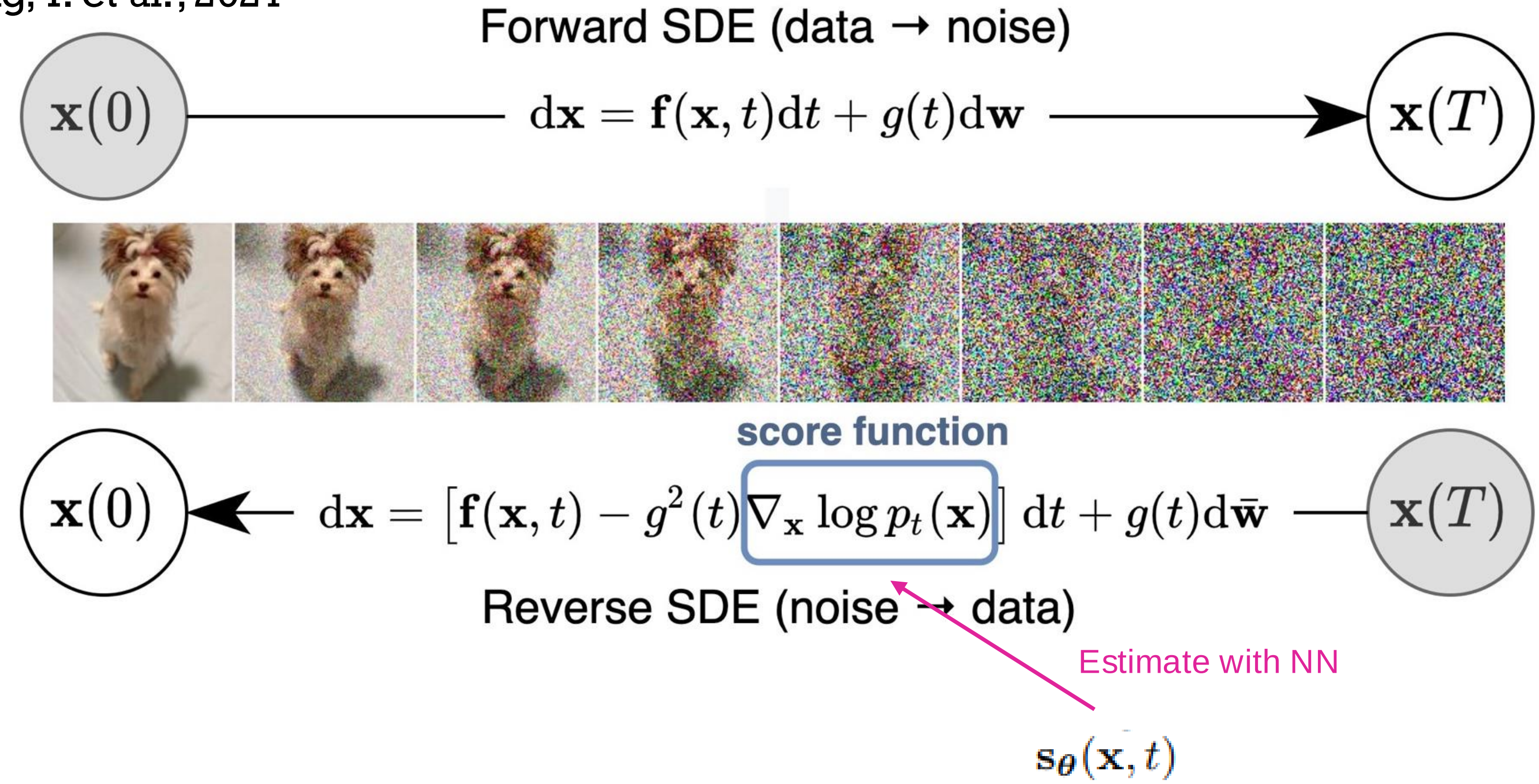
Approach

- Training: coarsened CPM variables → hi-res CPM precip
- Evaluating: coarsened CPM variables or GCM variables → hi-res CPM-like precip
- Use variables which are well-represented in GCMs and cause rainfall



