HClimRep: AI Climate Model for Capturing the Atmosphere, Ocean, and Sea Ice Interactions

Savvas Melidonis, Ankit Patnala, Asma Semcheddine and **Martin Schultz** *Forschungszentrum Jülich*
{s.melidonis,a.pantala,a.semcheddine,m.schultz}@fz-juelich.de
Julius Polz *Karlsruher Institut für Technologie* julius.polz@kit.edu
Kacper Nowak *Alfred Wegener Institute* kacper.nowak@awi.de

Abstract

Climate change poses a significant threat to ecosystems and human society. Accurate climate projections are crucial for developing effective policies for mitigating extreme weather events that are expected to increase due to global warming. However, traditional climate models have limitations, including biases and high computational costs. Under the Helmholtz Foundation Model Initiative (HFMI), we propose a new data-driven climate model, namely HClimRep, which uses foundation model principles and machine learning to analyze diverse climate datasets. This approach enables flexible and customizable outputs, providing a versatile tool for climate applications.

Keywords: Foundation Model, AI Model, Deep Learning, Climate Modelling, Climate Simulations

1 Introduction

Artificial Intelligence (AI) and Deep Learning (DL) have recently emerged as powerful alternatives to traditional numerical methods for weather forecasting. Since 2023, these models have demonstrated operational value, such as the Artificial Intelligence/Integrated Forecasting System (AIFS) [Lang et al., 2024], by achieving comparable or superior accuracy at far lower computational cost compared to conventional numerical models that are based on equation solvers. However, extending AI methods from medium-term weather prediction to long-term climate projections remains an open challenge. Unlike weather forecasts, climate projections require a consistent representation of very complex feedback processes across several components of the Earth system and additionally at very long timescales.

Several promising approaches highlight the feasibility of AI-driven climate modeling. For instance, cBottle model [Brenowitz et al., 2025] employs a U-Net architecture and a diffusion-based generative algorithm with super-resolution capabilities for high-resolution climate projections, while in the Deep Learning Earth System Model (DLESyM) [Cresswell-Clay et al., 2025] model, a U-Net architecture combined with an efficient training strategy to couple atmosphere and ocean has been employed. Spherical Fourier Neural Operators (SFNO) [Bonev et al., 2023] have been the backbone of the architectures in some works such as the Spherical DYffusion model [Cachay et al., 2024] that integrates SFNO to achieve thousands of years of stable climate simulation, and the lightweight uncoupled climate emulator LUCIE [Guan et al., 2025] that has been trained on as few as 2 years of training data. Despite these significant advances, most of these studies either omit feedback mechanisms involving atmospheric chemistry (such as greenhouse gases), the stratosphere, or sea ice, or they are limited to seasonal timescales. In addition, they are not trained on diverse sources of multi-modal data and observations but exclusively on biased climate data, precluding them from being considered foundation models.

Our early-stage research project HClimRep, as part of the Helmholtz Foundation Model Initiative, aims to deliver an AI foundation model for climate that complements the WeatherGenerator initiative (an EU-funded effort led by the European Centre for Medium-Range Weather Forecasts (ECMWF) with strong involvement of Forschungszentrum Jülich). The prototype is designed to ingest heterogeneous datasets — including reanalyses, climate simulations, satellite records, and in-situ observations — and to support downstream applications, such as hydrological downscaling, climate scenarios interpolation, a forecasting system for stratospheric tracers and warming as well as a simulation of counterfactual scenarios of recent marine heatwaves.

Key distinguishing features of HClimRep compared to the aforementioned works include the incorporation of stratospheric dynamics, tracers such as ozone, a fully interactive ocean component with deep layers, and explicit handling of external forcings (e.g., CO_2 pathways) while allowing forecasting on a longer horizon. This scope is unique worldwide and specifically addresses limitations in existing AI models. Our approach employs a two-stage training strategy: pretraining on high-quality observational and reanalysis datasets, followed by fine-tuning on climate simulations. This ensures fidelity in near-term dynamics while enabling the model to learn to distinguish among climate-forcing scenarios.

2 Model

The WeatherGenerator (WGen) architecture [Lessig et al., 2025] is designed to flexibly process diverse data sources while minimizing the need for external preprocessing, allowing the majority of the data to be used in its original form. Figure 1 illustrates the architecture of the model.

WGen consists of five core modules: stream (i.e., dataset) embedding, local assimilation engine, global assimilation engine, forecasting engine, and prediction heads. In the stream embedding stage, the model learns interactions between different state variables within the same data source. The local engine reorganizes data from different sources onto a common Hierarchical Equal Area isoLatitude Pixelization (HEALPix) [Górski et al., 2005] grid representation and learns cross-source interactions locally, i.e., within each cell. The global engine then assimilates information across all cells, capturing global interactions.

