

APPLICATION PAPER

Simulation of global sea surface temperature maps using Pix2Pix GAN

Deepayan Chakraborty  and Adway Mitra

Department of Artificial Intelligence, Indian Institute of Technology, Kharagpur, India

Corresponding author: Deepayan Chakraborty; Email: deepayan504@email.com

Received: 25 July 2024; **Accepted:** 28 September 2024

Keywords: CMIP6; generative adversarial network; global climate model; sea surface temperature

Abstract

Simulated data from the Coupled Model Intercomparison Project Phase 6 (CMIP6) has been very important for climate science research, as they can provide wide spatio-temporal coverage to address data deficiencies in both present and future scenarios. However, these physics-based models require a huge amount of high-performance computing (HPC) resources. As an alternative approach, researchers are exploring if such simulated data can be generated by Generative Machine Learning models. In this work, we develop a model based on Pix2Pix conditional Generative Adversarial Network (cGAN), which can generate high-resolution spatial maps of global sea surface temperature (SST) using comparatively less computing power and time. We have shown that the maps generated by these models have similar statistical characteristics as the CMIP6 model simulations. Notably, we trained and validated our cGAN model on completely distinct time periods across all ensemble members of the EC-Earth3-CC and CMCC-CM2-SR5 CMIP6 models, demonstrating satisfactory results and confirming the generalizability of our proposed model.

Impact Statement

Climate studies heavily rely on simulated climate data from the Coupled Model Intercomparison Project Phase 6 (CMIP6) due to its significant contribution in addressing data gaps. Nevertheless, the computational demands of physics-based models within CMIP6 necessitate huge HPC resources. To overcome this challenge, the proposed Pix2Pix GAN model enables the generation of global maps of SST, similar to the CMIP6 model, on a comparatively low hardware-based system and in a reduced timeframe. By combining domain knowledge with cutting-edge machine learning, this work represents an important step towards the accessibility of climate modeling tools and datasets.

1. Introduction

Climate models are computer simulations of the complex physical processes in the Earth's climatic system. These models are based on differential equations over climatic variables like temperature, humidity, and wind across a 3D grid representing the Earth System (Flato *et al.* 2014). Climate simulations are carried out by solving these equations at multiple spatiotemporal resolutions. In recent years there have emerged several families of such climate models, e.g., the Coupled Model Intercomparison Project version 6 (CMIP6; Eyring *et al.* 2016). These models have become vital tools for understanding climate change under various future scenarios. However, the computational demands of running these physics-

based models limit their accessibility and utility. These models are often used to simulate various geophysical parameters, e.g., sea surface temperature (SST) at both regional and global scales. Such simulations are often used by climate researchers for analyzing different geophysical phenomena (e.g., El-Niño Southern Oscillations [ENSO], Indian Ocean Dipole [IOD], etc.), regional rainfall forecasts at different timescales (Sharma *et al.* 2023, Le *et al.* 2023). However, such simulations are computationally very expensive and typically require supercomputers.

Machine learning (ML) methods like deep neural networks have emerged as powerful techniques for modeling complex systems and are being used in multiple scientific fields, including climate sciences (Bochenek & Ustrnul 2022). Deep neural networks are emerging as one of the most popular ML models in climate sciences due to their constantly improving performance (Sun *et al.* 2023). However, the main bottleneck of these models is the requirement for a huge amount of data to train. To address this challenge, we present a cost-effective alternative for generating synthetic climatic data, which can be used to train such models. Specifically, we propose to use a Generative Adversarial Network (GAN) for generating data such as SST. The proposed model is based on the Pix2pix conditional GAN, which takes observed SST anomaly data as the condition to calibrate the generated data with the observed one. Another contribution of the model is the generation of spatiotemporal maps. That means the model is able to generate spatial maps for a given timestamp. The model is validated for one CMIP6 model (EC-Earth3-CC). Hence, this approach is able to generate a map equivalent to a CMIP6 model with significantly less computational cost.

2. Literature survey

The coupled model intercomparison project (CMIP) was established in 1995 by the Joint Scientific Committee and CLIVAR (Core Project of the World Climate Research Programme) Working Group on Coupled Models to facilitate ensemble climate modeling. Now in its sixth iteration, CMIP6 enables coordinated simulations across complex physics-based models (Eyring *et al.* 2016). CMIP outputs like SST are valuable real-world surrogates. The ocean and atmosphere are intrinsically linked (Gill, A. E. 2016); hence simulated oceanic variables serve as predictors in diverse climate studies (Richter & Tokinaga 2020 and Rivera & Arnould (2020)). In particular, SSTs quantify various oceanic phenomena like Atlantic Multi-decadal Oscillation (Goswami *et al.* 2022), El-Niño Southern Oscillation (Karoly, D. J. 1989), North Atlantic Niño (Yadav *et al.* 2018), and so forth. They have a significant impact on the precipitation of East Africa (Yan *et al.* 2020, Yang *et al.* 2018, Stevenson *et al.* 2012, Endris *et al.* 2019), Maritime continent (Power *et al.* 2013, Xu *et al.* 2017), Southern and Eastern parts of Asia (Goswami *et al.* 2022, Wang *et al.* 2020, Yang *et al.* 2018), North America (Fasullo *et al.* 2018, Yang *et al.* 2018) and so forth. Hence CMIP6's 165-year (1850–2014) SST simulations have various utilities in the climate studies of various continental regions (Krishnamurthy & Krishnamurthy 2014).

