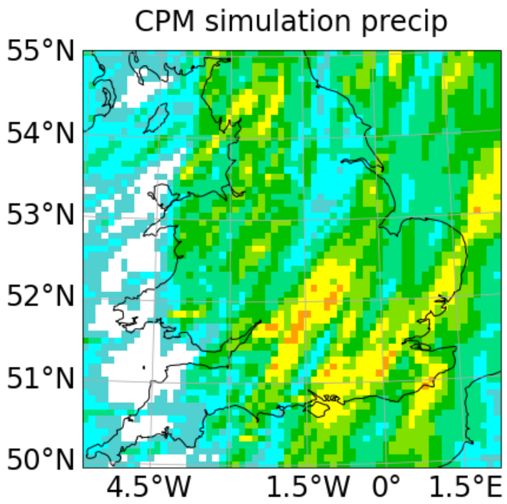
Machine learning emulation of a local-scale UK climate model

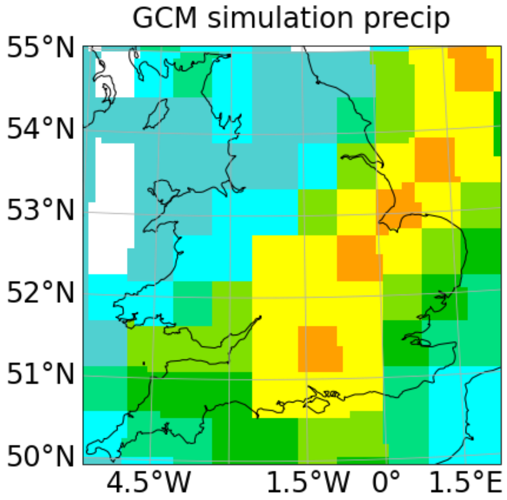
Tackling Climate Change with Machine Learning workshop at NeurIPS 2022

Henry Addison, Lizzie Kendon (Met Office), Suman Ravuri (Google DeepMind), Laurence Aitchison & Peter Watson

The problem

- **Can we learn hi-res precipitation from coarse, global climate simulation output?**
- High-resolution climate simulations are expensive
- New probabilistic ML methods could **complement** simulations with more hi-res samples



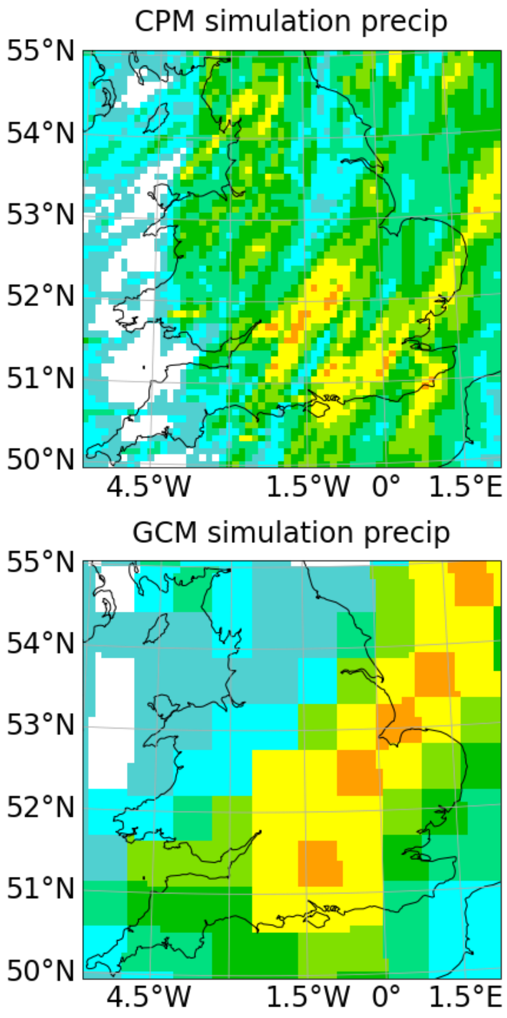


Dataset: Met Office UKCP18

- Projections over UK
  - Global Climate Model (**GCM**) @ **60km**
  - Convection Permitting Model (**CPM**) @ **2.2km**
- 12 CPM ensemble members
  - 60 years each (3 x 20 year chunks)
  - 1 emission scenario
  - **Daily** (and sub-daily available)

