Multi-Variable Hard Physical Constraints for Climate Model Downscaling

Jose González-Abad[1], Álex Hernández-García[2, 3], Paula Harder[4, 5], David Rolnick[2, 6], José Manuel Gutiérrez[1]

[1]Instituto de Física de Cantabria (IFCA) CSIC-UC
[2]Mila Quebec AI Institute
[3]University of Montreal
[4]Fraunhofer ITWM
[5]University of Kaiserslautern
[6]McGill University
gonzabad@ifca.unican.es

Abstract

Global Climate Models (GCMs) are the primary tool to simulate climate evolution and assess the impacts of climate change. However, they often operate at a coarse spatial resolution that limits their accuracy in reproducing local-scale phenomena. Statistical downscaling methods leveraging deep learning offer a solution to this problem by approximating local-scale climate fields from coarse variables, thus enabling regional GCM projections. Typically, climate fields of different variables of interest are downscaled independently, resulting in violations of fundamental physical properties across interconnected variables. This study investigates the scope of this problem and, through an application on temperature, lays the foundation for a framework introducing multi-variable hard constraints that guarantees physical relationships between groups of downscaled climate variables.

Introduction

Global Climate Models (GCMs) are physics-based models employed to simulate the spatio-temporal evolution of climate and to obtain future climate projections under different climate change scenarios. The resulting projections are crucial to develop adaptation and mitigation plans in many sectors. However, due to computational and physical limitations, the resolution of these models is coarse, which hinders their use in regional-to-local applications.

Statistical Downscaling (SD) techniques attempt to overcome this limitation by learning a relationship between large-scale (low-resolution) data and local-scale (high-resolution) variables of interest. The so-called Perfect Prognosis approach (PP-SD) (Maraun and Widmann 2018) aims at learning a relationship between large-scale predictors and local-scale predictand time-matched pairs from observational data. The set of predictors describes the state of the atmosphere, whereas the predictand corresponds to a surface variable of interest such as temperature or precipitation.

Deep Learning (DL) has recently emerged as a promising SD method, with great potential given its ability to handle spatio-temporal data and model non-linear relationships. DL models developed using historical and future GCM projections can produce actionable local-scale downscaled predictions for climate change studies (Baño-Medina et al. 2022).

Climate is a highly complex system and the variables involved in climate models are closely related to each other by physical links imposed by the constitutive equations, which take into account the interactions and feedback among them. However, most previous works in SD have proposed methods that downscale variables independently, ignoring potential relationships. This can result in inconsistencies and violations of basic physical properties between related groups of variables. Such violations raise important concerns about the reliability and adoption of DL-based downscaling models in climate change applications.

In this work, we first examine the violations of basic physical properties introduced by DL models for the PP-SD of temperature. Our analysis reveals that physical constraints are largely violated when variables are downscaled independently (univariate downscaling), which is especially pronounced in the attempt to generalize from present conditions to future climate scenarios. To address this limitation, we investigate the potential benefits of using a shared model to simultaneously downscale a group of variables. Going further, we establish the groundwork of a novel framework to incorporate multi-variable, physical, hard constraints in neural networks. Through our experimental setup with two architectures and evaluations in GCM projections of future climate, we demonstrate that our framework guarantees the physical constraints for the downscaling of temperature and can be flexibly adapted to other architectures. Our proposed multi-variable model strengthens the reliability of DL-based downscaling and hence facilitates its adoption by the climate science community for practical applications.

Background

DL Methods for Univariate Downscaling

Many works have explored the use of DL models for downscaling individual variables, most using the so-called super-resolution approach (Vandal et al. 2017; Stengel et al. 2020; Wang et al. 2021; Kumar et al. 2021; Passarella et al. 2022; Sharma and Mitra 2022). Inspired by super-resolution methods in computer vision (Wang, Chen, and Hoi 2020), these techniques use a low-resolution version of the target variable as a predictor (in contrast to PP, where the input and output represent different physical quantities). However, this approach is less suitable for downscaling GCMs, as the surface variables used as input are unreliable predictors due to the coarse resolution at which GCMs operate.

