Generating physically-consistent high-resolution climate data with hard-constrained neural networks

Paula Harder[1,2,3], Qidong Yang[1,4], Venkatesh Ramesh[1,5], Alex Hernandez-Garcia[1,5], Prasanna Sattigeri[6], Campbell D. Watson[6], Daniela Szwarcman[6], and David Rolnick[1,7]

[1]Mila Quebec AI Institute
[2]Fraunhofer ITWM
[3]New York University
[4]New York University
[5]University of Montreal
[6]IBM Research
[7]McGill University

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 can only make coarse resolution predictions. Statistical downscaling can provide an efficient method of upsampling low-resolution data. In this field, deep learning has been applied successfully, often using image super-resolution methods from computer vision. Despite achieving visually compelling results in some cases, such models often violate conservation laws when predicting physical variables. In order to conserve important physical quantities, we developed a deep downscaling method that guarantees physical constraints are satisfied, by adding a renormalization layer at the end of the neural network. Furthermore, the constrained model also improves the performance according to standard metrics. We show the applicability of our methods across different popular architectures and upsampling factors using ERA5 reanalysis data.

1 Introduction

Accurate modeling of weather and climate is critical for taking effective action to combat climate change. In addition to shaping global understanding of climate change, local and regional predictions shape adaptation decisions and provide impetus for action to reduce greenhouse gas emissions. Predicted and observed quantities such as precipitation, wind speed, and temperature impact decisions in fields like 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, resulting in coarse-resolution predictions, while observations are often sparsely available. 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 learning a mapping between a low-resolution (LR) input sample and a high-resolution (HR) version of that sample. Downscaling via established statistical methods—*statistical downscaling*— has been long used by the climate science community to increase the resolution of climate data Maraun and Widmann [2018]. In parallel, computer vision SR has evolved rapidly using various deep learning architectures, with such methods now including super-resolution convolutional neural

networks (CNNs) Dong et al. [2016], generative adversarial models (GANs) Wang et al. [2018], vision transformers Yang et al. [2020], and normalizing flows Lugmayr et al. [2020]. Increasing the temporal resolution via frame interpolation is also an active area of research for video enhancement Liu et al. [2017] that can be transferred to spatio-temporal climate data. Recently, deep learning approaches have been applied to a variety of climate and weather datasets, covering both model output data and observations. Climate super-resolution has mostly focused on CNNs Vandal et al. [2017], recently shifting towards GANs Stengel et al. [2020].

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 moisture. More broadly, there are numerous domains of ML for societal benefit in which satisfaction of physical constraints is fundamentally important. Examples include the discovery of new materials for energy and healthcare, aerodynamics simulations for efficient vehicles, and optimal control in industrial settings.

To the best of our knowledge, there is no published work on hard-constrained ML models for SR/downscaling, but there exist related work. In super-resolution of turbulent flows, Mesh-freeFlowNet Jiang et al. [2020] employs a physics-informed model adding partial differential equations as regularization terms to the loss function. A preprint Geiss and Hardin [2020] introduces an enforcement operator applied to multiple CNN architectures for scientific datasets while another more recent preprint applies that method to atmospheric chemistry model downscaling using a CNN Geiss et al. [2022].

In this work, we introduce a novel method to strictly enforce physical constraints between low-resolution (input) and high-resolution (output) images. Using water content data the ERA5 reanalysis product that describes the recent climate combining models and observations. We then build datasets to learn the upsampling with different factors ranging from 2 up to 8 times enhancement. We introduce a renormalization layer that is added as the last layer to any network architecture. Besides looking at spatial SR, we also include temporal SR or frame interpolation, demonstrating that our methodology works across all these architectures.

Our novel constraining methodology is not only applicable to SR tasks in scientific areas, but it can also constrain neural networks for emulation and prediction. Climate model emulation tasks as done by Beucler et al. [2021] and Harder et al. [2021] could benefit from using our constraining layer. Within the SR domain, our work could have implications beyond scientific datasets as the constraining methodologies can advance state-of-the-art SR more generally.

