Improving global high-resolution Earth system model simulations of precipitation with generative adversarial networks

Philipp Hess[1,2], Stefan Lange[2], and Niklas Boers[1,2,3]

[1]Earth System Modelling, School of Engineering & Design, Technical University of Munich, Munich, Germany
[2]Potsdam Institute for Climate Impact Research, Member of the Leibniz Association, Potsdam, Germany
[3]Global Systems Institute and Department of Mathematics, University of Exeter, Exeter, UK

Abstract

Precipitation extremes are expected to become stronger and more frequent in response to anthropogenic global warming. Accurately projecting the ecological and socioeconomic impacts is an urgent task. Impact models are developed and calibrated with observation-based data but rely on Earth system model (ESM) output for future scenarios. ESMs, however, exhibit significant biases in their output because they cannot fully resolve complex cross-scale interactions of processes that produce precipitation cannot. State-of-the-art bias correction methods only address errors in the simulated frequency distributions, locally at every individual grid cell. Improving unrealistic spatial patterns of the ESM output, which would require spatial context, has not been possible so far. Here we show that a post-processing method based on physically constrained generative adversarial networks (GANs) can correct biases of a state-of-the-art global ESM both in local frequency distributions and in the spatial patterns at once. While our method improves local frequency distributions similarly well to a gold-standard ESM bias-adjustment framework, it strongly outperforms existing methods in correcting spatial patterns. Our study highlights the importance of physical constraints in neural networks for out-of-sample predictions in the context of climate change.

1 Introduction

The accurate projection of precipitation changes due to a warming of the atmosphere is important for reliably assessing the impacts of anthropogenic greenhouse gas emissions. Precipitation and particularly its extremes have large socio-economic impact (Kotz et al., 2022). While precipitation and the frequency of extreme events are expected to increase with the water-holding capacity of the warming atmosphere globally, local trends show heterogeneous changes (Ali et al., 2018; Traxl et al., 2021).

Impact models are built to assess the hydrological, ecological, or socioeconomic impacts and are typically developed and tuned using observation-based data as input. For future projections, impact models rely on simulations from comprehensive Earth system model (ESM) simulations that integrate the governing physical equations on discretized grid. The numerical discretization, at comparable coarse grid size of typically $50-200$km is needed to keep computational costs of global long-term simulations within reasonable bounds. In turn, ESMs require approximations of processes on the smaller, unresolved spatial scales in the form of parameterizations (Balaji et al., 2022).

These approximations can lead to systematic errors in the model precipitation, which results from complex interactions of processes across many scales. The bias often manifests as

overly blurred images that do not resolve the characteristic high-frequency variability of precipitation both in space and in time, and as an under- or overestimation of extremes.

Statistical post-processing methods, such as quantile mapping (Cannon et al., 2015; Lange, 2019), can adjust biases in the frequency distribution for each model grid cell individually. Spatial patterns and correlations between grid cells are, however, not taken into account.

Neural networks, on the other hand, can use spatial context and have been successfully applied to weather prediction and post-processing tasks (Grönquist et al., 2020; Rasp and Thuerey, 2021; Ravuri et al., 2021; Pathak et al., 2022). Applying them to climate simulations, however, poses different challenges.

Model simulations and observations diverge rapidly due to the chaotic nature of the atmosphere, which limits weather forecast horizons. Nevertheless, in the context of numerical weather prediction, forecasts can be directly compared to corresponding ground truth. In the context of climate projections, however, simulated spatial fields of ESMs are *unpaired* with observations. There is no observational ground truth for a given snapshot of an ESM simulation, and one has to focus on statistical characteristics instead. This is one of the fundamental differences between weather prediction and climate modelling.

Given these unpaired samples, cycle-consistent generative adversarial networks (Zhu et al., 2017; Hess et al., 2022) have emerged as powerful tools that can translate the "style" (e.g. the small-scale variability) of precipitation fields from observations to the ESM output. The ESM output is thereby made more realistic while the large-scale patterns (or "content") are preserved (see Fig. A1 for examples).

Our work extends previous studies in two key regards: (i) we apply the GAN-processing to global high-resolution simulations from an ESM in the Coupled Model Intercomparison Project Phase 6 (CMIP6) that is used to inform policymakers (IPCC, 2021), (ii) we compare the performance against a state-of-the-art bias correction framework ISIMIP3BASD (Lange, 2019) based on quantile mapping (QM), which can be considered the operational state-of-the-art and thus represents a strong baseline.