Deep learning approaches for univariate PP-SD have been proposed by Pan et al. (2019); Baño-Medina, Manzanas, and Gutiérrez (2020); Sun and Lan (2021); Quesada-Chacón, Barfus, and Bernhofer (2022); Rampal et al. (2022). These models rely on large-scale atmospheric variables as predictors, which are better reproduced by GCMs since these do not depend on local-scale dynamics. These models show promising results in projecting plausible future climate change scenarios over Europe based on different GCMs (Baño-Medina, Manzanas, and Gutiérrez 2021; Baño-Medina et al. 2022). In order to expand the current set of validation tools and better address the concept of assurance (Batarseh, Freeman, and Huang 2021), recent work by González-Abad, Baño-Medina, and Gutiérrez (2023) has explored the development of diagnostics based on eXplainable Artificial Intelligence (XAI) techniques for SD models, uncovering physical inconsistencies in the relationships learned by the models for certain regions.

Approaches for Multivariate Downscaling

Non-DL-based SD models operating on multiple interrelated variables have been extensively explored in the literature. Typically, multivariate downscaling enhances multivariate properties such as cross-correlation (Jeong et al. 2012; Khalili, Van Nguyen, and Gachon 2013; Eum, Gupta, and Dibike 2020), although individual variables may be negatively impacted (Bhowmik et al. 2017). Multivariate DL-based SD approaches, by contrast, have not yet been well developed. Wang and Tian (2022) trained a multivariate DL model for downscaling minimum and maximum temperatures, but found that some predicted minimum temperatures were higher than the corresponding maximum temperatures, thus violating basic physical properties. More recently, Quesada-Chacón et al. (2023) trained a multivariate model to simultaneously downscale several surface variables. While this model exhibited reasonable performance with observational data, individual models achieved superior results, leading to its exclusion from GCM downscaling. No previous work has specifically addressed basic physical inconsistencies.

Constrained DL in Climate Modeling

Several works have attempted to enforce physical constraints in DL models for climate and weather, through both soft and hard constraints. Soft constraints are introduced by adding additional loss terms to the model (Esmaeilzadeh et al. 2020; Beucler et al. 2021; Harder et al. 2022a). Hard constraints are implemented by introducing modifications in the neural network architecture (Geiss and Hardin 2020; Harder et al. 2022a; Hess et al. 2022), ensuring that the constraints are satisfied during learning and inference. In Harder et al. (2022b); Geiss, Silva, and Hardin (2022), hard constraints are applied to DL models for super-resolution SD, enforcing physical relationships between the high-resolution predictand and its low-resolution predictor.

Experimental Framework

Region of Study and Data

Our experimental framework is based on the PP approach. Following previous work (Baño-Medina, Manzanas, and Gutiérrez 2020), in order to represent the atmospheric state, we choose five large-scale variables (geopotential height, zonal and meridional wind, air temperature, and specific humidity) at four different vertical levels (1000, 850, 700 and 500 hPa) as predictors. As predictand, we focus on three variables, namely, minimum, mean, and maximum near-surface air temperature. The predictor variables are obtained from the ERA-Interim reanalysis data set (Dee et al. 2011) at a 2° spatial resolution, while the predictand variables are extracted from the W5E5 observational data set (Lange 2019) at a 0.5° resolution, both at a daily timescale. We choose the European continent as the spatial domain to conduct our experiments.

We select the EC-Earth model run *r12i1p1* (Doblas Reyes et al. 2018) as the GCM to downscale. We focus on the period of 2006-2100 of the Representative Concentration Pathway 8.5 scenario (RCP8.5) (Schwalm, Glendon, and Duffy 2020). This scenario is selected as it represents the strongest climate change signal among those developed in the Coupled Model Intercomparison Project Phase 5 (CMIP5, Taylor, Stouffer, and Meehl (2012)).

Deep Learning Downscaling Models

We train two different DL models: UNet (Ronneberger, Fischer, and Brox 2015) and DeepESD (Baño-Medina et al. 2022). UNet is a fully-convolutional model inspired by an architecture widely used in image segmentation, which has also demonstrated outstanding performance in climate applications, including statistical downscaling (Quesada-Chacón, Barfus, and Bernhofer 2022; Wang and Tian 2022; Doury et al. 2023). DeepESD is a model developed for the PP-SD of temperature and precipitation over Europe. It is composed of a set of convolutional layers and a final fully-connected layer. This model has been validated for the downscaling of various GCMs under future climate scenarios (Baño-Medina, Manzanas, and Gutiérrez 2021; Baño-Medina et al. 2022). We choose these two distinct models as reflecting the current state-of-the-art DL models for downscaling.