Meanwhile, ML and deep learning (DL) offer versatile utilities for geoscientists and climatologists. It includes the prediction of rainfall (Endalie *et al.* 2022, Haq *et al.* 2021, Sharma *et al.* 2023), extreme events (Yan *et al.* 2022), and streamflow (Singh *et al.* 2023); downscaling of data (Niazkar *et al.* 2023, Sharma *et al.* 2022) and so forth. Apart from these prediction models, generative models are also used in climate studies. In recent years various variants of GAN have started to be used to model turbulent climate dynamics (Gupta *et al.* 2020), flood frequency estimation (Ji *et al.* 2024), tipping point discovery (Sleeman *et al.* 2023) and so forth. However, there is a lot to explore in this field where GAN can come up with more effective solutions compared to the traditional models.

Motivated thus, we develop a GAN architecture to generate CMIP6-like global SST fields based solely on observational data. This model has the potential to increase the accessibility of simulated data by replacing the physics-based model with a statistical model, hence reducing the hardware requirement.

3. Dataset and preprocessing

Our study uses two extensive datasets for model training and validation:

Observed Sea Surface Temperature (SST) data: Monthly global data spanning 1871–2010 was acquired from The Simple Oceanic Data Assimilation (SODA) v2.2.4 dataset (Giese *et al.* 2011).

Simulated SST data: Ten ensembles from the Coupled Model Intercomparison Project Phase 6 (CMIP6) dataset for the EC-Earth3-CC and CMCC-CM2-SR5 models were selected, covering the same period.

Both datasets were preprocessed and resampled to a common spatial resolution of $3.75^\circ \times 2.5^\circ$, resulting in a data dimension of 144×48 towards longitude and latitude, respectively, for each month.

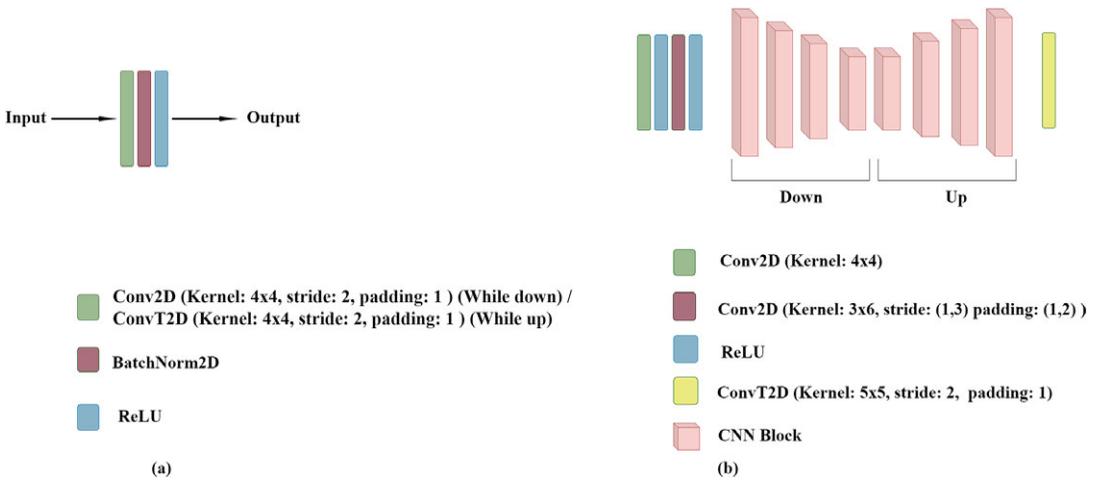


Figure 1. The architecture of the generator model. (a) The basic convolutional neural network (CNN) block is used in the generator; (b) The full generator model, where the combination of the CNN block is used.

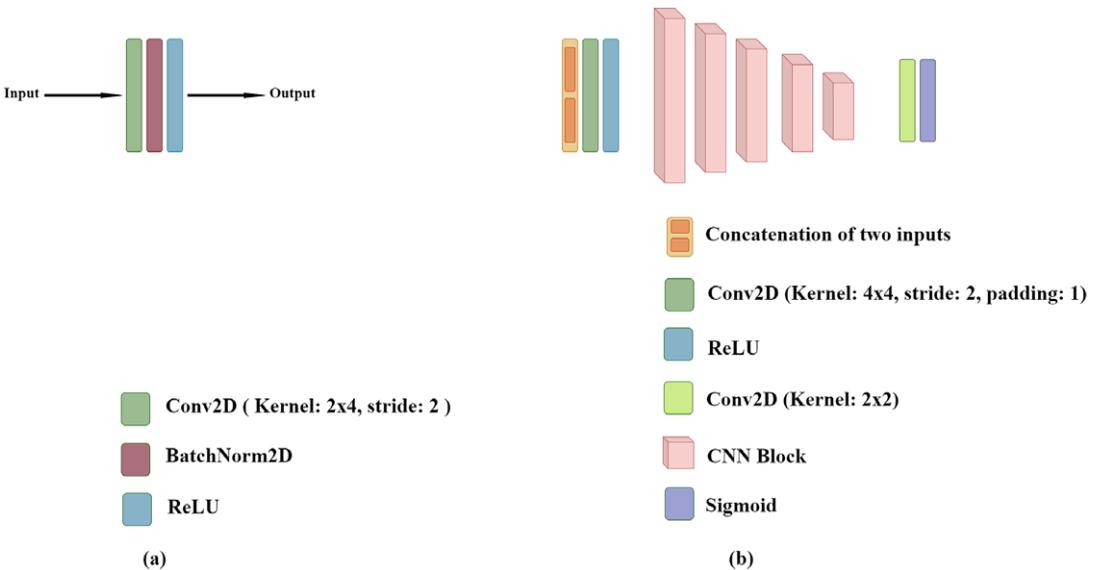


Figure 2. Architecture of the discriminator model. (a) The basic convolutional neural network (CNN) block is used in the discriminator; (b) The full discriminator model, where the combination of the CNN block is used.