2 Methodology

2.1 Training data

We use daily global precipitation output from the GFDL-ESM4 (Krasting et al., 2018) with 1° horizontal spatial resolution. The GFDL-ESM4 model simulates the Earth system with a general circulation model for the atmosphere and ocean and includes vegetation, biosphere and ice sheets as model components. The ground truth W5E5v2 reanalysis product (Lange et al., 2021) combines high-fidelity weather simulations with multiple observation sources through data assimilation routines. We split the dataset into three periods for training (1950-2000), validation (2001-2003), and testing (2004-2014).

2.2 Cycle-consistent Generative adversarial networks

We refer to Zhu et al. (2017); Hess et al. (2022) for a detailed description of the CycleGAN architecture. In a nutshell, the CycleGAN (Zhu et al., 2017) consists of two generator-discriminator network pairs. The generators G and F learn inverse transformations of images between two domains X and Y. The discriminators then have to classify whether images are generated or drawn from a target distribution (see Goodfellow et al. (2014; 2020) for details). The two generators allow defining a cycle-consistency regularization that constrains the training of the generators to preserve the overall content of the images, i.e., with

$$\mathcal{L}_{\text{cycle}}(G, F) = \mathbb{E}_{x\sim p_x(x)}[||F(G(x)) - x||_1] + \mathbb{E}_{y\sim p_y(y)}[||G(F(y)) - y||_1]. \quad (1)$$

It measures the error during a translation cycle of an image to the other domain and back. The full loss function and training procedure are outlined in appendix B.

2.3 Physical constraints

We constrain the generator network to preserve the global precipitation sum from the ESM during the CycleGAN-based post-processing. The constraint is implemented as an additional network layer that rescales the generator output with $\tilde{y}_i = y_i \left(\sum_{i=1}^{N_{\text{grid}}} x_i\right) / \left(\sum_{i=1}^{N_{\text{grid}}} y_i\right)$, where x_i and y_i are grid cell values of the generator input and output, respectively, and N_{grid} is the number of grid cells. To enforce precipitation not to have negative values, we additionally apply a ReLU activation function at the final generator layer.

2.4 Baselines

We compare the CycleGAN performance, with and without the physical constraint, to the ISIMIP3BASD bias correction framework (Lange, 2019) that is based on quantile mapping (see Cannon et al. (2015) for more details), on bias-correcting the GFDL-ESM4 model w.r.t. the W5E5 ground truth. We further construct a combination of the two methods, where the output of the physically constrained GAN is processed by the ISIMIP3BASD method.

3 Results

We evaluate the model performance on two tasks and time scales: (i) correcting summary statistics on decadal time scales with focus on frequency distributions and (ii) improving spatial fields of daily precipitation.

3.1 Correcting summary statistics

We first compare the performances using latitude profiles of precipitation averaged in time and over longitudes (see Fig. 1a). The analysis shows that the GFDL-ESM4 model exhibits a pronounced peak at around 7°S latitude - this double Intertropical Convergence Zone (ITCZ) is a well-known bias in many models (Tian and Dong, 2020). All post-processing methods are able to reduce this bias, with ISIMIP3BASD and the GAN combination achieving the best results. We further compute global histograms of relative frequencies, which show that the GFDL-ESM4 model overestimates the occurrence of extreme events (see Fig. 1b); note that this is relatively uncommon and past models typically underestimated extremes. All considered post-processing methods can correct the frequencies of extreme events in the tail. The global mean absolute error for annual and seasonal time series are reported in appendix C, where the ISIMIP3BASD slightly outperforms the GAN post-processing. The combined approach achieves the best overall results.

3.2 Improving spatial patterns

For evaluating improvements in spatial patterns, we compute radially averaged power spectral densities (RAPSDs) for each spatial field, and average over the test set following Ravuri et al. (2021) (see Fig. 2a). The RAPSDs show which spatial scales in the precipitation fields are represented realistically compared to the ground truth. To highlight the differences, we compute the absolute error between the logarithmically rescaled spectra (see Fig. 2b). The GFDL-ESM4 model shows large deviations at wavelengths below 400km. The GANs clearly outperform the ISIMIP3BASD baseline, the latter exhibiting larger errors towards smaller wavelengths.

We further compute the fractal dimension using box-counting (Lovejoy et al., 1987; Meisel et al., 1992) to evaluate the realisticness of the generated images (see Fig. A2). It essentially characterizes how "wiggly" or fractal the lines are that separate precipitation patterns - a key property of precipitation patterns. In agreement with the RAPSD analysis and qualitative comparisons (see Fig. A1), we find that the GAN-based methods strongly outperform the ISIMIP3BASD framework.