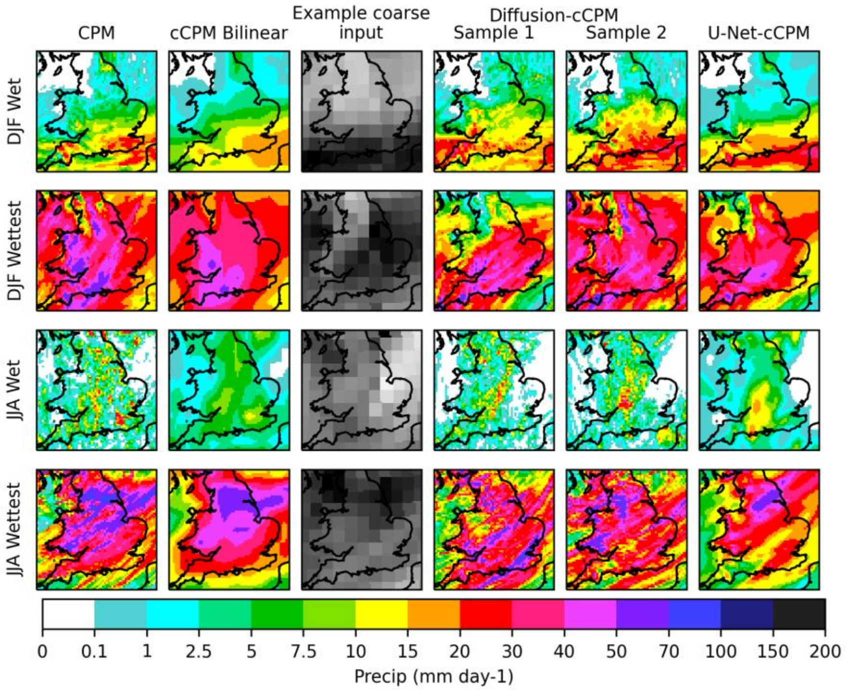
Score-Based Generative Modeling through Stochastic Differential Equations