The models are trained using the ERA-Interim and W5E5 data sets as predictor and predictand fields, respectively. Predictors are standardized to conform to a standard normal distribution, while predictands are normalized to take on values within the interval $[0, 1]$. We divide the observational data into a training (1980-2000) and a test set (2001-2005). Models are trained to minimize the Mean Squared Error (MSE) using the Adam optimizer with a standard learning rate of 10^{-5} and batch size of 64, using early stopping to prevent overfitting. Prior to passing GCM predictors to DL models for downscaling future scenarios, we perform a signal-preserving adjustment of the monthly mean and variance of the GCM, as suggested by Baño-Medina et al. (2022). This

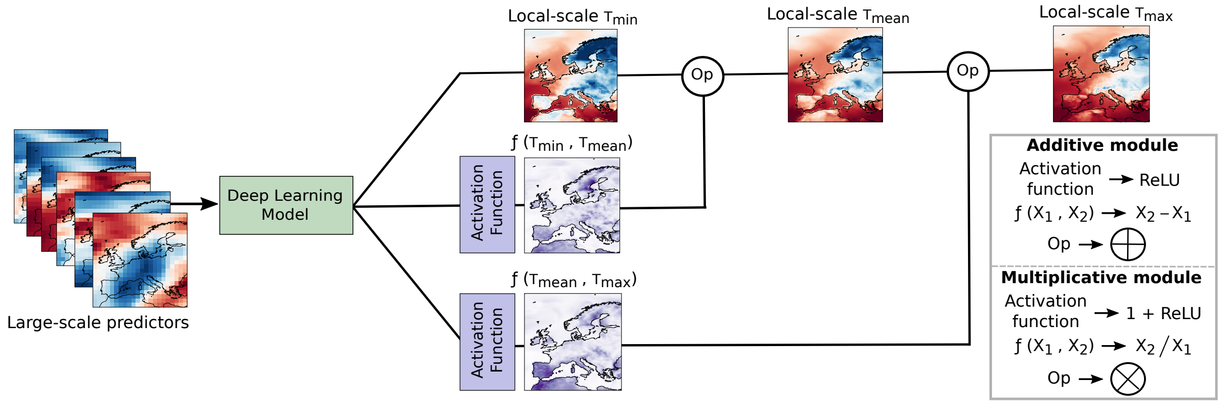

Figure 1: Schematic view of the proposed additive and multiplicative modules. The general architectural layout is depicted in the main area, while the specific operations and activation functions of the modules are indicated on the right side.

improves the extrapolation capabilities of the models by allowing for better similarity in the distributions of the GCM and the reanalysis data.

Multi-Variable Hard Physical Constraints

In this work, we introduce methods for enforcing hard physical constraints between interrelated variables in SD, in the form of simple modules that can be appended onto DL downscaling architectures.

The predictands used in our study are the near-surface minimum, mean, and maximum temperature. For these variables, the following constraints obviously hold:

$$T_{min} \leq T_{mean} \leq T_{max} \tag{1}$$

where T_{min}, T_{mean}, and T_{max} represent the minimum, mean, and maximum temperature. Any violation of these constraints would diminish not merely the average accuracy of an SD method but also its plausibility, significantly reducing the likelihood that the method would be used in practice. While prior work (Wang and Tian 2022) predicted T_{min} and T_{max} jointly, they still suffered from significant constraint violation. We propose two different options for predicting these three variables together while enforcing physical constraints. Figure 1 presents a schematic illustration of our proposed additive and multiplicative modules.

In the *additive* approach, we predict T_{min} and $T_{mean} - T_{min}$ and add these values to obtain T_{mean}, likewise computing T_{max} by predicting $T_{max} - T_{mean}$ and adding it to the previously estimate for T_{mean}. This method ensures that all constraints are met by using *ReLUs* to enforce the nonnegativity of the predicted T_{min}, $T_{mean} - T_{min}$, and $T_{max} - T_{mean}$. In the *multiplicative* approach, a similar method is applied, but instead, we predict T_{min} using *ReLU* activation and the values T_{mean}/T_{min} and T_{max}/T_{mean} using the activation function $1 +$ *ReLU*. Multiplying these predicted values gives the desired quantities while ensuring constraint satisfaction.