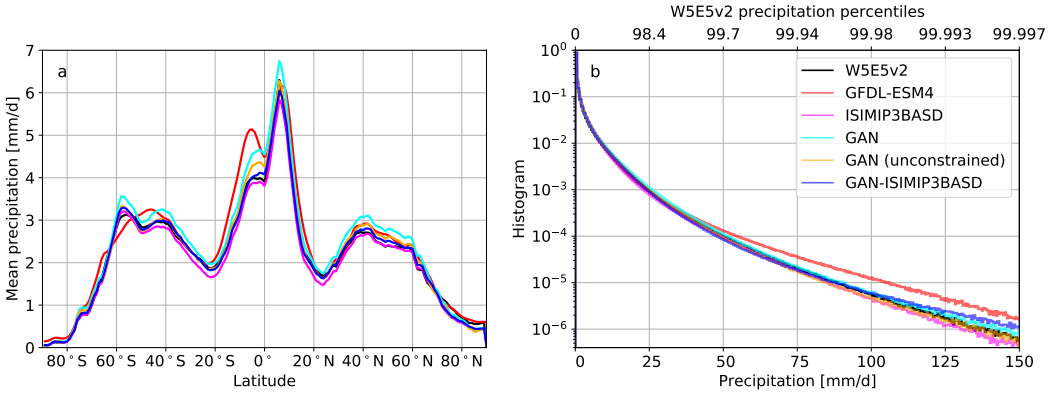

Figure 1: Evaluating temporal summary statistics. (a) The precipitation mean in time and over longitudes is shown for the W5E5v2 ground truth, GFDL-ESM4, ISIMIP3BASD, GAN, unconstrained GAN, and the GAN-ISIMIP3BASD combination. (b) Histograms of relative frequencies are shown for the different models and the ground truth. The GFDL-ESM4 model shows an over-estimation of extreme events, in the distribution's tail.

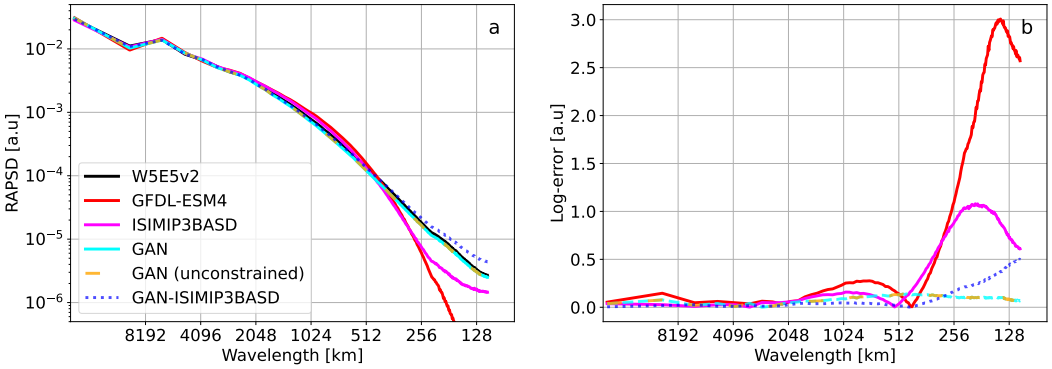

Figure 2: Evaluating spatial precipitation patterns. (a) The radially averaged power spectral density (RASPD) of precipitation fields is shown as a mean over the test set for the W5E5v2 ground truth, GFDL-ESM4, ISIMIP3BASD, GAN, unconstrained GAN, and the GAN-ISIMIP3BASD combination. (b) We compute the log-error between the W5E5v2 ground and modelled RAPSDs to show where the largest deviations occur. The CycleGANs can capture the RAPSDs accurately over all resolved scales.

4 Discussion

We apply a novel bias correction method based on physically constrained CycleGANs to global high-resolution ESM precipitation output. It can improve the simulation in the spatial dimension on short times scales and in temporal summary statistics over long periods. The former is not possible with state-of-the-art statistical post-processing methods.

We have introduced a physical constraint to preserve the global precipitation sum in the ESM model. Enforcing this consistency is motivated by the observation that large-scale average trends can be expected to be modelled well in GFDL-ESM4. As shown in Fig. A3, the constraint enables the GAN to generalize to predictions outside the training distribution, realistically reproducing trends in an extreme warming scenario. Without the constraint, physically unrealistic trends of decreasing precipitation averages in the tropics are found. The summary statistics show that the unconstrained GAN performed better over the historical period, perhaps due to the lack of constraint. Overall the combination of GAN and quantile mapping has been found to perform better than either method alone.

The main advantage of the GAN-based post-processing is that spatial patterns, like the characteristic small-scale precipitation variability, can be skillfully corrected. This is particularly important for localized extreme events. We believe there is great potential in generative methods of unpaired image-to-image translations to very similar tasks of post-processing Earth system model simulations.

