ClimForGe: A Diffusion-based Forcing–Response Climate Emulator on Daily Timescales

Jack Kai Lim   Duncan Watson-Parris   Salva Rühling Cachay

University of California San Diego, USA

Introduction

Climate models are indispensable for understanding and projecting climate change, but their high computational costs limit the exploration of diverse scenarios and extreme weather events. Machine learning–based emulators offer a potential solution, however, they often oversimplify complex processes, operate at coarse temporal resolutions, or lack uncertainty quantification.

We introduce ClimForGe — a diffusion-based force-response climate emulator that learns global daily-scale temperature and precipitation responses to realistic CMIP6 forcings, including greenhouse gases, aerosols and shortwave radiation (RSDT). Built on the EDM diffusion framework (Karras et al., 2022).

Our Key Contributions

- **Diffusion-based climate emulator**: Introduces a novel application of diffusion models for global climate forcing–response learning.
- **Rapid scenario generation:** Enables fast, uncertainty-aware exploration of future climate conditions at daily resolution.

Methodology

Background

Diffusion models excel at generating high-quality data samples through an iterative process that gradually adds random noise to data and trains a neural network to reverse this corruption. Hence, given the inherent stochastic nature of diffusion models, we believe that they represent a natural and promising choice for modeling and sampling complex climate systems.

Data

Input Variables - CO_2, CH_4, SiO_2, BC, RSDT*

Output Variables - Surface temperature and Precipitation

Data Source

- Input4MIPs within CMIP6 (Inputs)
- CESM2 (Outputs + RSDT*)

Training Data - SSP126, SSP585 (2015 - 2100) + SSP370 (2015 - 2060).

Validation Data - SSP370 (2060-2100)

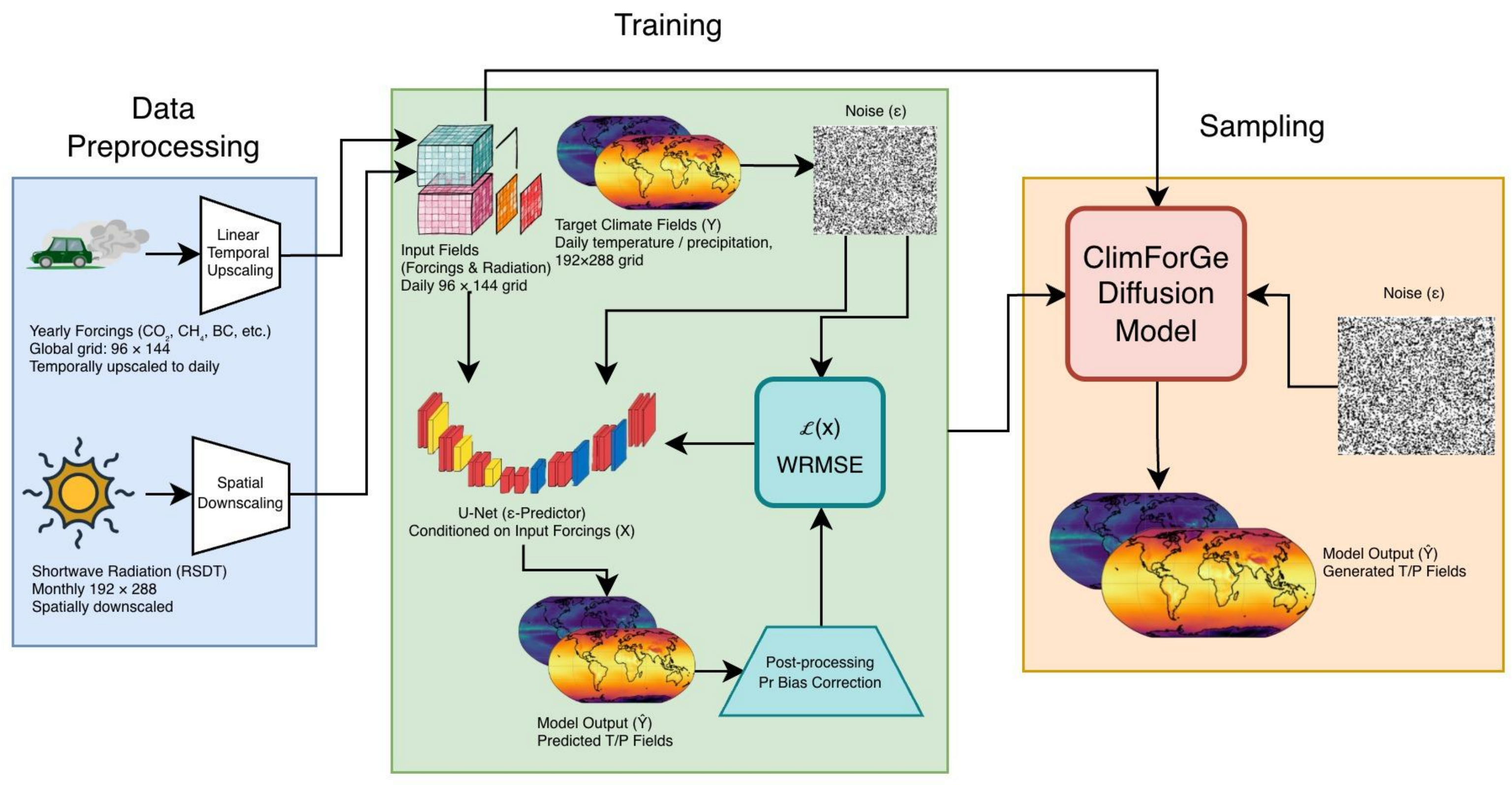
Test Data - SSP245 (2015 - 2100)