In order to assess the efficacy of the proposed modules, we conducted four experiments. In accordance with established methodologies, the first experiment (referred to as *single*) entails the independent modeling of each variable through a distinct DL model for each one. The second experiment (referred to as *shared*) involves training a shared model for the three variables in a multivariate approach. In the case of the shared UNet, the sole discrepancy from the single model is the last convolutional layer, which now features three output channels rather than one. In the DeepESD model, the shared approach entails the utilization of three parallel fully-connected layers instead of one. In the third (*additive*) and fourth (*multiplicative*) experiments, multi-variable hard-constrained models are trained with the two approaches presented above.

Results

Figure 2 shows the Root Mean Squared Error (RMSE) values on the test set obtained for the two DL models trained, namely UNet and DeepESD, for the different experiments (*single*, *shared*, *additive* and *multiplicative*). For T_{min} and T_{max} all combinations of models and experiments display a higher error than for T_{mean}. The underlying cause for this is the extreme values associated with T_{min} and T_{max}, which make modelling them harder (Kjellström et al. 2007).

The UNet model exhibits comparable RMSE values across all experiments. However, for the DeepESD model, a notable improvement in RMSE performance occurs for both the additive and multiplicative hard-constrained models as compared to the single and shared versions. This improvement may be attributed to the presence of fully-connected layers in the architecture of DeepESD, which give rise to the phenomenon known as *dying ReLUs* (Lu et al. 2019), thereby resulting in a degradation of the model's predictive performance. In the hard-constrained modules, the *ReLU* activation outputs are incorporated into T_{min} and T_{mean} through addition or multiplication, thereby mitigating the possible adverse effects of this phenomenon.

Figure 3 presents the annual percentage of violations of the constraints introduced in Section for the downscaled variables obtained from the UNet and DeepESD models. For each of these, the figure displays the results for the *single* and *shared* experiments (solid and dashed lines, respectively), along with the 95% confidence interval computed by applying bootstrapping to 10 independent executions. Violations for the *additive* and *multiplicative* experiments are not shown since, by construction, they are zero.

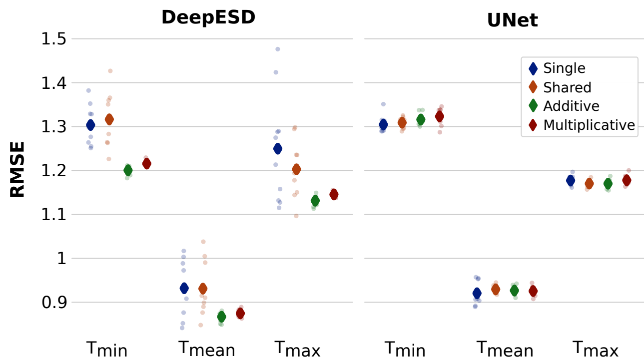

Figure 2: RMSE on the test set for the four experiments (*single*, *shared*, *additive* and *multiplicative*) of the two different models (UNet and DeepESD) trained to downscale minimum (T_{min}), mean (T_{mean}) and maximum (T_{max}) temperature. The values for 10 independent executions and their means are shown.

In the *single* experiments, both architectures show a similar trend in the amount of violations for the training period (1980-2000), maintaining a relatively low and constant amount of violations. However, during the test period (2001-2005), the models must extrapolate to out-of-distribution data, resulting in a sudden spike in the percentage of violations. As the models are applied to the GCM predictors (2006-2100), this issue becomes more pronounced, particularly as we move forward in time, reflecting the difficulties encountered while dealing with the extreme climate conditions associated with the RCP8.5 scenario. The confidence interval of the DeepESD model remains stable over time; however, UNet variability increases, thus being more susceptible to these extreme conditions.

In the *shared* experiment, UNet demonstrates a marked reduction in the percentage of violations, whereas DeepESD continues to exhibit a significant amount. This disparity is attributed to the architectural composition of DeepESD, which allocates a substantial portion of its parameters in the final fully-connected layer. Consequently, the shared model replicates this layer across the downscaled variables, with each variable retaining a considerable amount of parameters dedicated solely to itself, thus not being too different from the single version. Unlike DeepESD, UNet consists entirely of convolutional layers; therefore, in its shared version, each variable's allocation of parameters remains minimal. Despite the evident decline in violations in the *shared* experiments, they do not completely disappear, and more importantly, there are no guarantees regarding the behavior of these models under diverse future scenarios and various GCMs. What is more, as in the *single* experiment, the variability in the number of violations of the UNet model increases with time. Only the multi-variable hard physically constrained models can assure no violations in any scenario.