Acknowledgements

The authors would like to thank the reviewers for the valuable feedback. NB and PH acknowledge funding by the Volkswagen Foundation, as well as the European Regional Development Fund (ERDF), the German Federal Ministry of Education and Research and the Land Brandenburg for supporting this project by providing resources on the high-performance computer system at the Potsdam Institute for Climate Impact Research. N.B. acknowledges funding by the European Union's Horizon 2020 research and innovation programme under grant agreement No 820970 and under the Marie Sklodowska-Curie grant agreement No. 956170, as well as from the Federal Ministry of Education and Research under grant No. 01LS2001A. SL acknowledges funding from the European Union's Horizon 2022 research and innovation programme under grant agreement no. 101081193 Optimal High Resolution Earth System Models for Exploring Future Climate Changes (OptimESM).

Appendices

A Qualitative comparison

For a qualitative comparison, precipitation fields from the reanalysis ground truth W5E5v2, GFDL-ESM4 model, CycleGAN and the ISIMIP3BASD baseline are shown in Fig. A1. The fields are shown for the same date (14th October 2014) to illustrate the unpaired nature of the training samples. The global fields have been cropped to the shown region to visualize the small-scale variability. Besides some coincidental overlap of precipitation in central Africa, the GFDL-ESM4 model precipitation does not agree with the W5E5v2 ground truth on a grid cell level - as expected. The raw GFDL-ESM4 and the ISIMIP3BASD fields appear too smooth and blurry compared to the W5E5v2 ground truth. The GAN-processed precipitation field, on the other hand, shows realistic variability on small scales that is visually indistinguishable from the observation-based W5E5v2 data.

B CycleGAN training

The generator and discriminator networks are fully convolutional and described in detail in Zhu et al. (2017). The generator uses ReLU activation functions, instance normalization, and reflection padding. The discriminator uses leaky ReLU activations with slope 0.2 instead and instance normalization. We have changed the number of residual layers in the generator network from 6 to 7.

Besides the cycle-consistency constraint in Eq. 1, another regularization term is added to the generator loss, in order to enforce an identity mapping with,

$$\begin{aligned}\mathcal{L}_{\text{ident}}(G,F) = \mathbb{E}_{x\sim p_x(x)}[||G(x)-x||_1] \\ + \mathbb{E}_{y\sim p_y(y)}[||F(y)-y||_1].\end{aligned} \tag{2}$$

Combining the two regularization terms, the full generator loss then reads

$$\begin{aligned}\mathcal{L}_{\text{Generator}} = \mathbb{E}_{x\sim p_x(x)}[(D_X(G(x))-1)^2] \\ + \mathbb{E}_{y\sim p_y(y)}[(D_Y(F(y))-1)^2] \\ + \lambda\mathcal{L}_{\text{cycle}}(G,F) + \tilde{\lambda}\mathcal{L}_{\text{ident}}(G,F),\end{aligned} \tag{3}$$

where, D_X and D_Y are the two discriminator networks, and λ and $\tilde{\lambda}$ are two hyperparameters, set to 10 and 5, respectively. The corresponding loss term for the discriminator

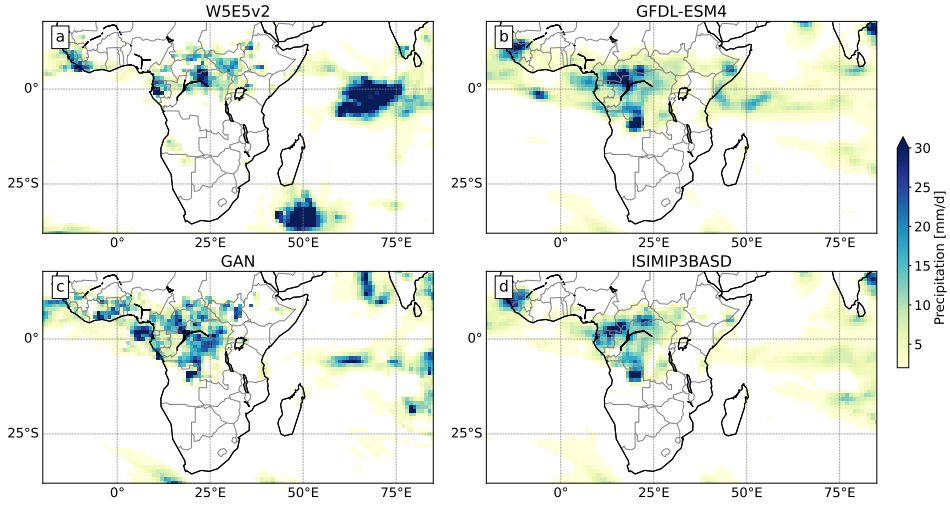

Figure A1: A qualitative comparison of daily precipitation fields for the same date (14th October 2014) . The precipitation fields from (a) the reanalysis ground truth W5E5v2, (b) GFDL-ESM4 model, (c) GAN and (d) ISIMIP3BASD baseline are shown for a cropped region to visualize the difference in small-scale variability. As expected from a climate simulation, the local "weather" on daily time scales does not match the observation-based ground truth W5E5v2, i.e., the fields are *unpaired*. The GAN transforms the ESM output (b) into a much more realistic representation (c) that preserves the large-scale structures ("content") but improves the small-scale variability ("style").