Song, Y. et al., 2021

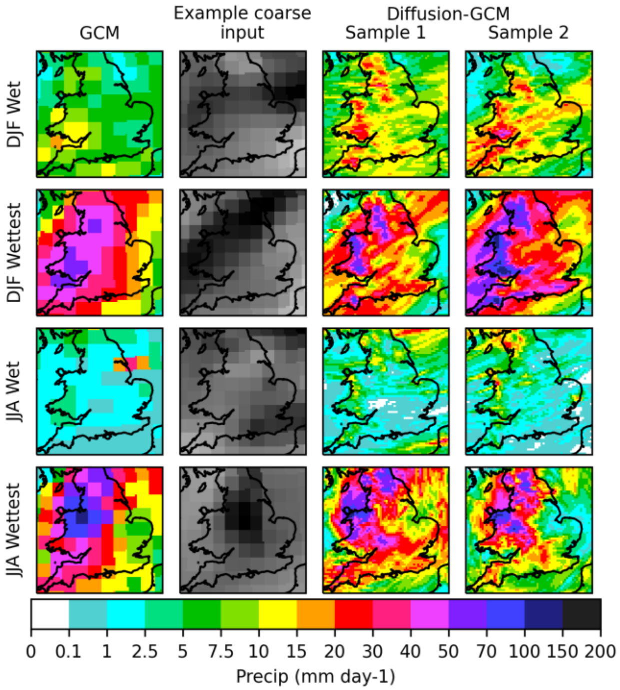


Samples

Coarsened CPM → 8.8km CPM rainfall

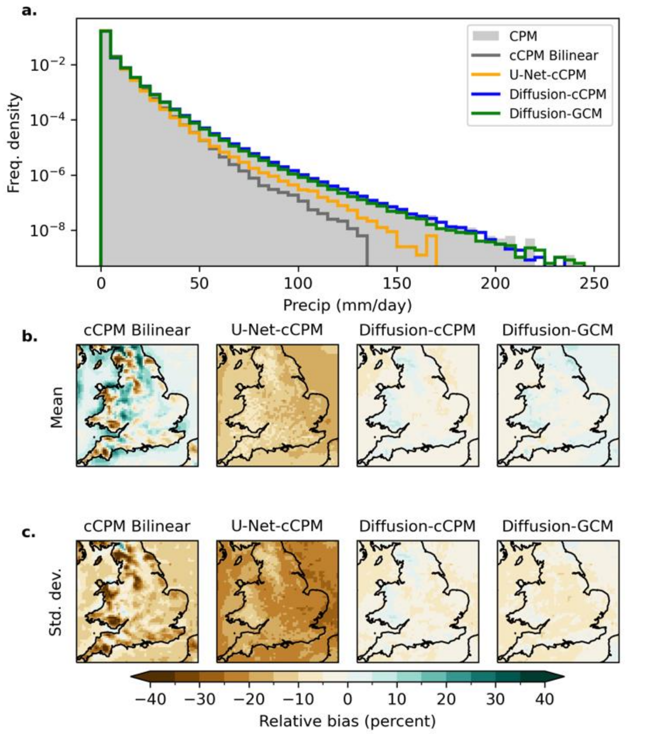


GCM → 8.8km CPM rainfall



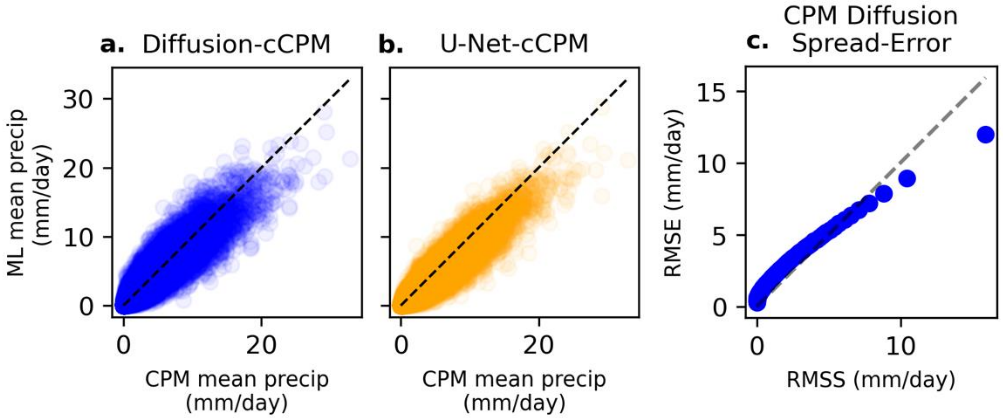
Distribution of precipitation

- Both coarsened CPM and GCM inputs to emulator produce samples with grid-box distribution matching CPM precipitation
- Little bias in mean or std dev across the domain



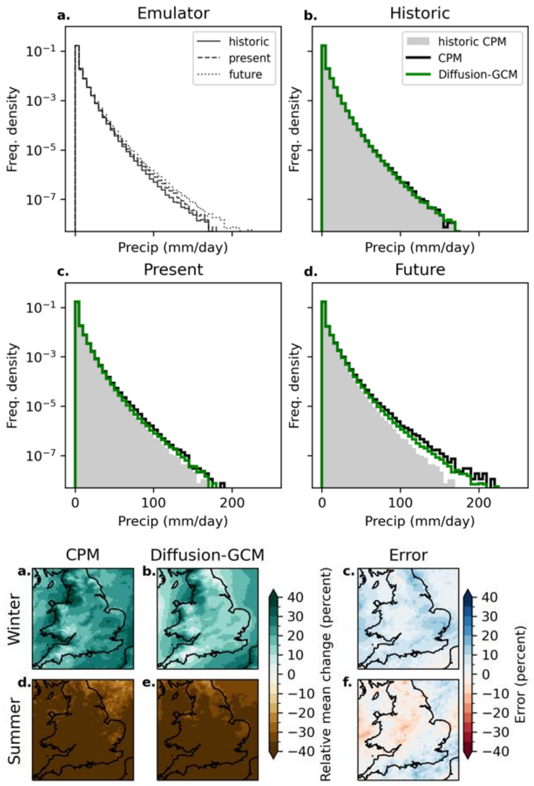
Dependence on input and uncertainty

- Spread between samples matches error in expected way
- There may be more spread overall compared to deterministic U-Net, this represents the stochastic component of rainfall well