Conclusions

In this work, we have analyzed the violations of physical constraints commited by DL models for the PP-SD of temperature. Our results reveal that current approaches result in a large number of violations, particularly in the GCM domain. To address this, we have proposed a shared model for the desired variables as a partial solution. However, to ensure multi-variable physical constraints, we have introduced a simple and flexible framework that satisfies hard constraints and achieves the same performance or better than standard approaches.

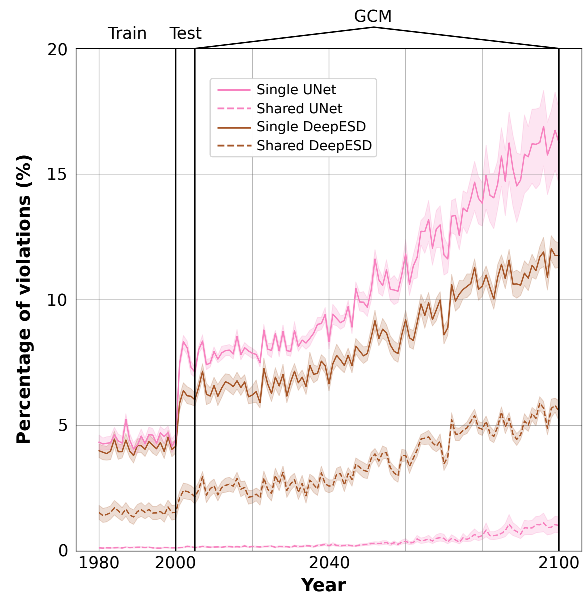

Figure 3: Annual percentage of violations of the physical constraints for the two different DL models intercompared (UNet and DeepESD), for the *single* and *shared* experiments. For each series, the 95% confidence interval is shown.

Preserving fundamental physical properties of GCMs in DL models is crucial for their reliability and hence the adoption by the climate science community. For instance, in the Sixth IPCC Assessment Report, certain indices relying on daily minimum, mean and maximum temperature are employed to asses the impact of climate change in energy demand (Gutiérrez et al. 2021). In this study, we set the groundwork for a novel framework that enables DL models to mimic some underlying physics properties of GCMs, enhancing the reliability of PP-SD models. This work represents a significant effort in integrating DL and GCM models.

As multivariable aspects may be necessary for accurately simulating and downscaling GCMs, in future research, we plan to expand our approach to include new variables such as temperature and precipitation, which may exhibit more complex covariation patterns (Maraun and Widmann 2018). To enhance the reliability of these models, in addition to employing multivariable evaluation metrics, we aim to explore XAI-based diagnostics. These diagnostics allow us to assess the physical consistency of DL models, particularly in challenging scenarios, such as the downscaling of surface variables at higher resolutions, where more complex DL models like transformers (Vaswani et al. 2017) may be required.

Acknowledgments

J. González-Abad would like to acknowledge the support of the funding from the Spanish Agencia Estatal de Investigación through the Unidad de Excelencia María de Maeztu with reference MDM-2017-0765.

References

Baño-Medina, J.; Manzanas, R.; Cimadevilla, E.; Fernández, J.; González-Abad, J.; Cofiño, A. S.; and Gutiérrez, J. M. 2022. Downscaling multi-model climate projection ensembles with deep learning (DeepESD): contribution to CORDEX EUR-44. *Geoscientific Model Development*, 15(17): 6747–6758.

Baño-Medina, J.; Manzanas, R.; and Gutiérrez, J. M. 2020. Configuration and intercomparison of deep learning neural models for statistical downscaling. *Geoscientific Model Development*, 13(4): 2109–2124.

Baño-Medina, J.; Manzanas, R.; and Gutiérrez, J. M. 2021. On the suitability of deep convolutional neural networks for continental-wide downscaling of climate change projections. *Climate Dynamics*, 57: 2941–2951.

Batarseh, F. A.; Freeman, L.; and Huang, C.-H. 2021. A survey on artificial intelligence assurance. *Journal of Big Data*, 8(1): 60.

Beucler, T.; Pritchard, M.; Rasp, S.; Ott, J.; Baldi, P.; and Gentine, P. 2021. Enforcing analytic constraints in neural networks emulating physical systems. *Physical Review Letters*, 126(9): 098302.