networks is given by

$$\mathcal{L}_{\text{Discriminator}} = \mathbb{E}_{y\sim p_y(y)}[(D_Y(y)-1)^2] + \mathbb{E}_{x\sim p_x(x)}[(D_X(G(x)))^2] \tag{4}$$

$$+ \mathbb{E}_{x\sim p_x(x)}[(D_X(x)-1)^2] + \mathbb{E}_{y\sim p_y(y)}[(D_Y(F(y)))^2]. \tag{5}$$

We use the ADAM (Kingma and Ba, 2015) optimizer with a learning rate of $2e^{-4}$ and alternatively update the networks. We train the network for 350 epochs and a batch size of 1 (following (Zhu et al., 2017)), saving model checkpoints every other epoch.

C Global error statistics

We evaluate the post-processing methods to improve the mean error bias of the ESM, i.e., with

$$\text{ME}(\hat{y}, y) = \frac{1}{T}\sum_{t=1}^{T}\hat{y}_t - \frac{1}{T}\sum_{t=1}^{T} y_t, \tag{6}$$

where $\hat{y}_t$ and y_t are the modelled and ground truth precipitation values for a single grid cell at time t. The ME metric can be applied to the unpaired time series since it only compares long-term averages. The resulting pixel-wise ME bias is summarized in Tab. 1 as globally averaged absolute mean errors. The constrained GAN can significantly reduce the ESM bias for annual and seasonal time series. Removing the constraint improves the GAN performance, achieving result close to the ISIMIP3BASD baseline. The combination of the two methods outperforms each individual method.

D Fractal dimensions of precipitation fields

Quantifying how realistic generated spatial precipitation fields are is challenging (Ravuri et al., 2021), especially given their unpaired nature in the climate modelling context of this study. Typical scores from image generation, such as the Fréchet inception distance (FID) (Salimans et al., 2016; Heusel et al., 2017), cannot straightforwardly be applied to geospatial data. Here, we use the fractal dimension of binary precipitation fields to measure

Table 1: The globally averaged absolute value of the mean error (Eq. 6) between the W5E5v2 ground truth the different post-processing methods for annual and seasonal time series (in [mm/day]). The relative improvement over the raw GFDL-ESM4 climate model output is shown as percentages for each method.

Season	GFDL-ESM4	ISIMIP3-BASD	%	GAN	%	GAN (unconst.)	%	GAN-ISIMIP3-BASD	%
Annual	0.535	0.217	59.4	0.328	38.7	0.265	50.5	**0.195**	**63.6**
DJF	0.634	0.321	49.4	0.395	37.7	0.371	41.5	**0.308**	**51.4**
MAM	0.722	0.314	56.5	0.419	42.0	0.378	47.6	**0.285**	**60.5**
JJA	0.743	0.289	61.1	0.451	39.3	0.357	52.0	**0.280**	**62.3**
SON	0.643	0.327	49.1	0.409	36.4	0.362	43.7	**0.306**	**52.4**

the characteristic small-scale variability of the fields. The global precipitation fields are first converted into binary data using a quantile threshold. The fractal dimension D_{fractal} is then computed using the box-counting method (Lovejoy et al., 1987; Meisel et al., 1992), i.e., by counting the number of squares N of side length s that cover the boundary separating the pattern,

$$D_{\text{fractal}} = \frac{\log(N_{\text{squares}})}{\log(1/s)}. \tag{7}$$

The results are shown in Fig. A2 for different quantile thresholds. We find that the GANs accurately capture the fractal dimension across all quantiles, strongly outperforming the ISIMIP3BASD baseline.