Missing values were replaced with zeros, and observed data anomalies were calculated by subtracting the long-term mean.

For training, the first 130 years (1871–2000) of data were used. To match the dimensionality of the 10 CMIP6 ensembles, the observed data were replicated tenfold. This allowed training the model on each input data point separately for each ensemble. Validation was performed for the remaining 14 years (2001–2014). For each year, the mean of the 10 ensembles of the simulated data were used, ensuring the performance of the model was evaluated against the average ensemble behavior, instead of individual variations. This data processing and preparation strategy ensures thorough training on diverse simulated scenarios while evaluating the generalizability of the model against the average ensemble response.

4. Methodology

We develop a conditional generative adversarial network (cGAN) based on the Pix2pix architecture to generate simulated SST data based on the observed anomaly data as the condition. The cGAN contains

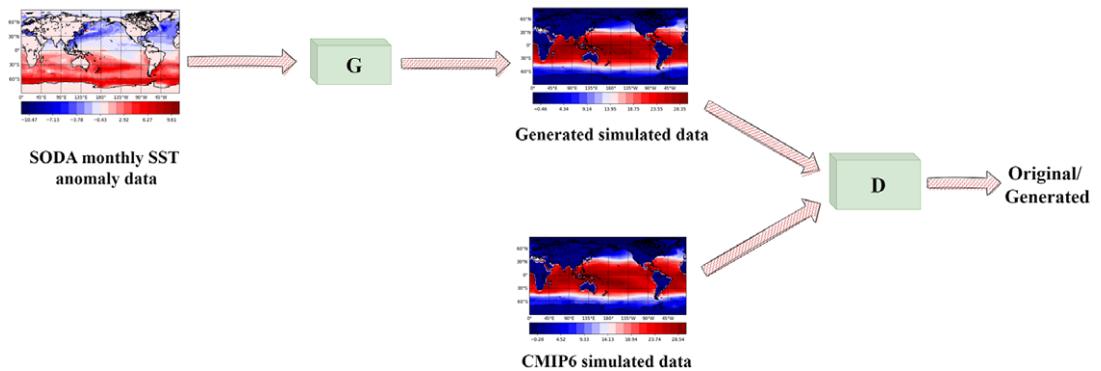


Figure 3. Architecture for adversarial learning for the proposed model.

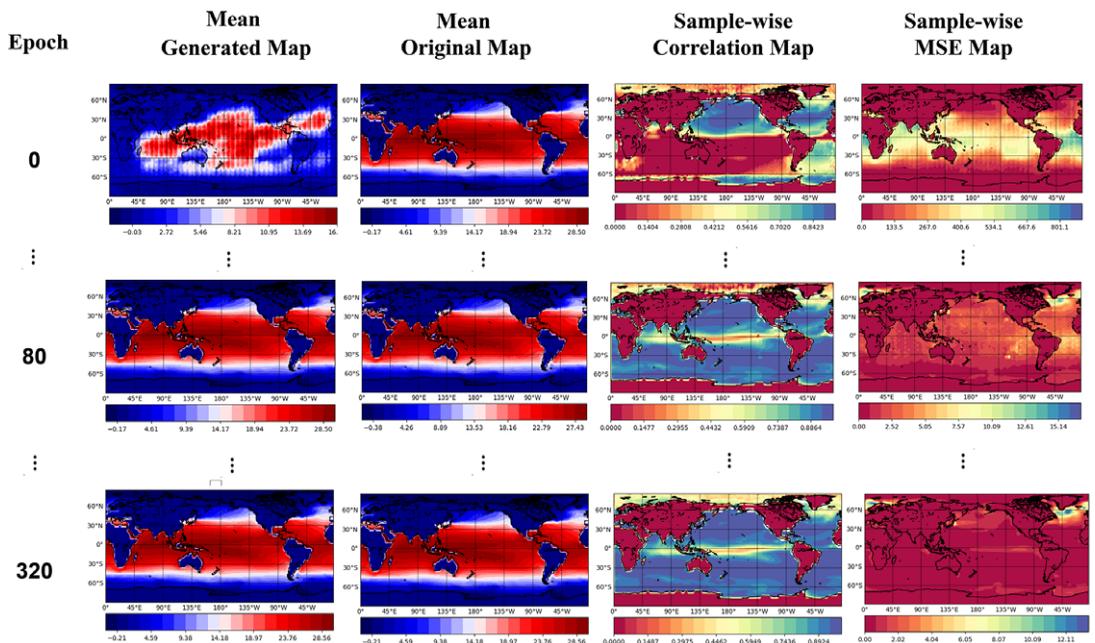


Figure 4. Evolution of the generator on training data for the EC-Earth3-CC GCM model.

two key components—a generator model to produce synthetic SST data, and a discriminator model to differentiate real versus generated SST.

