Emulating Multiple Variables Across Multiple Earth System Models with Diffusion

Michelle Fast[1], Luke Dojan[1], Katie Christensen[1], Seth Bassetti[2], Brian Hutchinson[1,3], Claudia Tebaldi[3]

[1]Computer Science Department, Western Washington University
[2]Computer Science Department, Utah State University
[3]Joint Global Change Research Institute, Pacific Northwest National Laboratory

Overview

Motivation: Earth System Models (ESMs) are important tools used to analyze possible future climate conditions. However, using them to explore possible scenarios requires intensive energy, time, and money.

Goal: Create an efficient emulator capable of generating multiple daily climate variables in the style of multiple ESMs.

Approach: Train a video diffusion model across data from multiple ESMs to generate month-long daily sequences of climate predictions on a grid of the Earth, conditioned on monthly means.

Background

Earth System Models:

- Generate realistic daily climate data, conditioned on emissions

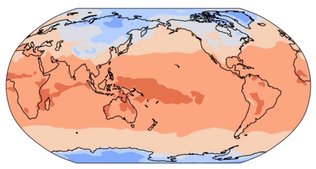

Figure: An example single day of temperature data

Diffusion:

- Generative model: $\boldsymbol{P(x)}$
- Forward process: iteratively add noise to sample (e.g., a 28-day sequence of daily climate values)
- Backward process: iteratively denoise sample by estimating noise and stepping toward denoised estimate

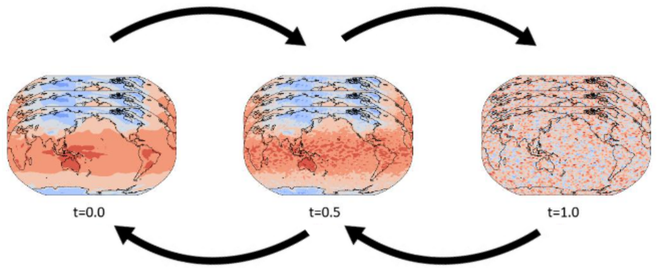

Figure: Diffusion forward and reverse processes

DiffESM:

- Denoising generative probabilistic diffusion model for climate
- Supports only one variable and one ESM
- S Bassetti, B Hutchinson, C Tebaldi, B Kravitz. DiffESM: Conditional Emulation of Temperature and Precipitation in Earth System Models with 3D Diffusion Models. JAMES. Volume 16, no. 10, e2023MS004194, 2024

Methods

- Extend DiffESM to condition on and generate for several ESMs
- For each ESM, we jointly generate daily temperature and precipitation over a global grid for 28-day "months"

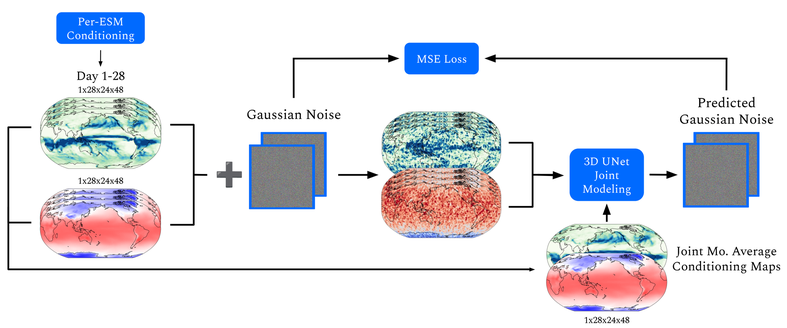

Figure: Training loop

Results

- Performance measured using held-out datasets from **RCP4.5** and **RCP8.5** scenarios across **2080 – 2100**
- For each ESM, **Generated (Gen)** is a novel DiffESM realization conditioned on the monthly means from the **Held Out 1 (HO1)** realization
- The **Held Out 2 (HO2)** realization is included to show inter-realization variability

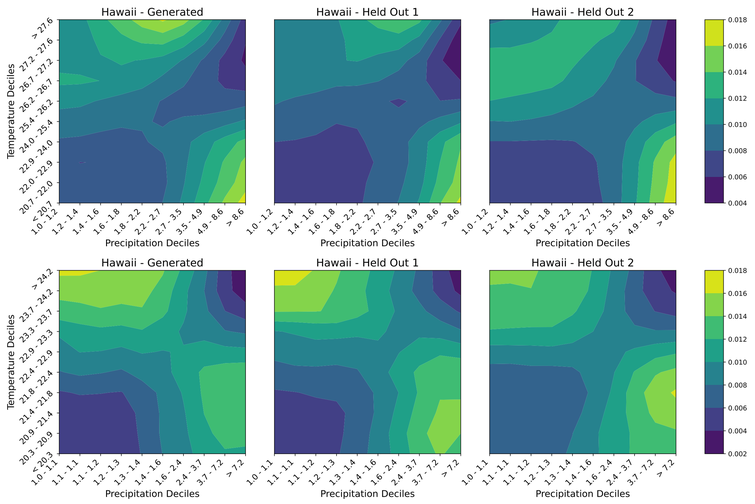

Figure: Joint distribution of temperature and precipitation on wet days (Hawaii) for IPSL (top) and MIROC6 (bottom)