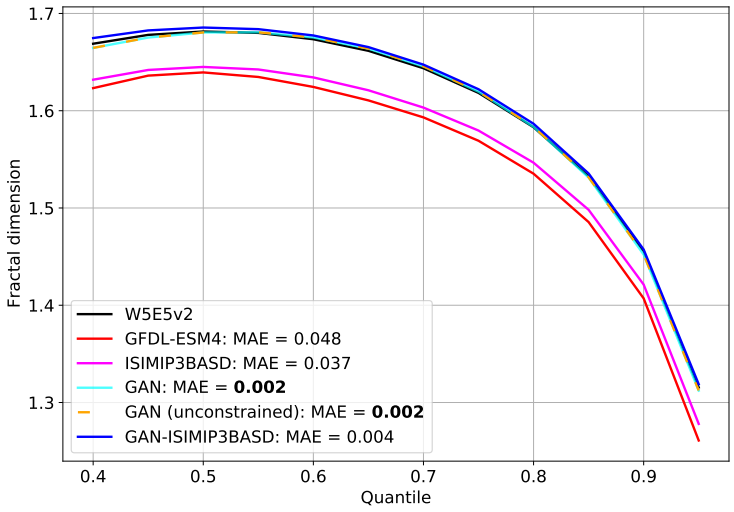

Figure A2: The fractal dimension of binary global precipitation fields is compared as averages for different quantile thresholds. The GAN can accurately reproduce the fractal dimension of the W5E5v2 ground truth over all quantile thresholds, clearly outperforming the ISIMIP3BASD baseline. A higher fractal dimension corresponds to more "wiggly" boundaries of the binary precipitation fields. The mean absolute error w.r.t. the ground truth W5E5 is shown in the legend.

E Climate non-stationarity

To investigate the ability of the generator network to generalize to new transient climate states not seen during the training on historical observations, we use the SSP5-8.5 emission scenario from 2015 to 2100 for the GFDL-ESM4 model. This scenario is one of the most extreme ones typically considered for future projections.

As seen in Fig. A3, the ESM projects a significant increase in global mean precipitation in the scenario for this century. This can be expected due to increased water saturation vapour pressure with higher temperatures (Berg et al., 2013; Guerreiro et al., 2018). By design, the constraint lets the generator follow the global mean of the ESM. Without the constraint, on the other hand, the generator is not able to accurately capture the future trend in global mean precipitation (see Fig. A3a).

On a more regional level in the tropics (from 23°S to 23°N), the constrained generator has the freedom to correct the ESM simulations (see Fig. A3b). As shown in Fig. 1a, the CycleGAN reduces the mean rainfall in the tropics where the ESM overestimates the double-ITCZ rainband. The unconstrained network - contrary to the GFDL-ESM4 model - projects a physically implausible scenario of decreasing precipitation in the tropics.

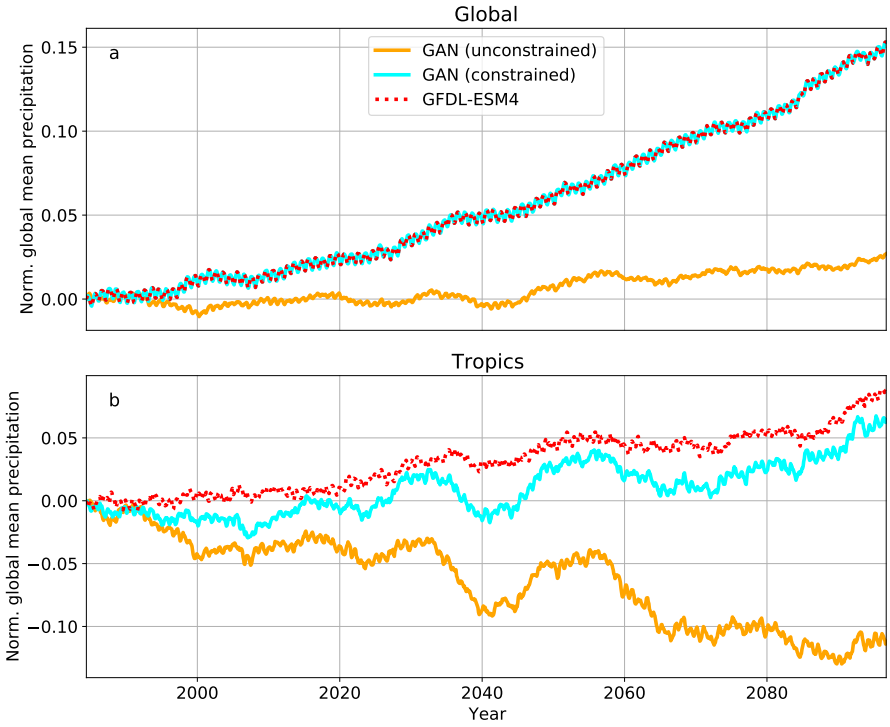

Figure A3: Evaluating the generalization of the GAN to a warming scenario (CMIP6, SSP5-8.5) unseen during training. (a) The precipitation averaged globally and over two consecutive years is shown for GFDL-ESM4, the constrained and the unconstrained GAN. The time series have been normalized to start at zero for better comparisons of the trends. The unconstrained network underestimates the trend of increasing global mean precipitation. (b) Same as (a) but only for spatial averages over the tropics from 23°S to 23°N. The unconstrained model predicts decreasing precipitation sums which is physically inconsistent with the expansion of the water-holding capacity in the warming atmosphere.