Climate change signal

- Captures shift in distribution from historic period (1981-2000) through present (2021-2040) to future (2061-2080)
- Captures some of the drier summers and wetter winters signal



Summary

Using SOTA diffusion model to emulate Met Office’s hi-res UK climate model

- Reproduces realistic spatial structure and variability of rainfall
- Good match in distribution and well-calibrated stochastic component, even using GCM-based inputs
- Captures most of the 21st century climate change signal in the CPM

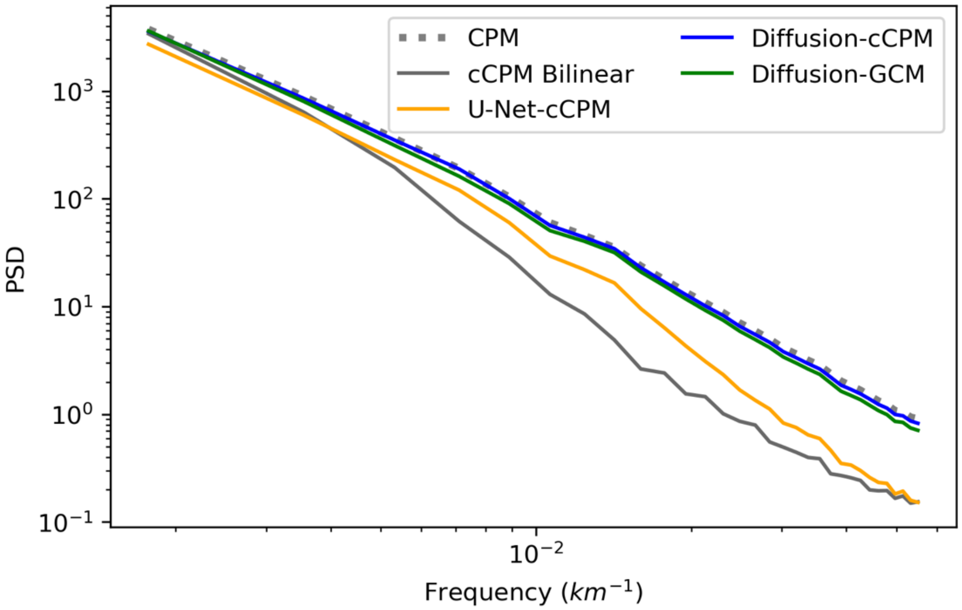
Any questions or suggestions?

henry.addison@bristol.ac.uk

Or read the pre-print

Spatial structure

- RAPSD show emulator samp contain similar amounts of variability across the range of different spatial scales



Future work

- Multivariate predictions
- Generalize to (large ensembles of) other climate models, time periods, locations
- More extreme Extremes: 1-in-100 years
- Sub-daily frequency and temporal sequences (video)
- Flood modelling applications

Key references

- Gutiérrez, J. M. et al. 2019. 'An intercomparison of a large ensemble of statistical downscaling methods over Europe: Results from the VALUE perfect predictor cross-validation experiment', *International Journal of Climatology, 39: 3750-85.*
- Kendon, E. J. et al. 2021. 'Update to the UKCP Local (2.2km) projections: Science report', Met Office Hadley Centre, Exeter, UK.
- Luo, C. 2022. "Understanding diffusion models: A unified perspective." arXiv preprint arXiv:2208.11970.
- Murphy, J., G. et al. 2018. 'UKCP18 land projections: science report', Met Office Hadley Centre, Exeter, UK.
- Song, Y. et al., 2021. Score-Based Generative Modeling through Stochastic Differential Equations', *International Conference on Learning Representations*.
- Vandal, T. et al. 2018. 'Intercomparison of machine learning methods for statistical downscaling: the case of daily and extreme precipitation', *Theoretical and Applied Climatology, 137: 557-70.*