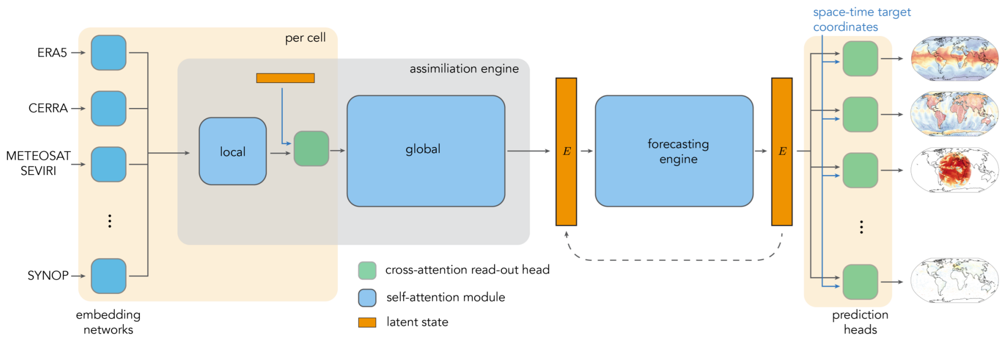


Figure 1: The HClimRep prototype using WeatherGenerator's architecture.

The forecasting engine performs dynamic stepping by predicting future latent states, while the prediction heads are target-agnostic, offering the flexibility to specify target coordinates and data sources for which the user requires predictions. This modular design enables multiple applications, including fine-tuning on heterogeneous datasets, zero-shot and supervised downscaling (from global to regional scales), nonlinear interpolation, and atmospheric state reconstruction from data gaps (e.g., satellite observations). These features distinguish WGen from existing models.

The HClimRep prototype builds directly on the WGen but differs in its objectives. While WGen is primarily designed for medium-term weather forecasting, HClimRep extends the scope to seasonal and decadal timescales. Achieving skillful predictions on these horizons requires coupling across multiple components of the Earth system, i.e., atmosphere, stratosphere, ocean, and sea ice, thereby enabling the model to learn energy exchanges and feedbacks among these interacting subsystems.

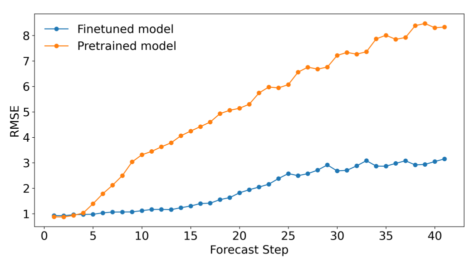
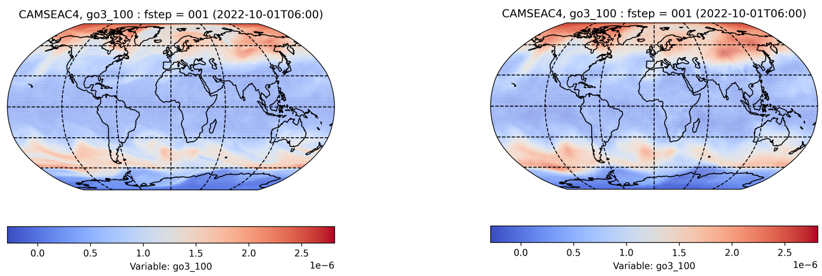


Figure 2: Forecasting over 10 days (i.e., 40 6h-steps) for the 2m-temperature.

3 Results

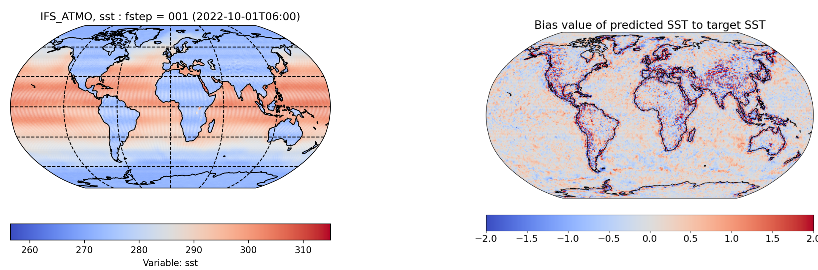
The HClimRep project is currently focusing on advancing each Earth system component individually. We are conducting and testing three major tasks in parallel as part of our early development: extending the WeatherGenerator



(a) Ground truth of global O_3 on the 100^{th} pressure level with 6h-lead.

(b) Prediction of global O_3 on the 100^{th} pressure level with 6h-lead.

Figure 3: Current progress on capturing stratospheric tracers.



(a) Predicted SST

(b) Prediction bias on SST

Figure 4: Forecast map of sea surface temperature values and bias plot.

model forecast horizon on ERA5 atmospheric reanalysis data [Hersbach et al., 2020], integrating a unique dataloader for Finite-Element/volumE Sea ice-Ocean Model (FESOM) oceanic data [Danilov et al., 2017], and studying the model's capability to capture stratospheric tracer dynamics.

We train the model on 40 years of ERA5 atmospheric data (1979-2020) for the atmospheric component, 200 years of oceanic simulation data for the oceanic component, and 20 years of CAMS data (2003-2024) for the stratospheric component. We pretrain the model with two 6h-forecast steps (over 24 hours 64 epochs) and finetuning with 9 6h-forecast steps (over 36 hours 32 epochs) on JUWELS Booster on 2 nodes with 2xIntel Xeon Platinum 12 CPUs, and 8 NVIDIA A100 GPUs. The code is published and can be found in the Weather-Generator GitHub repository. The trainable parameters of the model are half a billion parameters.