References

Maximilian Kotz, Anders Levermann, and Leonie Wenz. The effect of rainfall changes on economic production. *Nature*, 601(7892):223–227, January 2022. ISSN 1476-4687. doi: 10.1038/s41586-021-04283-8. URL https://www.nature.com/articles/s41586-021-04283-8$. Number: 7892 Publisher: Nature Publishing Group.

Haider Ali, Hayley J. Fowler, and Vimal Mishra. Global Observational Evidence of Strong Linkage Between Dew Point Temperature and Precipitation Extremes. *Geophysical Research Letters*, 45(22):12,320–12,330, 2018. ISSN 1944-8007. doi: 10.1029/2018GL080557. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2018GL080557. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2018GL080557.

Dominik Traxl, Niklas Boers, Aljoscha Rheinwalt, and Bodo Bookhagen. The role of cyclonic activity in tropical temperature-rainfall scaling. *Nature Communications*, 12(1):6732, November 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-27111-z. URL https://www.nature.com/articles/s41467-021-27111-z. Number: 1 Publisher: Nature Publishing Group.

V. Balaji, Fleur Couvreux, Julie Deshayes, Jacques Gautrais, Frédéric Hourdin, and Catherine Rio. Are general circulation models obsolete? *Proceedings of the National Academy of Sciences*, 119(47):e2202075119, November 2022. doi: 10.1073/pnas.2202075119. URL https://www.pnas.org/doi/10.1073/pnas.2202075119. Publisher: Proceedings of the National Academy of Sciences.

Alex J. Cannon, Stephen R. Sobie, and Trevor Q. Murdock. Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes? *Journal of Climate*, 28(17):6938–6959, September 2015. ISSN 0894-8755, 1520-0442. doi: 10.1175/JCLI-D-14-00754.1. URL https://journals.ametsoc.org/view/journals/clim/28/17/jcli-d-14-00754.1.xml. Publisher: American Meteorological Society Section: Journal of Climate.

Stefan Lange. Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0). *Geoscientific Model Development*, 12(7):3055–3070, 2019. ISSN 19919603. doi: 10.5194/gmd-12-3055-2019.

Peter Grönquist, Chengyuan Yao, Tal Ben-Nun, Nikoli Dryden, Peter Dueben, Shigang Li, Torsten Hoefler, Eth Zürich, Chengyuan Yao, Tal Ben-Nun, Nikoli Dryden, Peter Dueben, Shigang Li, and Torsten Hoefler. Deep Learning for Post-Processing Ensemble Weather Forecastse, May 2020. URL https://cds.climate.copernicus.eu/.

Stephan Rasp and Nils Thuerey. Data-driven medium-range weather prediction with a Resnet pretrained on climate simulations: A new model for WeatherBench. *Journal of Advances in Modeling Earth Systems*, February 2021. ISSN 1942-2466. doi: 10.1029/2020ms002405. arXiv: 2008.08626 Publisher: American Geophysical Union (AGU).

Suman Ravuri, Karel Lenc, Matthew Willson, Dmitry Kangin, Remi Lam, Piotr Mirowski, Megan Fitzsimons, Maria Athanassiadou, Sheleem Kashem, Sam Madge, Rachel Prudden, Amol Mandhane, Aidan Clark, Andrew Brock, Karen Simonyan, Raia Hadsell, Niall Robinson, Ellen Clancy, Alberto Arribas, and Shakir Mohamed. Skilful precipitation nowcasting using deep generative models of radar. *Nature*, 597(7878):672–677, 2021. ISSN 14764687. doi: 10.1038/s41586-021-03854-z. URL http://arxiv.org/abs/2104.00954. arXiv: 2104.00954 Publisher: Springer US.

Jaideep Pathak, Shashank Subramanian, Peter Harrington, Sanjeev Raja, Ashesh Chattopadhyay, Morteza Mardani, Thorsten Kurth, David Hall, Zongyi Li, Kamyar Azizzadenesheli, Pedram Hassanzadeh, Karthik Kashinath, and Animashree Anandkumar. FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators, February 2022. URL http://arxiv.org/abs/2202.11214. arXiv:2202.11214 [physics].

Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. Unpaired Image-To-Image Translation Using Cycle-Consistent Adversarial Networks. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2223–2232, 2017. URL https://openaccess.thecvf.com/content_iccv_2017/html/Zhu_Unpaired_Image-To-Image_Translation_ICCV_2017_paper.html.

