Can we use machine learning to predict global patterns of climate change?

Laura Mansfield

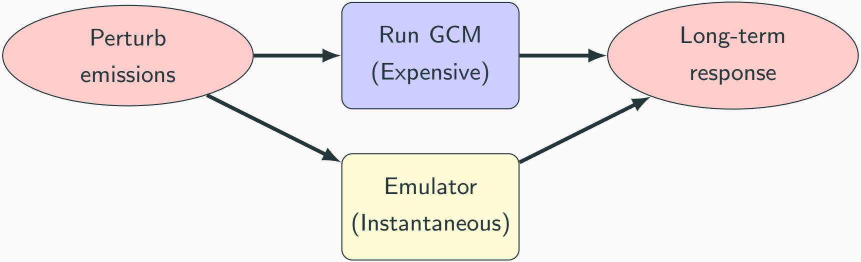
laura.mansfield@pgr.reading.ac.uk

EGU2020 Sharing Geoscience Online

Laura Mansfield[1,2], Peer Nowack[2,3], Matt Kasoar[2], Richard Everitt[4], William Collins[1], Apostolos Voulgarakis[2,5]

1 - University of Reading; 2 - Imperial College London; 3 - University of East Anglia; 4 - University of Warwick; 5 - Technical University of Crete

Predicting long-term response to a change in emissions

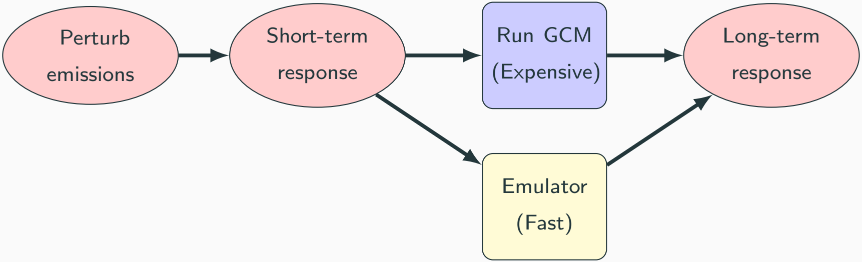


We're interested in the **long-term climate response patterns** (80+ years) to a range of pollutants

- Long-lived (e.g. CO_2) and Short-lived (e.g. SO_4)
- Global and Regional perturbations

Typically we run a perturbed General Circulation Model (GCM). These are expensive. Can we build a machine learning emulator to make fast predictions?

Predicting long-term response to a change in short-term response



We re-use a unique dataset from previous studies, all run in HadGEM3 [Kasoar et al., 2018, Baker et al., 2015, Myhre et al., 2017]

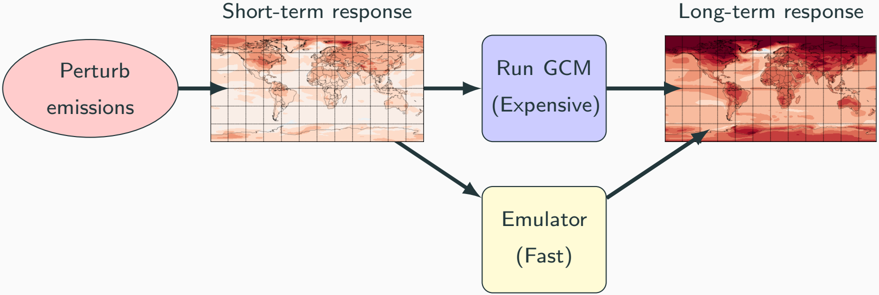
Only $N = 21$ samples!

Some pollutants only have 1 or 2 samples, so we cannot use emissions as inputs to this emulator.

We choose the **inputs** to be **short-term response map** to remove dependence on emission type.

Focus on **surface temperature responses**

Predicting long-term response to a change in short-term response



We re-use a unique dataset from previous studies, all run in HadGEM3 [Kasoar et al., 2018, Baker et al., 2015, Myhre et al., 2017]

Only $N = 21$ **samples!**

Some pollutants only have 1 or 2 samples, so we cannot use emissions as inputs to this emulator.

We choose the **inputs** to be **short-term response map** to remove dependence on emission type.

