Don’t Waste Data: Transfer Learning to Leverage All Data for Machine-Learnt Climate Model Emulation

Raghul Parthipan[1,2], Damon Wischik[1]

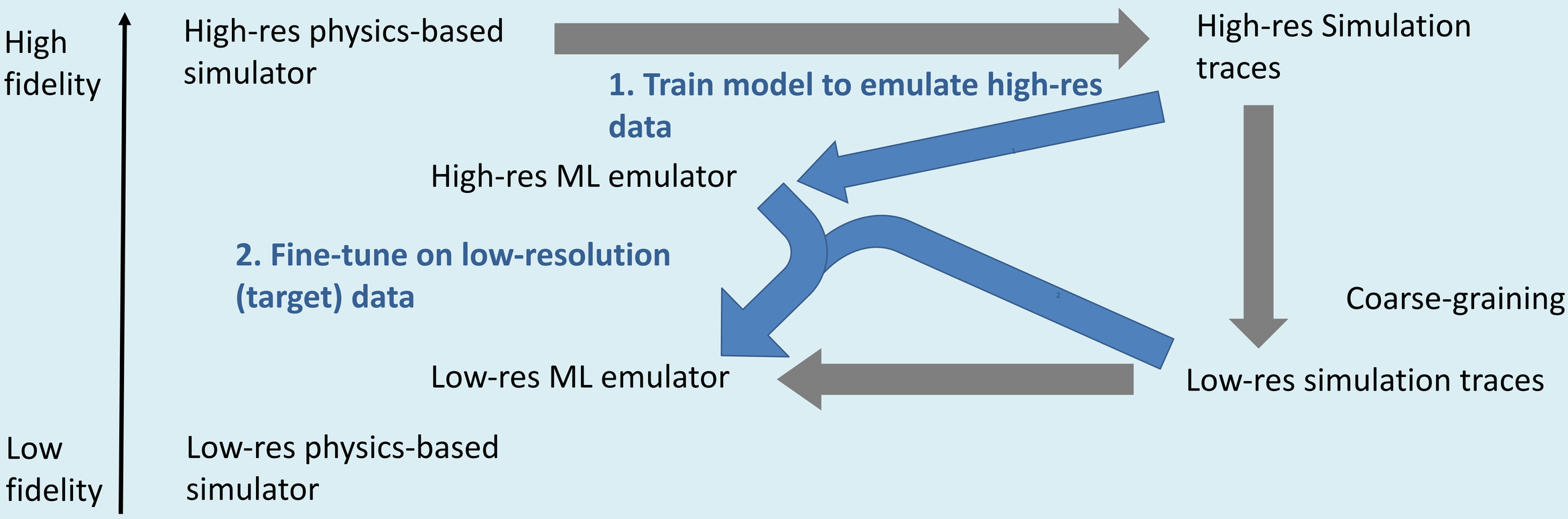
1. Computer Laboratory, University of Cambridge
2. British Antarctic Survey, Cambridge

Correspondence address: rp542@cam.ac.uk

Goal: Create Better ML Emulators of Climate Models

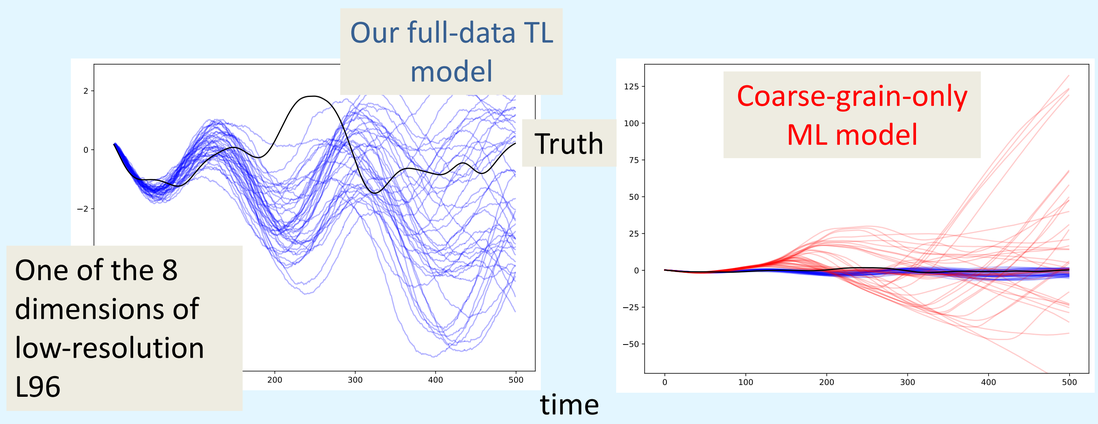
- Often, we have a high-resolution physics-based model which we trust and want to use it to **create a lower-cost emulator** of simulator accuracy. There is much work using ML to do this.
- The **standard approach** is to use **coarse-grained high-resolution** data as training data. The high-resolution data is averaged onto the lower-resolution grid and treated as source data. The aim is to then match the evolution of the coarse-grained data using the lower-resolution model.
- But the **averaging procedure** means much **high-resolution data is thrown away.**
- **Our novelty** is showing that climate model emulation can be framed as a **Transfer Learning** task. We can **use all** the **high-resolution data** as an auxiliary task to learn the low-resolution emulator. And without any extra cost at simulation time.

Our approach is a **two-step process** which uses all data



Better forecasting skill

(L96, KS & Brusselator)



Takeaways

Approach performs particularly well in **data-scarce** scenario, acting as a **regularizer.**

Can be **easily tested on various problems**. Now must be tested on operational models.