Philipp Hess, Markus Drüke, Stefan Petri, Felix M. Strnad, and Niklas Boers. Physically constrained generative adversarial networks for improving precipitation fields from Earth system models. *Nature Machine Intelligence*, 4(10):828–839, October 2022. ISSN 2522-5839. doi: 10.1038/s42256-022-00540-1. URL https://www.nature.com/articles/s42256-022-00540-1. Number: 10 Publisher: Nature Publishing Group.

IPCC. Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Technical report, Cambridge University Press, 2021. URL https://www.ipcc.ch/report/sixth-assessment-report-working-group-i/.

John P Krasting, Jasmin G John, Chris Blanton, Colleen McHugh, Serguei Nikonov, Aparna Radhakrishnan, Kristopher Rand, Niki T Zadeh, V Balaji, Jeff Durachta, Christopher Dupuis, Raymond Menzel, Thomas Robinson, Seth Underwood, Hans Vahlenkamp, Krista A Dunne, Paul P G Gauthier, Paul Ginoux, Stephen M Griffies, Robert Hallberg, Matthew Harrison, William Hurlin, Sergey Malyshev, Vaishali Naik, Fabien Paulot, David J Paynter, Jeffrey Ploshay, Brandon G Reichl, Daniel M Schwarzkopf, Charles J Seman, Levi Silvers, Bruce Wyman, Yujin Zeng, Alistair Adcroft, John P Dunne, Raphael Dussin, Huan Guo, Jian He, Isaac M Held, Larry W Horowitz, Pu Lin, P C D Milly, Elena Shevliakova, Charles Stock, Michael Winton, Andrew T Wittenberg, Yuanyu Xie,

and Ming Zhao. NOAA-GFDL GFDL-ESM4 model output prepared for CMIP6 CMIP, 2018. URL https://doi.org/10.22033/ESGF/CMIP6.1407.

Stefan Lange, Christoph Menz, Stephanie Gleixner, Marco Cucchi, Graham Weedon, Alessandro Amici, Bellouin Nicolas, Hannes Schmied, Hans Hersbach, and Chiara Cagnazzo. WFDE5 over land merged with ERA5 over the ocean (W5E5 v2.0), 2021. URL https://doi.org/10.48364/ISIMIP.342217. Version Number: 2.0.

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative Adversarial Nets. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014. URL https://proceedings.neurips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf.

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. *Communications of the ACM*, 63(11):139–144, October 2020. ISSN 0001-0782, 1557-7317. doi: 10.1145/3422622. URL https://dl.acm.org/doi/10.1145/3422622.

Baijun Tian and Xinyu Dong. The Double-ITCZ Bias in CMIP3, CMIP5, and CMIP6 Models Based on Annual Mean Precipitation. *Geophysical Research Letters*, 47(8):1–11, 2020. ISSN 19448007. doi: 10.1029/2020GL087232.

S. Lovejoy, D. Schertzer, and A. A. Tsonis. Functional Box-Counting and Multiple Elliptical Dimensions in Rain. *Science*, 235(4792):1036–1038, February 1987. doi: 10.1126/science.235.4792.1036. URL https://www.science.org/doi/abs/10.1126/science.235.4792.1036. Publisher: American Association for the Advancement of Science.

L. V. Meisel, Mark Johnson, and P. J. Cote. Box-counting multifractal analysis. *Physical Review A*, 45(10):6989–6996, May 1992. doi: 10.1103/PhysRevA.45.6989. URL https://link.aps.org/doi/10.1103/PhysRevA.45.6989. Publisher: American Physical Society.

Diederik P. Kingma and Jimmy Lei Ba. Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pages 1–15, 2015. arXiv: 1412.6980.

Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, Xi Chen, and Xi Chen. Improved Techniques for Training GANs. In D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc., 2016. URL https://proceedings.neurips.cc/paper/2016/file/8a3363abe792db2d8761d6403605aeb7-Paper.pdf.

Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL https://proceedings.neurips.cc/paper/2017/hash/8a1d694707eb0fefe65871369074926d-Abstract.html.

Peter Berg, Christopher Moseley, and Jan O. Haerter. Strong increase in convective precipitation in response to higher temperatures. *Nature Geoscience*, 6(3):181–185, March 2013. ISSN 17520894. doi: 10.1038/ngeo1731. URL www.nature.com/naturegeoscience. Publisher: Nature Publishing Group.

Selma B. Guerreiro, Hayley J. Fowler, Renaud Barbero, Seth Westra, Geert Lenderink, Stephen Blenkinsop, Elizabeth Lewis, and Xiao Feng Li. Detection of continental-scale intensification of hourly rainfall extremes. *Nature Climate Change*, 8(9):803–807, September 2018. ISSN 17586798. doi: 10.1038/s41558-018-0245-3. URL https://doi.org/10.1038/s41558-018-0245-3. Publisher: Nature Publishing Group.