4.1. Generator model

The generator is a convolutional neural network (CNN) that takes observed SST anomaly data as input and outputs corresponding simulated SST data. The Generator consists of a combination of a basic CNN block. As shown in Figure 1(a), the basic CNN block comprises convolutional (Conv2D) and transposed convolutional (ConvT2D) layers to down-scale (down) and up-scale (up) the spatial dimensions of the features, respectively. This layer is followed by batch normalization (BatchNorm2D) and ReLU activation.

The input first undergoes initial feature extraction through two Conv2D and ReLU activation layers. The primary structure of the generator consists of four CNN blocks for downsampling features and an additional four blocks for upsampling. The final output layer is a ConvT2D layer to produce the simulated SST output. This architecture enables the generator to learn the spatial variability between input and output data. The full generator model is visually illustrated in Figure 1(b).

Table 1. Epoch-wise updation of the correlations and mean squared error (approximated till 3 decimal places) for training and validation dataset for EC-Earth3-CC GCM model (the highest values for each column have been highlighted in bold)

Training				
Epoch	Average sample-wise correlation	Average sample-wise MSE	Average spatial correlation	Average spatial MSE
Training				
0	0.238	128.829	0.664	138.699
40	0.498	1.440	0.992	2.282
80	0.509	1.154	0.997	1.296
120	0.526	0.475	0.998	0.562
160	0.531	0.487	0.998	0.564
200	0.533	0.446	0.998	0.560
240	0.537	0.430	0.998	0.508
280	0.541	0.430	0.998	0.501
320	0.542	0.420	0.998	0.485
360	0.542	0.413	0.998	0.480
400	0.544	0.411	0.998	0.475
Validation				
0	0.306	120.821	0.704	131.276
40	0.571	1.122	0.993	1.946
80	0.578	1.298	0.998	1.445
120	0.597	0.249	0.999	0.298
160	0.603	0.181	0.999	0.220
200	0.604	0.223	0.999	0.303
240	0.608	0.208	0.999	0.258
280	0.612	0.232	0.999	0.260
320	0.612	0.186	0.999	0.222
360	0.611	0.246	0.999	0.288
400	0.611	0.234	0.999	0.264

4.2. Discriminator model

The discriminator is a CNN that performs binary classification on SST data distinguishing between real and generated samples. As illustrated in Figure 2(a), it shares a similar CNN block structure as the generator, using Conv2D layers only to downsample the input.

As in Figure 2(b), the discriminator takes both the input SST anomaly and either the real or generated SST data as input. These are concatenated and passed through initial Conv2D and ReLU layers, followed by five CNN blocks. The final Conv2D layer and sigmoid activation output the probability of the SST data being real or fake. By competing against the generator, the discriminator improves its ability to better differentiate between real and fake samples.

4.3. Adversarial learning

The generator and discriminator models are trained jointly through an adversarial learning approach. The objective of the generator is to synthesize increasingly realistic simulated SST data to fool the discriminator, while the discriminator aims to distinguish between real and generated SST. As the training

Table 2. Epoch-wise updation of the correlations and mean squared error (approximated till 3 decimal places) for training and validation dataset for the CMCC-CM2-SR5 model

Epoch	Training			
	Average sample-wise correlation	Average sample-wise MSE	Average spatial correlation	Average spatial MSE
	Training			
0	0.493	6.219	0.975	8.954
40	0.571	0.424	0.999	0.489
80	0.575	0.260	0.999	0.294
120	0.576	0.266	0.100	0.299
160	0.576	0.269	0.999	0.306
200	0.576	0.266	0.999	0.298
240	0.574	0.306	0.999	0.345
260	0.577	0.263	0.999	0.295
280	0.576	0.274	0.999	0.311
320	0.574	0.278	0.999	0.311
360	0.576	0.252	0.999	0.284
400	0.574	0.353	0.999	0.430
	Validation			
0	0.527	7.809	0.974	11.008
40	0.612	0.782	0.999	0.890
80	0.614	0.336	0.999	0.390
120	0.613	0.269	0.999	0.321
160	0.614	0.347	0.999	0.402
200	0.613	0.242	0.999	0.291
240	0.613	0.398	0.999	0.445
260	0.615	0.197	0.999	0.235
280	0.614	0.227	0.999	0.259
320	0.612	0.260	0.999	0.306
360	0.613	0.336	0.999	0.389
400	0.613	0.223	0.999	0.285

progresses, both networks improve through this adversarial competition (Figure 3). Each input SST image is associated with 10 different ensemble members of simulated SST data. This aids the generator in learning to capture the variability across ensembles given the same input.

Both models are optimized using the Adam optimizer with a learning rate of 0.00008. The discriminator loss is the binary cross-entropy between predictions and true labels for real–fake data. The generator loss combines binary cross-entropy for fooled predictions and Huber loss between generated and real SST. A batch size of 64 is used. To leverage multiple ensembles, we adopt a cyclic training strategy. Each epoch trains on a different ensemble of batch size 64, enabling the model to incrementally learn the distinct features of each. This mimics the transfer learning strategy to capture the inter-ensemble variability.

Training continues for 400 epochs(as after that overfitting starts), with model weights being saved at every 20 epochs. The best weights are selected by evaluating the temporal correlation, spatial correlation, and discriminator accuracy on the validation set.

5. Experimental results

To evaluate model generalization, we have trained and validated our model on mutually exclusive training and validation datasets. Training ensemble members and validating the ensemble mean also test the generalization ability across the ensembles.