In practice currently:

- 64x64 patch for reasons of computational resources
- 8.8km resolution to cover all of England and Wales



Approach

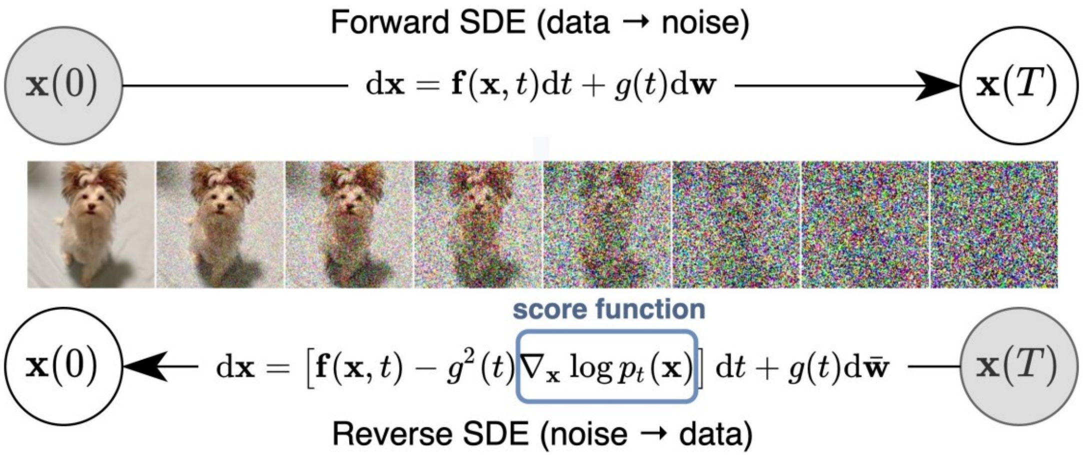
- Training: coarsened CPM variables → hi-res CPM precip
- Sampling: coarsened CPM **OR** GCM variables → hi-res CPM-like precip
- **Use variables which are well-represented in GCMs and physically drive rainfall (without being strongly affected by it)**
  - Wind
  - Temperature
  - Humidity

Diffusion Models

AKA Score-based Generative Models

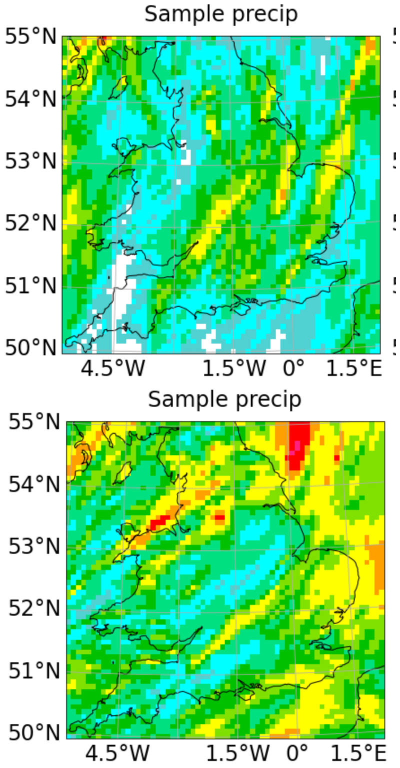
Offer good trade-offs in:

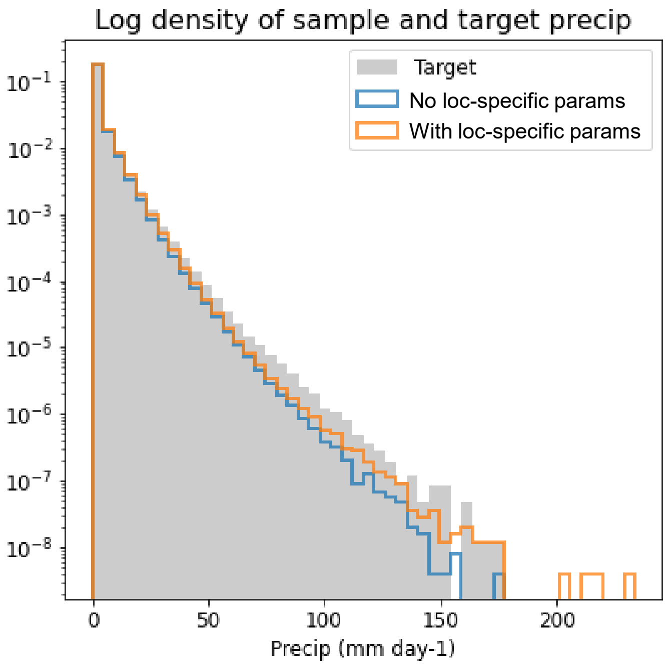
- Sample diversity
- Sample sharpness
- Sampling cost

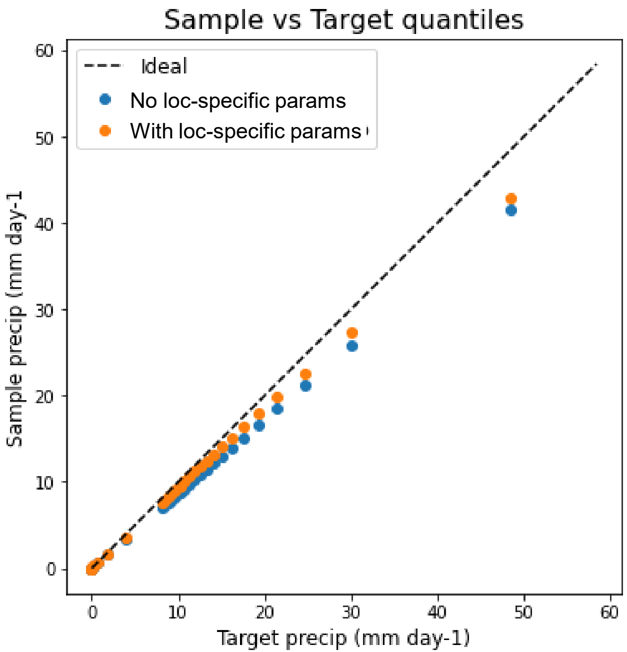


Song, Y. et al., 2021. Score-Based Generative Modeling through Stochastic Differential Equations', *International Conference on Learning Representations.*