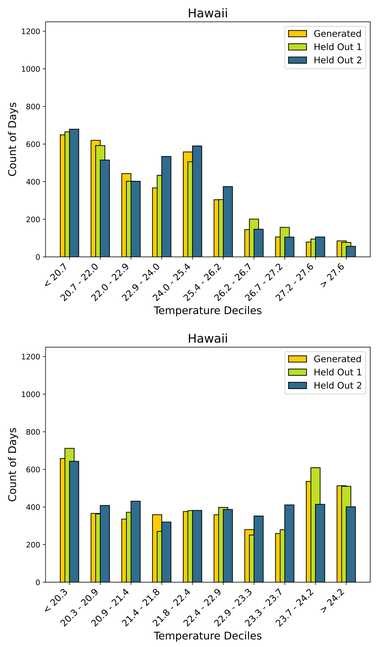

Figure: Temperature distribution on dry days (Hawaii) for IPSL (top) and MIROC6 (bottom)

- DiffESM replicates the relationship between climate variables for both ESMs (IPSL and MIROC6)

Acknowledgments

This research was supported by the U.S. Department of Energy, Office of Science, as part of the GCIMS project under the MultiSector Dynamics, Earth and Environmental System Modeling Program. The authors would also like to acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups at the National Center for Atmospheric Research (USA), the Atmostphere and Ocean Research Institute from the Japan Agency for Marine-Earth Science and Technology (MIROC6, University of Tokyo), and Institute Pierre Simon Laplace (IPSL, France) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

Results (cont.)

- For each month, we compute some climate metric (e.g., number of days exceeding the 90th percentile – "hot days") per-location.
- We plot the average Gen-HO2 and HO1-HO2 differences as maps, and as histograms over spatial locations

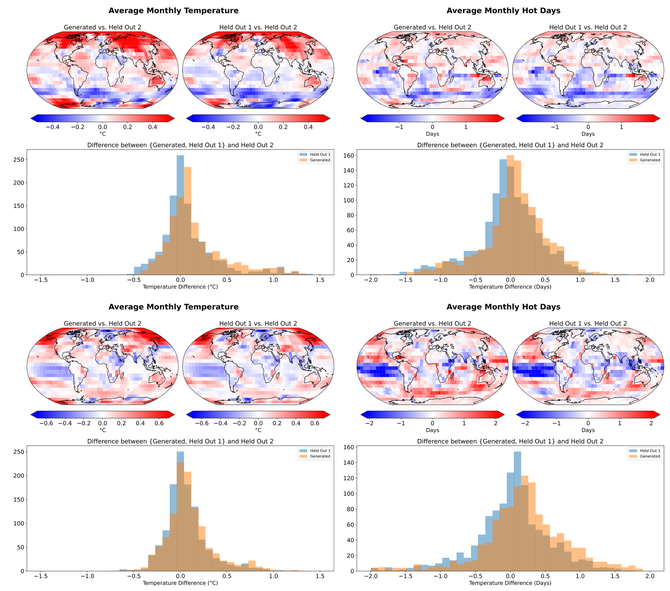

Figure: Avg difference maps for temperature metrics for IPSL (top) and MIROC6 (bottom)

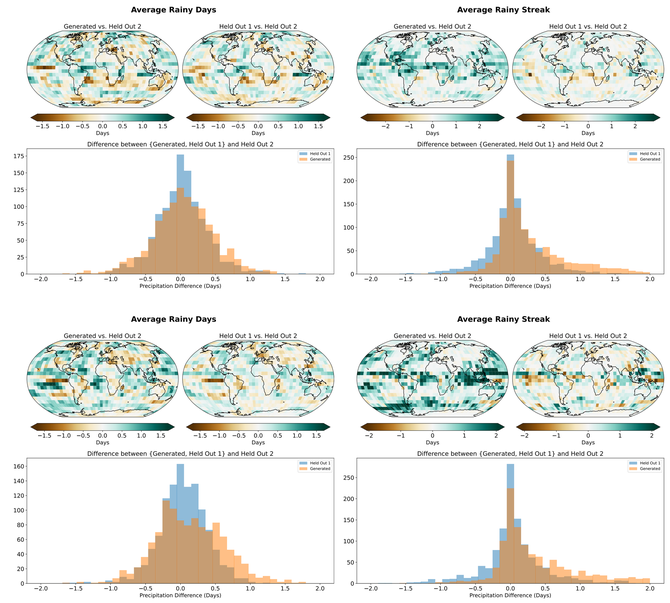

Figure: Average difference maps for precipitation metrics for IPSL (top row) and MIROC6 (bottom row)

- Due to inter-realization variability, we expect differences in both maps/histograms, but expect the both sets of differences to be similar to each other; ideally, the orange histogram will have identical spread to the blue
- The variability of Gen-HO2 is similar to that of HO1-HO2, suggesting that DiffESM is capable of producing new realizations that largely behave like new realizations from the ESM

Conclusions & Future Work

- Our extended DiffESM can effectively and efficiently emulate multiple variables and multiple ESMs simultaneously, producing samples statistically similar to the underlying ESMs
- We plan to replicate these findings on higher spatial resolutions
- We plan to replicate these findings on a larger set of ESMs