Figures 4 and 6 show the evolution of the training data generation for the EC-Earth3-CC and CMCC-CM2-SR5 models till epoch numbers 320 and 260, respectively, as our GAN model shows the highest correlation in these epochs. Here, mean maps are calculated by taking the grid-wise mean of all the individual SST maps. The correlation maps are generated by calculating the grid-wise Pearson correlation coefficient between the generated and original maps, while the mean squared error (MSE) maps are generated by calculating the MSE for each grid cell between the generated and original maps. As shown in Tables 1 and 2, metric values are calculated by averaging the values across all grid cells of the respective metric maps. In contrast, the spatial metric values are computed by considering the correlation and MSE between the mean generated and original maps, effectively evaluating the overall spatial patterns. According to the tables, in early epochs, the generated SST exhibits a low correlation with real data, as

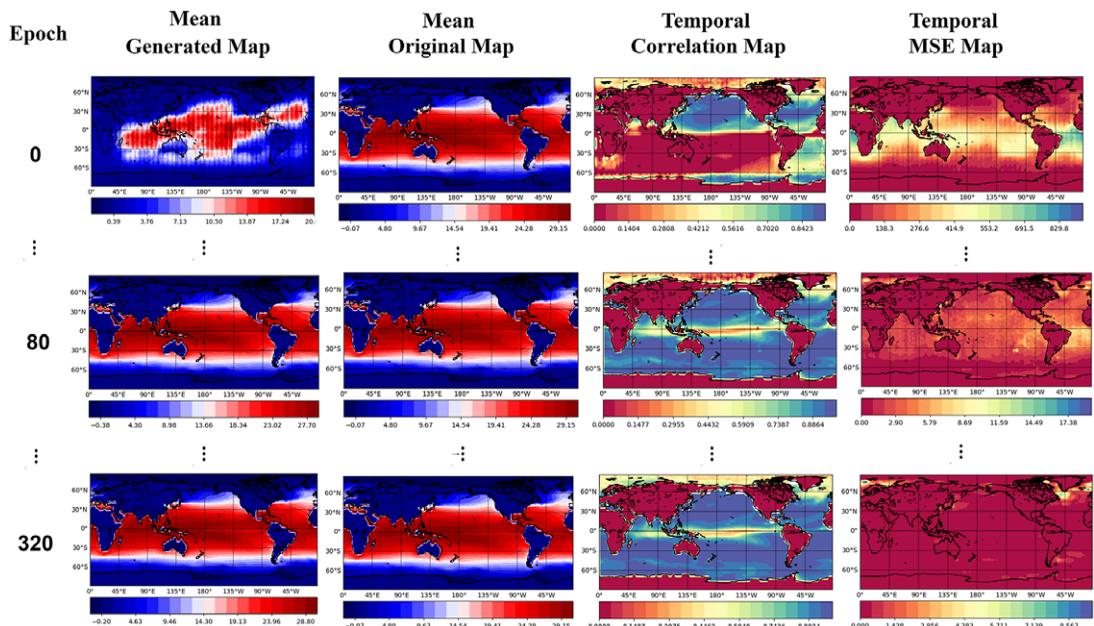


Figure 5. Evolution of the generator on validation data for the EC-Earth3-CC GCM model.

evident from the high values of the MSE between the generated and the original maps. This can also be understood visually by observing the sample-wise correlation and MSE maps. The random initialization of the weights improves substantially by epoch 80 although as per the correlation map, the prediction of the equatorial and polar regions remains challenging. This is because, in the equatorial region, all the waves of the tropical regions come together because of the Inter Tropical Convergence Zone (ITCZ; Schneider *et al.* (2014)) varying the characteristics of that region from the other regions of the ocean. For

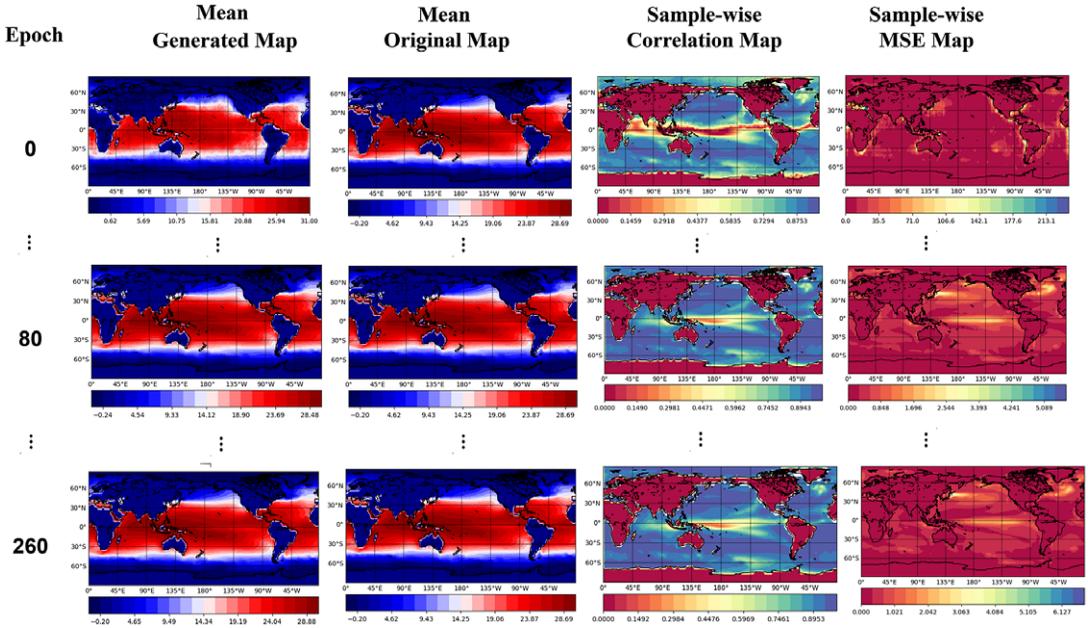


Figure 6. Evolution of the generator on training data for CMCC-CM2-SR5 GCM model.