Focus on **surface temperature responses**

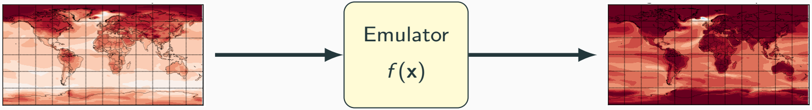
Supervised Learning

A supervised learning problem with

- Small N: Only $N = 21$ training simulations to learn from
- Big p: Take entire response maps (both inputs and outputs) with 145 latitudes, 192 longitudes $p = 145 \times 192 = 27840$

Learn mapping from inputs $\mathbf{x}\,(N \times p)$ to outputs $\mathbf{y}\,(N \times p)$

Inputs $\mathbf{x}\,(N \times p)$ Outputs $\mathbf{y}\,(N \times p)$



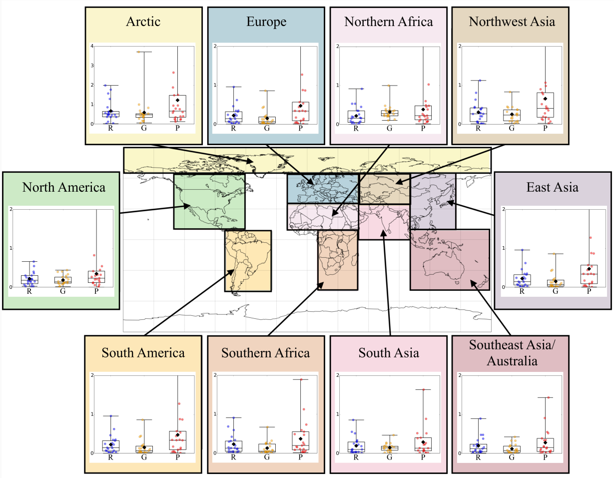
We compare **Ridge regression** and **Gaussian process regression** against a standard approach, **Pattern Scaling**

We train the regression models on all-but-one simulation and predict on the remaining.

[Hoerl and Kennard, 1970, Rasmussen, 2004, Mitchell, 2003]

Performance: Prediction Errors

Errors averaged over broad regions are shown here for **Ridge (R)**, **Gaussian process (G)** and **Pattern Scaling (P)**. We mostly see lower errors for **Gaussian process** predictions.



Summary

- We have explored the use of machine learning emulators to quickly predict **long-term surface temperature response** to **long- and short-lived pollutants**.
- Even with limited data, we find machine learning methods (**Ridge**, **Gaussian Process**) predict response more accurately than the standard approach, **Pattern Scaling**. Global and regional variability is also captured better.
- We also explored a variety of methods (e.g. elastic net, random forest), different input variables (e.g. temperature, geopotential height) and dimension reduction (e.g. physical regions, PCA).
- Could the predictions be improved with **additional data**? Data sharing and collaborations could help us test this.
- Next, we consider emulation of the **short-term response** given the **emissions perturbation**.

Thank you for reading. Happy to take questions, feedback, comments in the chat or via email at laura.mansfield@pgr.reading.ac.uk

References

Baker, L. H., Collins, W. J., et al. (2015).
Climate responses to anthropogenic emissions of short-lived climate pollutants.
Atmospheric Chemistry and Physics, 15(14):8201–8216.

Hoerl, A. E. and Kennard, R. W. (1970).
Ridge Regression: Applications to Nonorthogonal Problems.
Technometrics.

Kasoar, M., Shawki, D., and Voulgarakis, A. (2018).
Similar spatial patterns of global climate response to aerosols from different regions.
npj Climate and Atmospheric Science, 1(1):12.

Mitchell, T. D. (2003).
Pattern Scaling: An Examination of the Accuracy of the Technique for Describing Future Climates.
Climatic Change, 60(3):217–242.

Myhre, G., Forster, P. M., et al. (2017).
PDRMIP: A precipitation driver and response model intercomparison project-protocol and preliminary results.
In *Bulletin of the American Meteorological Society*, volume 98 of *EGU General Assembly Conference Abstracts*, pages 1185–1198.

Rasmussen, C. E. (2004).
Gaussian Processes in machine learning.
Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).