Coarsened CPM vorticity@850hPa → 8.8km CPM rainfall

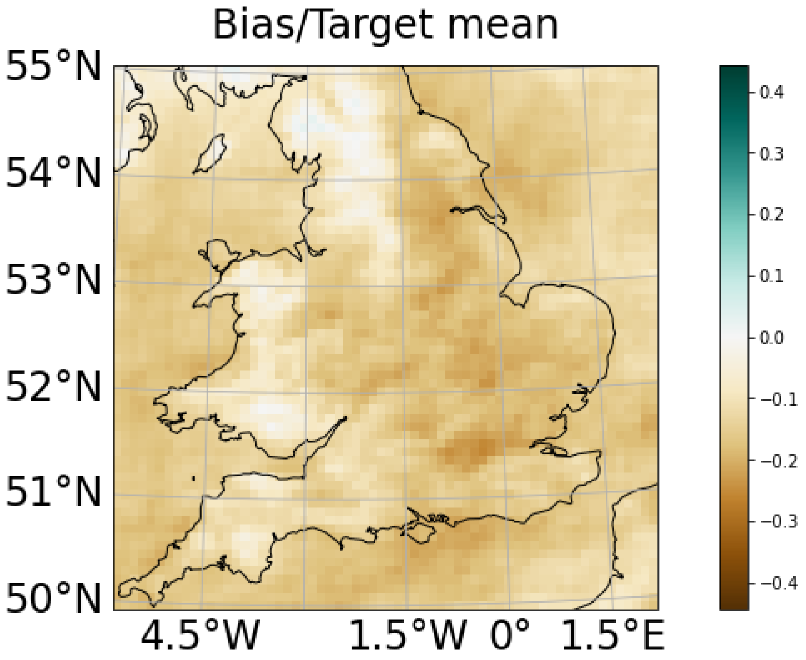




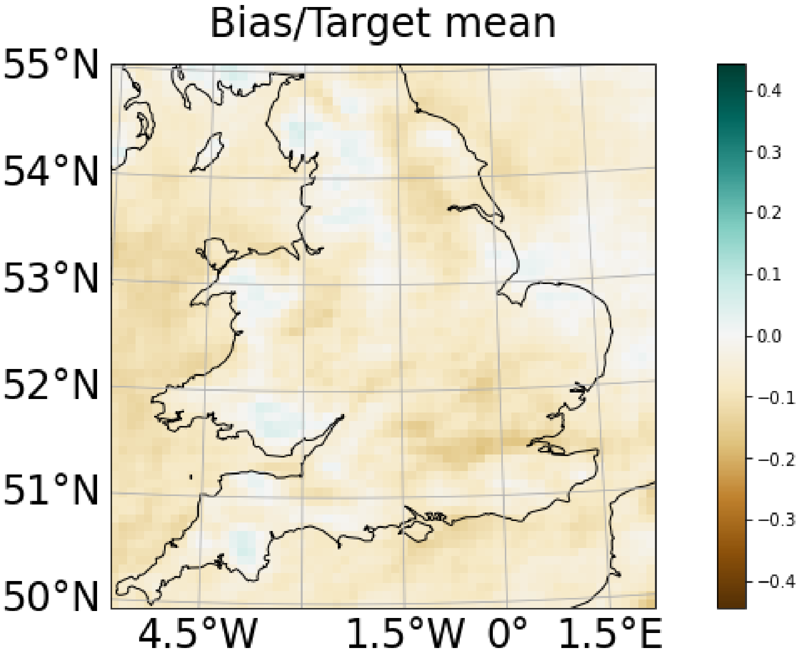


Based on a validation set of 4,320 randomly selected days

Coarsened CPM vorticity@850hPa → 8.8km CPM rainfall



No location-specific parameters



8-channel location-specific parameters

Future work

- More conditioning input variables
- Improve performance when transferring to GCM inputs (e.g. better transforming of inputs)
- Samples from large ensembles of coarse climate models
- Sub-daily frequency and temporal sequences
- More extreme Extremes: 1-in-100 years
- Long-range spatial correlations

Summary

- Hi-res simulations are expensive
- Use much cheaper stochastic ML to **complement** hi-res projections like UKCP18
- First demonstration of diffusion models to predict CPM precipitation

Any questions or suggestions? `henry.addison@bristol.ac.uk`

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.
- Ravuri, Suman et al. 2021. 'Skillful Precipitation Nowcasting using Deep Generative Models of Radar'.
- 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.*
- Met Office Hadley Centre, UKCP18 Local Projections at 2.2km Resolution for 1980-2080, Centre for Environmental Data Analysis, 2019. [Online]. Available: https://catalogue.ceda.ac.uk/uuid/d5822183143c4011a2bb304ee7c0baf7 (visited on 09/06/2021).
- Met Office Hadley Centre, UKCP18 Global Projections at 60km Resolution for 1900-2100, Centre for Environmental Data Analysis, 2018. [Online]. Available: https://catalogue.ceda.ac.uk/uuid/97bc0c622a24489aa105f5b8a8efa3f0 (visited on 09/06/2021).