Physics-Constrained Deep Learning for Downscaling

Paula Harder [1 2] **Venkatesh Ramesh** [1 3] **Alex Hernandez-Garcia** [1 3] **Qidong Yang** [1 4] **Prasanna Sattegeri** [5]
Daniela Szwarcman [5] **Campbell Watson** [5] **David Rolnick** [1 6]

Abstract

The availability of reliable, high-resolution climate and weather data is important to inform long-term decisions on climate adaptation and mitigation and to guide rapid responses to extreme events. Forecasting models are limited by computational costs and, therefore, often generate coarse-resolution predictions. Statistical downscaling, including super-resolution methods from deep learning, can provide an efficient method of upsampling low-resolution data. However, despite achieving visually compelling results in some cases, such models frequently violate conservation laws when predicting physical variables. In order to conserve physical quantities, we develop a method that guarantees physical constraints are satisfied by a deep learning downscaling model while also improving their performance according to traditional metrics.

1. Introduction

Accurate modeling of weather and climate is critical f taking effective action to combat climate change. In a dition to shaping global understanding of climate change, local and regional predictions guide adaptation decisions and provide impetus for action to reduce greenhouse gas emissions (Gutowski & et al., 2020). Predicted and observed quantities such as precipitation, wind speed, and temperature impact decisions in sectors such as agriculture, energy, and transportation. While these quantities are often required at a fine geographical and temporal scale to ensure informed decision-making, most climate and weather models are extremely computationally expensive to run (sometimes taking months even on super-computers), resulting in coarse-resolution predictions. Thus, there is a need for fast methods that can generate high-resolution data based on the low-resolution models that are commonly available.

The terms *downscaling* in climate science and *super-resolution* (SR) in machine learning (ML) refer to a map from low-resolution (LR) input data to high-resolution (HR) versions of that same date; the high-resolution output is referred to as the super-resolved (SR) data. Generating high-resolution data with machine learning can produce realistic-looking images and good predictive accuracy. However, a major obstacle often encountered when applying ML to a physical system such as the Earth's atmosphere is that the predicted output values can violate physical laws such as conservation of energy, momentum, and mass. In this work, we introduce novel methods to strictly enforce physical constraints between low-resolution (input) and high-resolution (output) images.

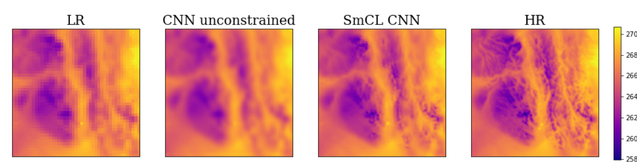

Figure 1. A random prediction for the WRF temperature dataset. We compare unconstrained and softmax-constrained predictions.

2. Methodology and Data

Consider the case of downscaling by a factor of N in each linear dimension, and let $n := N^2$. Let $y_i, i = 1, \dots, n$ be the values in the predicted high-resolution patch that correspond to low-resolution pixel x. The set $\{y_i\}$ for $i = 1, \dots, n$ is also referred to as a super-pixel. Then, a conservation law takes the form of the following constraint:

$$\frac{1}{n}\sum_{i=1}^{n} y_i = x. \tag{1}$$

We will use $\tilde{y}_i, i = 1, \dots, n$ to denote the intermediate outputs of the neural network before the constraint layer and $y_i, i = 1, \dots, n$ to be the final outputs after applying the

[1]Mila Quebec AI Institute, Montreal, Canada [2]Competence Center High Performance Computing, Fraunhofer Institute for Industrial Mathematics, Kaiserslautern, Germany [3] University of Montreal, Montreal, Canada [4]New York University, New York, USA [5]IBM Research, New York, USA [6]McGill University, Montreal, Canada. Correspondence to: paula.harder@itwm.fraunhofer.de <
>.

Table 1. Metrics for different constraining methods applied to the SR CNN applied on the WRF temperature data, calculated over 10,000 test samples. The mean is taken over 3 runs. The best scores are highlighted in bold.

DATA	MODEL	CONSTRAINT	RMSE	MAE	MS-SSIM	CONSTR. VIOL.
ERA5 WC	BICUBIC	NONE	0.322	0.137	99.90	0.066
ERA5 WC	CNN	NONE	0.251	0.105	99.95	0.026
ERA5 WC	CNN	SOFT	0.301	0.137	99.23	0.016
ERA5 WC	CNN	SMCL	**0.215**	**0.094**	**99.96**	**0.000**
T2 WRF	BICUBIC	NONE	0.322	0.137	99.90	0.066
T2 WRF	CNN	NONE	0.952	0.618	94.92	0.181
T2 WRF	CNN	SOFT	1.020	0.660	94.57	0.032
T2 WRF	CNN	SMCL	**0.950**	**0.592**	**95.25**	**0.000**

constraints. We use a softmax multiplied by the corresponding input pixel value x:

$$y_j = \exp(\tilde{y}_j) \cdot \frac{x}{\frac{1}{n}\sum_{i=1}^{n} \exp(\tilde{y}_i)}. \quad (2)$$

This Softmax Constraint Layer (SmCL) satifies (1) by construction and additionally enforces $y_i \geq 0, i = 1, \ldots, n$.

To evaluate our proposed method we consider two datasets: one uses ERA5 reanalysis (Hersbach, 2020) water content data and the other data from the Weather and Research Model (WRF) forecast (Auger et al., 2021) of temperature. While the water content data creates the low-resolution counterparts synthetically, the WRF dataset uses two independent simulations for low-resolution and high-resolution pairs.

3. Experiments

We add our SmCL to a standard CNN for downscaling and compare it to bicubic interpolation, the same CNN without the constraint layer and soft constraining. Soft constraining does not change the architecture but adds a penalizing term to the loss function:

$$\text{Loss} = (1 - \alpha) \cdot \text{MSE} + \alpha \cdot \text{Constraint violation}, \quad (3)$$

with a tunable parameter α, here chosen as 0.99.

4. Results and Conclusion

This work presents a novel methodology to incorporate physics-based constraints into neural network architectures for climate downscaling. We demonstrate its skill both on a standard downscaling dataset and on data created by independent simulations. Our constrained models are not only guaranteed to satisfy conservation laws such as mass conservation, but also increase predictive performance across metrics and use cases, as shown in Table 1. An example of the WRF dataset is visualized in Figure 1, where it can be seen that constraining increases the visual quality of the prediction significantly. Future work could extend the application of our constraint layer to other tasks than downscaling. Climate model emulation for example could strongly benefit from a reliable and performance-enhancing method to enforce physical laws.

References

Auger, G. A. R., Watson, C. D., and Kolar, H. R. The influence of weather forecast resolution on the circulation of lake george, ny. *Water Resources Research*, 57(10): e2020WR029552, 2021. doi: https://doi.org/10.1029/2020WR029552. e2020WR029552 2020WR029552.

Gutowski, W. J. and et al. The ongoing need for high-resolution regional climate models: Process understanding and stakeholder information. *Bulletin of the American Meteorological Society*, 101(5):E664 – E683, 2020. doi: 10.1175/BAMS-D-19-0113.1.

Hersbach, H. e. a. The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730): 1999–2049, 2020. doi: https://doi.org/10.1002/qj.3803.