Training

- We train 2 models, one for generating tas and one for pr.
- We use a area-weight MSE loss and a U-net backbone training for 100 epochs.
- Diffusion noise schedulers and preconditioning hyperparameters following recommendations from (Karras et al.2022).
- Each model required approximately 4 days to train on 8 Tesla V100-SXM2-32GB GPU
- ClimForGe is able to generate 50 years of full global daily temperature and precipitation fields in approximately 1h 40m on a single Nvidia A100 GPU.

Post Processing

During evaluation we observed a positive bias in precipitation predictions which were traced to CESM2, non-zero values in "dry" areas. To counteract this, we performed a post-processing function which sets all precipitation values below a certain threshold to 0.



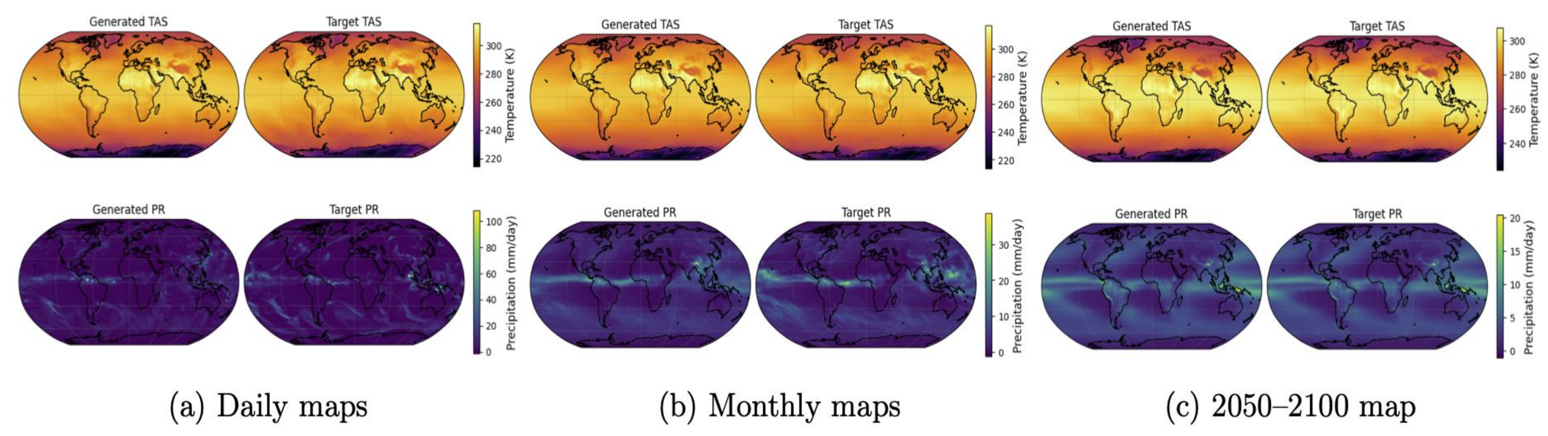


Figure 1. *Global climatological maps comparing ClimForGe and CESM2 under SSP2-4.5, averaged across (a) June 1 2099 instantaneous fields, (b) June 2099 monthly means, and (c) 2050–2100 long-term means. ClimForGe effectively captures key spatial patterns such as global warming regions and the tropical precipitation belt, demonstrating demonstrating strong short- and long-term consistency.*

Results

Table 1. Area-weighted RMSE and Continuous Ranked Probability Score (CRPS) are computed over 20-year means following ClimateBench protocols. Compared to Gaussian Process (GP) and Convolutional Neural Network (CNN) baselines, ClimForGe achieves competitive RMSE and surpasses GP in CRPS—despite operating at 2× higher spatial and 365× higher temporal resolution showcasing its strength as a daily-scale probabilistic emulator.