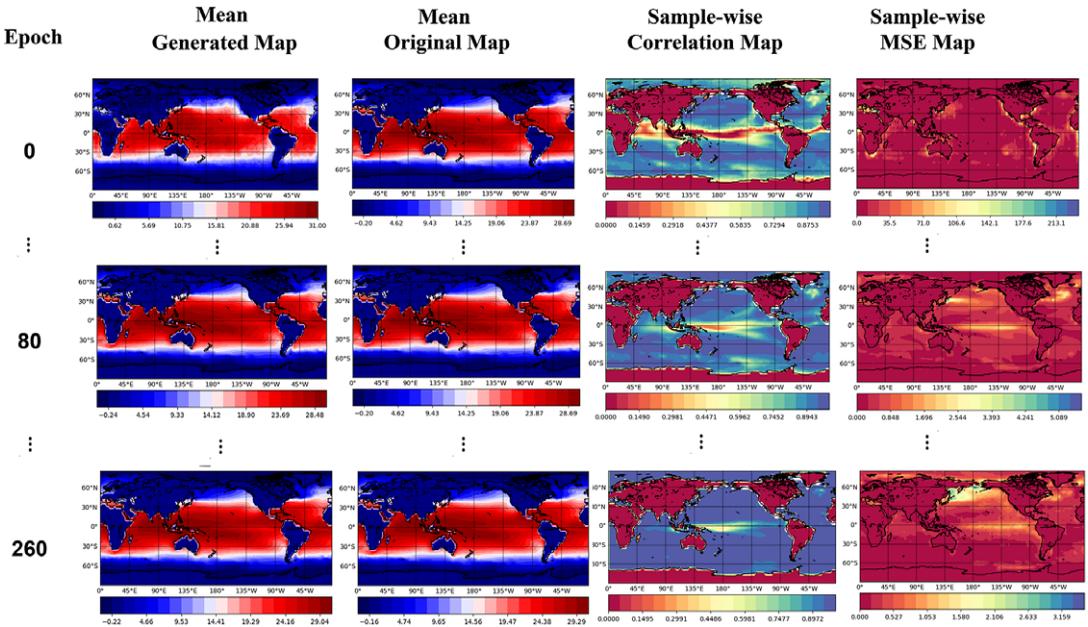


Figure 7. Evolution of the generator on validation data CMCC-CM2-SR5 GCM model.

the polar regions, very low SST makes the region hard to predict. The magnitude differences also decrease drastically. By epoch 320 (for EC-Earth3-CC) and 260 (for CMCC-CM2-SR5), when the model peaks at the correlations 0.542 (0.577) and 0.998 (0.999), even these difficult regions achieve the correlation value nearly to 0.5. The grid-wise magnitude differences also come down to almost 0 in most of the grid locations, making the average MSEs 0.420 (0.263) and 0.485 (0.295), respectively.

The evolution of the validation set in Figures 5 and 7 follow a similar trajectory, with temporal and spatial correlation improving from 0.306 (0.527) and 0.70 (0.974) at epoch 0 to 0.612 (0.615) and 0.999 (0.999) at epochs 320 and 260 respectively. Even the MSE maps have also followed the same by improving the temporal and spatial values. Unlike the training set, in the validation set the equatorial and polar correlations reach a value of more than 0.5 by the best epochs, indicating the ability of the model to learn the capture ensemble variability. We have also shown some sample SST maps (January, February, and March months of the year 2001) from the EC-Earth3-CC model and our proposed GAN-based model in Figure 8.