Progress has been made on the first task, gradually extending forecast skill from an initial 2-day horizon to 14 days for the atmosphere and 1 year for the ocean over the course of 4–5 months of continuous development, see Figure 2. We are currently focusing on significantly extending the forecast of the atmospheric component, which is more challenging.

Experiments that target modelling of the dynamics of the stratospheric tracers have also been successful. Specifically, we have now a dedicated stream has been integrated into the model to capture Copernicus Atmosphere Monitoring Service (CAMS) greenhouse reanalysis data [Inness et al., 2019]. Finetuning on those components is still in progress but the first training experiments have been completed, see Figure 3.

The development of the ocean component focuses on processing ocean data, which inherently have more challenging spatial structures than atmospheric data. A dedicated stream embedding component has been developed and integrated into the model, allowing training on oceanic inputs in parallel with atmospheric inputs (e.g., from ERA5). Prediction runs of ocean states have demonstrated promising results, with precise capture of SST, see Figure 4.

The next steps in the project will be to increase the forecast length to several months and enable stable rollouts over at least a decade. The various compartments (i.e., ocean, troposphere, and stratosphere) will be coupled together to allow for the investigation and evaluation of various climate feedback processes.

Acknowledgements

This work was supported by the Helmholtz Association within the framework of the Helmholtz Foundation Model Initiative (HClimRep). We are grateful for the substantial contributions of many individuals and teams, whose efforts have been integral to this work. In particular, we would like to thank Christian Lessig and the WeatherGenerator development team for their great efforts and meticulous organisation of the software development. We also acknowledge all the HCLimRep partners for their collaborative support and helpful discussions, which have significantly advanced this line of research.

References

B. Bonev, T. Kurth, C. Hundt, J. Pathak, M. Baust, K. Kashinath, and A. Anandkumar. Spherical Fourier Neural Operators: Learning stable dynamics on the sphere, 2023. URL `https://arxiv.org/abs/2306.03838`.

N. D. Brenowitz, T. Ge, A. Subramaniam, P. Manshausen, A. Gupta, D. M. Hall, M. Mardani, A. Vahdat, K. Kashinath, and M. S. Pritchard. Climate in a Bottle: Towards a generative foundation model for the kilometer-scale global atmosphere, 2025. URL `https://arxiv.org/abs/2505.06474`.

S. R. Cachay, B. Henn, O. Watt-Meyer, C. S. Bretherton, and R. Yu. Probabilistic emulation of a global climate model with spherical DYffusion, 2024. URL `https://arxiv.org/abs/2406.14798`.

N. Cresswell-Clay, B. Liu, D. R. Durran, Z. Liu, Z. I. Espinosa, R. A. Moreno, and M. Karlbauer. A deep learning earth system model for efficient simulation of the observed climate. *AGU Advances*, 6(4):e2025AV001706, 2025.

S. Danilov, D. Sidorenko, Q. Wang, and T. Jung. The finite-volumE sea ice–ocean model (FESOM2). *Geoscientific Model Development*, 10(2):765–789, 2017. doi: 10.5194/gmd-10-765-2017. URL `https://gmd.copernicus.org/articles/10/765/2017/`.

K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. HEALPix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere. *The Astrophysical Journal*, 622(2):759–771, 2005. doi: 10.1086/427976. URL `https://doi.org/10.1086/427976`.

H. Guan, T. Arcomano, A. Chattopadhyay, and R. Maulik. LUCIE: A lightweight uncoupled climate emulator with long-term stability and physical consistency for O(1000)-member ensembles, 2025. URL `https://arxiv.org/abs/2405.16297`.

H. Hersbach, B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, et al. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020. doi: 10.1002/qj.3803. URL `https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.3803`.

A. Inness, M. Ades, A. Agustí-Panareda, J. Barré, A. Benedictow, A.-M. Blechschmidt, J. J. Dominguez, R. Engelen, H. Eskes, J. Flemming, et al. The cams reanalysis of atmospheric composition. *Atmospheric Chemistry and Physics*, 19(6):3515–3556, 2019.

S. Lang, M. Alexe, M. Chantry, J. Dramsch, F. Pinault, B. Raoult, M. C. A. Clare, C. Lessig, M. Maier-Gerber, L. Magnusson, Z. B. Bouallègue,

A. P. Nemesio, P. D. Dueben, A. Brown, F. Pappenberger, and F. Rabier. AIFS – ECMWF's data-driven forecasting system, 2024. URL https://arxiv.org/abs/2406.01465.

C. Lessig, I. Luise, T. Hunter, K. Tezcan, S. Grasse, S. Melidonis, M. Langguth, K. Nowak, M. Schultz, and P. Dueben. The Weathergenerator: A foundation model for weather and climate. In *CESOC Conference Proceedings*, Bonn, Germany, 2025. URL https://cesoc.net/wp-content/uploads/2025/08/40_Lessig.pdf. Presented at MLESM 2025.