Bhowmik, R. D.; Sankarasubramanian, A.; Sinha, T.; Patskoski, J.; Mahinthakumar, G.; and Kunkel, K. E. 2017. Multivariate downscaling approach preserving cross correlations across climate variables for projecting hydrologic fluxes. *Journal of Hydrometeorology*, 18(8): 2187–2205.

Dee, D. P.; Uppala, S. M.; Simmons, A. J.; Berrisford, P.; Poli, P.; Kobayashi, S.; Andrae, U.; Balmaseda, M.; Balsamo, G.; Bauer, d. P.; et al. 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society*, 137(656): 553–597.

Doblas Reyes, F.; Acosta Navarro, J. C.; Acosta Cobos, M. C.; Bellprat, O.; Bilbao, R.; Castrillo Melguizo, M.; Fuckar, N.; Guemas, V.; Lledó Ponsati, L.; Menegoz, M.; et al. 2018. Using EC-Earth for climate prediction research. *ECMWF Newsletter*, (154): 35–40.

Doury, A.; Somot, S.; Gadat, S.; Ribes, A.; and Corre, L. 2023. Regional Climate Model emulator based on deep learning: concept and first evaluation of a novel hybrid downscaling approach. *Climate Dynamics*, 60(5-6): 1751–1779.

Esmaeilzadeh, S.; Azizzadenesheli, K.; Kashinath, K.; Mustafa, M.; Tchelepi, H. A.; Marcus, P.; Prabhat, M.; Anandkumar, A.; et al. 2020. Meshfreeflownet: A physics-constrained deep continuous space-time super-resolution framework. In *SC20: International Conference for High Performance Computing, Networking, Storage and Analysis*, 1–15. IEEE.

Eum, H.-I.; Gupta, A.; and Dibike, Y. 2020. Effects of univariate and multivariate statistical downscaling methods on climatic and hydrologic indicators for Alberta, Canada. *Journal of Hydrology*, 588: 125065.

Geiss, A.; and Hardin, J. C. 2020. Strict Enforcement of Conservation Laws and Invertibility in CNN-Based Super Resolution for Scientific Datasets. *arXiv preprint arXiv:2011.05586*.

Geiss, A.; Silva, S. J.; and Hardin, J. C. 2022. Downscaling atmospheric chemistry simulations with physically consistent deep learning. *Geoscientific Model Development*, 15(17): 6677–6694.

González-Abad, J.; Baño-Medina, J.; and Gutiérrez, J. M. 2023. Using Explainability to Inform Statistical Downscaling Based on Deep Learning Beyond Standard Validation Approaches. *arXiv preprint arXiv:2302.01771*.

Gutiérrez, J. M.; Ranasinghe, R.; Ruane, A. C.; and Vautard, R. 2021. *Annex VI: Climatic Impact-driver and Extreme Indices*, 2205–2214. Cambridge University Press.

Harder, P.; Watson-Parris, D.; Stier, P.; Strassel, D.; Gauger, N. R.; and Keuper, J. 2022a. Physics-informed learning of aerosol microphysics. *Environmental Data Science*, 1: e20.

Harder, P.; Yang, Q.; Ramesh, V.; Sattigeri, P.; Hernandez-Garcia, A.; Watson, C.; Szwarcman, D.; and Rolnick, D. 2022b. Generating physically-consistent high-resolution climate data with hard-constrained neural networks. *arXiv preprint arXiv:2208.05424*.

Hess, P.; Drüke, M.; Petri, S.; Strnad, F. M.; and Boers, N. 2022. Physically constrained generative adversarial networks for improving precipitation fields from Earth system models. *Nature Machine Intelligence*, 4(10): 828–839.

Jeong, D. I.; St-Hilaire, A.; Ouarda, T. B.; and Gachon, P. 2012. A multivariate multi-site statistical downscaling model for daily maximum and minimum temperatures. *Climate research*, 54(2): 129–148.

Khalili, M.; Van Nguyen, V. T.; and Gachon, P. 2013. A statistical approach to multi-site multivariate downscaling of daily extreme temperature series. *International Journal of Climatology*, 33(1): 15–32.