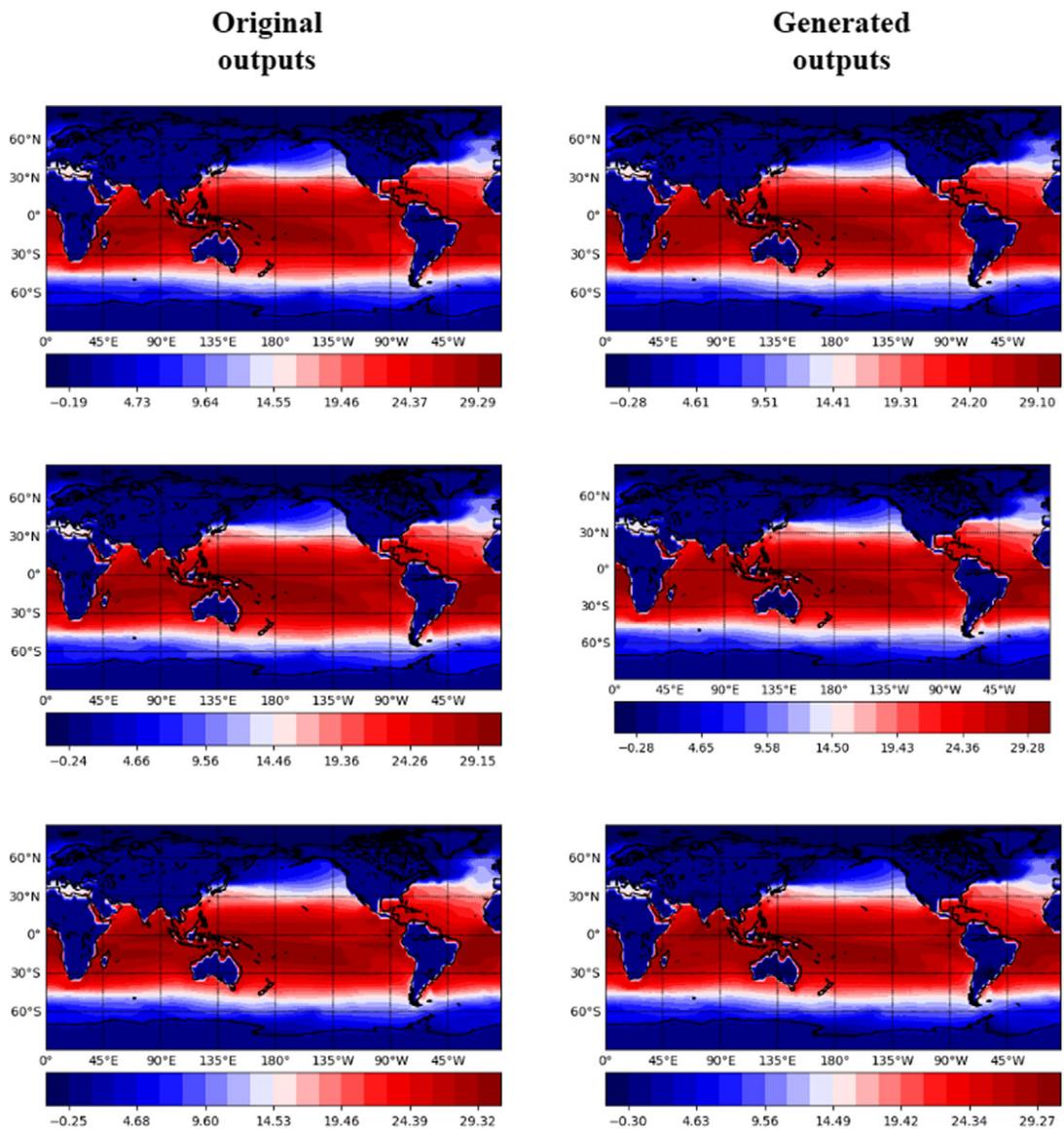


Figure 8. Sample outputs from EC-Earth3-CC GCM model at epoch 320.

In summary, adversarial training enables the model to incrementally generate increasingly realistic SST data over epochs, as quantified by the improved values of spatial and temporal/sample-wise correlation and MSEs (Tables 1 and 2). Validating on an unseen time period and the ensemble mean demonstrates the ability of the model to generalize across both time and ensembles. The proposed training strategy effectively captures ensemble variability in simulated SST generation.

6. Conclusion

Physics-based climate models like those in the CMIP6 have become indispensable tools for climate studies by providing extensive simulated datasets to address critical data gaps. To overcome its hardware limitations, we have developed a data-driven generative modeling approach for simulating SST based on Pix2pix cGAN. Our proposed conditional GAN model learns to synthesize realistic SST data simply from observed SST, bypassing the need to explicitly solve complex physical equations. Critically, we have trained and validated the GAN on entirely distinct time periods across all ensemble members of the EC-Earth3-CC and CMCC-CM2-SR5 GCM models. The skill of the model in generalizing across both temporal and spatial dimensions underscores the viability of using generative machine learning models for efficient and accessible climate simulation. This approach can be extended in the future to generate other oceanic variables like Isothermal layers, sea level pressure, *and so forth* furthermore this model can be further extended to develop a predictive approach for simulated data.

Open peer review. To view the open peer review materials for this article, please visit <http://doi.org/10.1017/eds.2024.38>.

Acknowledgments. We are grateful to Devabrat Sharma (Research Scholar at IASST, Guwahati) and Prof. Bhupendra Nath Goswami (Cotton University) for technical discussions related to climate science and CMIP6 simulations.

Author contribution. The first author has performed the experiments and first draft of the paper. The second author has defined the problem, provided technical advice, and vetted the paper.

Competing interest. The authors declare no competing interests exist.

Data availability statement. The SODA 2.2.4 observed data and CMIP6 dataset are freely available in the following links, apdrc.soest.hawaii.edu/datadoc/soda2.2.4.php, esgf-data.dkrz.de/projects/cmip6-dkrz/ respectively.

Funding statement. This research has been partially funded by the ISIRD Grant to Adway Mitra by Sponsored Research and Industrial Consultancy (SRIC), IIT Kharagpur, Grant No. IIT/SRIC/ISIRD/2020–2021/11.