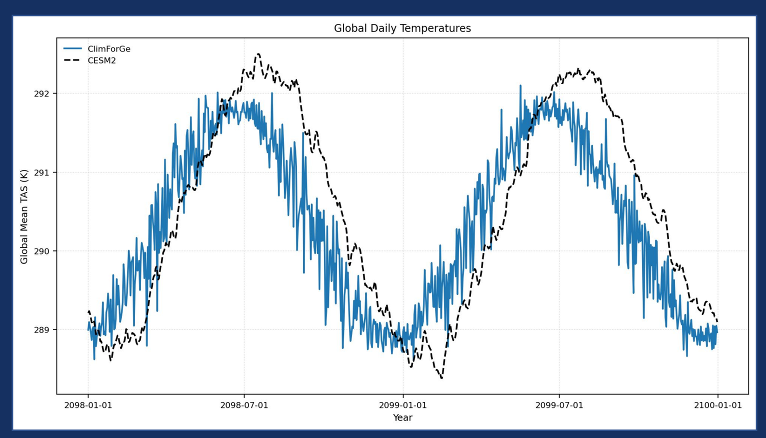


Figure 2. *Global Means on a Daily Scale, Generated from ClimForGe vs CESM2 targets for temperature, showcasing ClimForGe's generation ability on a daily scale compared to CESM2.*

	Temperature		Precipitation	
Model	**RMSE (K)**	**CRPS (K)**	**RMSE (mm/day)**	**CRPS (mm/day)**
Gaussian Process	**0.225**	**0.477**	**0.153**	**1.075**
CNN	**0.222**	**-**	**0.139**	**-**
ClimForGe (Ours)	**0.377**	**0.282**	**0.239**	**0.128**

Table 1. *RMSE and CRPS of global and time means for the years 2080-2100 of SSP245 for ClimForGe and Climatebench models. Note that because Climatebench emulates NorESM, these metrics are not directly comparable but the metrics provide a useful reference point for relative model skill.*

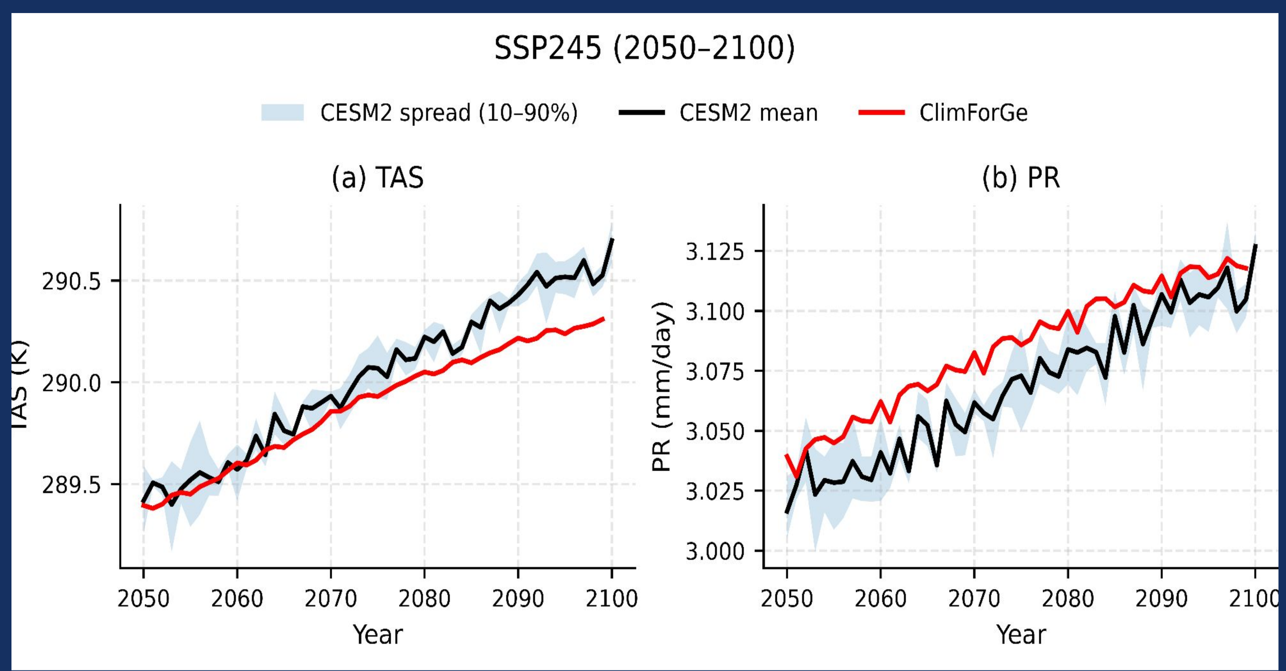


Figure 3. *Comparison of ClimForGe's global annual mean temperature and precipitation against CESM2 ensemble means (10–90% spread across three CESM2 members) for 2050–2100 under SSP2-4.5. ClimForGe closely tracks CESM2 trends, accurately capturing the positive long-term warming and precipitation patterns.*

Discussion

We demonstrate that a diffusion-based architecture can effectively emulate complex forcing–response climate dynamics, enabling rapid exploration of diverse socioeconomic scenarios while accurately capturing the uncertainties and internal variability of Earth System Models.

However, several limitations remain — including the independent generation of temperature and precipitation and the lack of temporal coherence across days due to the i.i.d. sampling process. Future work will focus on joint variable generation and diffusion guidance techniques to allow for better and more consistent exploration of rare and extreme events, further improving realism and predictive skill for climate risk assessment.

https://climate-analytics-lab.github.io

jklim3

jklim@ucsd.edu

Climate Analytics Lab

SCRIPPS INSTITUTION OF OCEANOGRAPHY

UC San Diego

HALICIOĞLU DATA SCIENCE INSTITUTE