Kjellström, E.; Bärring, L.; Jacob, D.; Jones, R.; Lenderink, G.; and Schär, C. 2007. Modelling daily temperature extremes: recent climate and future changes over Europe. *Climatic Change*, 81(Suppl 1): 249–265.

Kumar, B.; Chattopadhyay, R.; Singh, M.; Chaudhari, N.; Kodari, K.; and Barve, A. 2021. Deep learning–based downscaling of summer monsoon rainfall data over Indian region. *Theoretical and Applied Climatology*, 143: 1145–1156.

Lange, S. 2019. WFDE5 over land merged with ERA5 over the ocean (W5E5). *GFZ Data Services*.

Lu, L.; Shin, Y.; Su, Y.; and Karniadakis, G. E. 2019. Dying relu and initialization: Theory and numerical examples. *arXiv preprint arXiv:1903.06733*.

Maraun, D.; and Widmann, M. 2018. *Statistical downscaling and bias correction for climate research*. Cambridge University Press.

Pan, B.; Hsu, K.; AghaKouchak, A.; and Sorooshian, S. 2019. Improving precipitation estimation using convolutional neural network. *Water Resources Research*, 55(3): 2301–2321.

Passarella, L. S.; Mahajan, S.; Pal, A.; and Norman, M. R. 2022. Reconstructing high resolution ESM data through a novel fast super resolution convolutional neural network (FSRCNN). *Geophysical Research Letters*, 49(4): e2021GL097571.

Quesada-Chacón, D.; Baño-Medina, J.; Barfus, K.; and Bernhofer, C. 2023. Downscaling CORDEX through deep learning to daily 1 km multivariate ensemble in complex terrain. *Earth's Future*, 11(8): e2023EF003531.

Quesada-Chacón, D.; Barfus, K.; and Bernhofer, C. 2022. Repeatable high-resolution statistical downscaling through deep learning. *Geoscientific Model Development*, 15(19): 7353–7370.

Rampal, N.; Gibson, P. B.; Sood, A.; Stuart, S.; Fauchereau, N. C.; Brandolino, C.; Noll, B.; and Meyers, T. 2022. High-resolution downscaling with interpretable deep learning: Rainfall extremes over New Zealand. *Weather and Climate Extremes*, 38: 100525.

Ronneberger, O.; Fischer, P.; and Brox, T. 2015. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, 234–241. Springer.

Schwalm, C. R.; Glendon, S.; and Duffy, P. B. 2020. RCP8. 5 tracks cumulative CO2 emissions. *Proceedings of the National Academy of Sciences*, 117(33): 19656–19657.

Sharma, S. C. M.; and Mitra, A. 2022. ResDeepD: A residual super-resolution network for deep downscaling of daily precipitation over India. *Environmental Data Science*, 1: e19.

Stengel, K.; Glaws, A.; Hettinger, D.; and King, R. N. 2020. Adversarial super-resolution of climatological wind and solar data. *Proceedings of the National Academy of Sciences*, 117(29): 16805–16815.

Sun, L.; and Lan, Y. 2021. Statistical downscaling of daily temperature and precipitation over China using deep learning neural models: Localization and comparison with other methods. *International Journal of Climatology*, 41(2): 1128–1147.

Taylor, K. E.; Stouffer, R. J.; and Meehl, G. A. 2012. An overview of CMIP5 and the experiment design. *Bulletin of the American meteorological Society*, 93(4): 485–498.

Vandal, T.; Kodra, E.; Ganguly, S.; Michaelis, A.; Nemani, R.; and Ganguly, A. R. 2017. Deepsd: Generating high resolution climate change projections through single image super-resolution. In *Proceedings of the 23rd acm sigkdd international conference on knowledge discovery and data mining*, 1663–1672.

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; and Polosukhin, I. 2017. Attention is all you need. *Advances in neural information processing systems*, 30.

Wang, F.; and Tian, D. 2022. On deep learning-based bias correction and downscaling of multiple climate models simulations. *Climate dynamics*, 59(11-12): 3451–3468.

Wang, J.; Liu, Z.; Foster, I.; Chang, W.; Kettimuthu, R.; and Kotamarthi, V. R. 2021. Fast and accurate learned multiresolution dynamical downscaling for precipitation. *Geoscientific Model Development*, 14(10): 6355–6372.

Wang, Z.; Chen, J.; and Hoi, S. C. 2020. Deep learning for image super-resolution: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 43(10): 3365–3387.