References

- Bochenek B and Ustrnul Z** (2022) Machine learning in weather prediction and climate analyses—applications and perspectives, *Atmosphere* 13(2), 180.
- Endalie D, Haile G and Taye W** (2022) Deep learning model for daily rainfall prediction: case study of Jimma, Ethiopia, *Water Supply* 22(3), 3448–3461.
- Endris HS, Lennard C, Hewitson B, et al.** (2019) Future changes in rainfall associated with ENSO, IOD and changes in the mean state over Eastern Africa, *Climate Dynamics* 52, 2029–2053.
- Eyring V, Bony S, Meehl GA, et al.** (2016) Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geoscientific Model Development* 9(5), 1937–1958.
- Fasullo JT, Otto-Bliesner BL and Stevenson S** (2018) ENSO's changing influence on temperature, precipitation, and wildfire in a warming climate, *Geophysical Research Letters* 45(17), 9216–9225.
- Flato G, Marotzke J, Abiodun B, et al.** (2014) *Evaluation of climate models, Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 741–866.
- Gill AE** (2016) Atmosphere–ocean dynamics, *Elsevier*
- Giese BS and Ray S** (2011) El Niño variability in simple ocean data assimilation (SODA), 1871–2008, *Journal of Geophysical Research: Oceans* 116(C2).
- Goswami BN, Chakraborty D, Rajesh PV, et al.** (2022) Predictability of South-Asian monsoon rainfall beyond the legacy of Tropical Ocean Global Atmosphere program (TOGA), *npj Climate and Atmospheric Science* 5(1), 58.
- Gupta R, Mustafa M and Kashinath K** (2020) Climate-Style GAN: Modeling Turbulent Climate Dynamics Using Style-GAN, *NeurIPS AI for Earth Science Workshop* pp. 1–5.

- Haq DZ, Novitasari DCR, Hamid A**, et al. (2021) Long short-term memory algorithm for rainfall prediction based on El-Nino and IOD data, *Procedia Computer Science* 179, 829–837.
- Ji HK, Mirzaei M, Lai SH**, et al. (2024) Implementing generative adversarial network (GAN) as a data-driven multi-site stochastic weather generator for flood frequency estimation, *Environmental Modelling and Software* 172, 105896.
- Karoly DJ** (1989) Southern hemisphere circulation features associated with El Niño-southern oscillation events, *Journal of Climate* 2(11), 1239–1252.
- Krishnamurthy L and Krishnamurthy VJCD** (2014) Influence of PDO on South Asian summer monsoon and monsoon–ENSO relation, *Climate Dynamics* 42, 2397–2410.
- Le PV, Randerson JT, Willett R**, et al. (2023) Climate-driven changes in the predictability of seasonal precipitation, *Nature communications* 14(1),3822.
- Niazkar M, Goodarzi MR, Fatehifar A**, et al. (2023) Machine learning-based downscaling: Application of multi-gene genetic programming for downscaling daily temperature at Dogonbadan, Iran, under CMIP6 scenarios, *Theoretical and Applied Climatology* 151(1–2), 153–168.
- Power S, Delage F, Chung C**, et al. (2013) Robust twenty-first-century projections of El Niño and related precipitation variability, *Nature* 502(7472), 541–545.
- Richter I and Tokinaga H** (2020) An overview of the performance of CMIP6 models in the tropical Atlantic: mean state, variability, and remote impacts, *Climate Dynamics* 55, 2579–2601.
- Rivera JA and Arnould G** (2020) Evaluation of the ability of CMIP6 models to simulate precipitation over Southwestern South America: climatic features and long-term trends (1901–2014), *Atmospheric Research* 241,104953.
- Schneider T, Bischoff T and Haug GH** (2014) Migrations and dynamics of the intertropical convergence zone, *Nature* 513(7516), 45–53.
- Sharma SCM and Mitra A** (2022) ResDeepD: A residual super-resolution network for deep downscaling of daily precipitation over India, *Environmental Data Science*, 1, e19.
- Sharma D, Das S, Chakraborty D**, et al. (2023) Physics-based 18-month lead forecast of Indian Summer Monsoon Rainfall, *Physical Sciences - Article* doi: <https://doi.org/10.21203/rs.3.rs-2923543/v1>
- Singh D, Vardhan M, Sahu R**, et al. (2023) Machine-learning-and deep-learning-based streamflow prediction in a hilly catchment for future scenarios using CMIP6 GCM data, *Hydrology and Earth System Sciences* 27(5), 1047–1075.
- Sleeman J, Chung D, Gnanadesikan A**, et al. (2023) A generative adversarial network for climate tipping point discovery (TIP-GAN), *arXiv preprint arXiv: 2302. 10274*.
- Sun M, Chen L, Li T**, et al. (2023) CNN-based ENSO forecasts with a focus on SSTA zonal pattern and physical interpretation, *Geophysical Research Letters* 50(20), e2023GL105175.
- Stevenson S, Fox-Kemper B, Jochum M**, et al. (2012) Will there be a significant change to El Niño in the twenty-first century? *Journal of Climate* 25(6), 2129–2145.
- Wang B, Luo X and Liu J** (2020) How robust is the Asian precipitation–ENSO relationship during the industrial warming period (1901–2017)? *Journal of Climate* 33(7), 2779–2792.
- Xu K, Tam CY, Zhu C**, et al. (2017) CMIP5 projections of two types of El Niño and their related tropical precipitation in the twenty-first century, *Journal of Climate* 30(3), 849–864.
- Yadav RK, Srinivas G and Chowdary JS** (2018) Atlantic Niño modulation of the Indian summer monsoon through Asian jet, *npj Climate and Atmospheric Science* 1(23).
- Yan Z, Wu B, Li T**, et al. (2020) Eastward shift and extension of ENSO-induced tropical precipitation anomalies under global warming, *Science Advances* 6(2), eaax4177.
- Yang S, Li Z, Yu JY**, et al. (2018) El Niño–Southern Oscillation and its impact in the changing climate, *National Science Review* 5(6), 840–857.
- Yan Y, Wang H, Li G**, et al. (2022) Projection of future extreme precipitation in china based on the cmip6 from a machine learning perspective, *Remote Sensing* 14(16), 4033.