2 Enforcing Constraints

When modeling physical quantities such as precipitation or water mass, principled relationships such as mass conservation can naturally be established between low-resolution and high-resolution samples. Here, we introduce a new methodology to incorporate these constraints within a neural network architecture.

2.1 Setup

Consider the case of N times downscaling and let $n := N^2$. Let $y_i, i = 1, \ldots, n$ be the values in the predicted high-resolution data point that correspond to low-resolution pixel x. The downscaling/conservation constraint is given by the following:

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

As we model mass per area, conserving the mean in the SR image means conserving the overall mass.

2.2 Constraints Layer

Let $\tilde{y}_i, i = 1, \ldots, n$ be the intermediate outputs of the neural network before the constraints layer. At the end of a network architecture we take these outputs and rescale them using the corresponding

input value x:

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

We call this operation the constraints layer (CL), which guarantees Equation 2 is satisfied and can be added to the end of a neural network.

For predicting quantities like atmospheric water content, we also want to enforce the output to be positive and, therefore, physically valid. Here, we use a softmax multiplied by the corresponding input pixel value x:

$$y_j = \exp\left(\tilde{y}_j\right) \cdot \frac{x}{\frac{1}{n} \sum_{i=1}^{n} \exp\left(\tilde{y}_i\right)}. \tag{3}$$

We call this layer the softmax constraints layer (SMCL). It not only enforces our equality constraints, but also our positivity constraints: $y_i \geq 0, i = 1, \ldots, n$. Including the exponential function also changes the distribution, making it more extreme/peakier, which can help with the common problem of having too smooth predictions while optimizing the MSE.

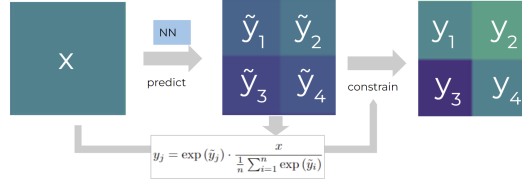The constraints are applied for each pair of input pixel x and the corresponding SR $N \times N$ patch.

Figure 1: Our constraining methodology shown for one input pixel x and the softmax constraining on the corresponding predicted 2×2 patch for the case of 2 times upsampling.

3 Experimental Setup

We use ERA5 water content data. To obtain our high-resolution data points we extract a random 128×128 pixel image from each available time step (each time step is 721×1440 and there are roughly 60,000 time steps available). The low-resolution counterparts are generated by average downsampling. Our CNN network consists of convolutional layers with 3×3 kernels and ReLU activations. The upsampling is done by a transpose convolution followed by residual blocks. We also tested on different architectures like GANs and RNNs, details can be found in the supplementary material.

Our models were trained with the Adam optimizer, learning rate of 0.001, batch size of 256 and MSE loss between predicted SR image and true HR image. We trained for 200 epochs, which took about 3—6 hours on a single NVIDIA A100 Tensor Core GPU, depending on the architecture. All models use the MSE as their criterion, the GAN additionally uses its discriminator loss term.

We also run experiments with the enforcement operator, introduced by Geiss and Hardin [2020], which we refer to as *G-H*. Their constraining operator, tested in the context of SR CNNs only, is as follows:

$$y_j = \tilde{y}_j + (x - \frac{1}{n} \sum_{i=1}^{n} \tilde{y}_i) \cdot \frac{\sigma + \tilde{y}_i}{\sigma + \frac{1}{n} \sum_{i=1}^{n} \tilde{y}_i}, \tag{4}$$

with $\sigma := \text{sign}(\frac{1}{n} \sum_{i=1}^{n} \tilde{y}_i - x)$.

4 Results

As standard SR metrics, we report the peak signal-to-noise ratio (PSNR) and the structural similarity index measure (SSIM) for all our experiments. Additionally, we report how much mass conservation is violated.

Table 1: Metrics for different constraining methods applied to an SR CNN, calculated over 10,000 test samples. Best scores are highlighted in bold, second best in blue. The metrics shown are PSNR/SSIM/Mass violation.

MODEL	CONSTRAINT	2x	4x	8x
BICUBIC	-	$52.2/99.88/6.6e^{-2}$	$44.3/99.29/1.7e^{-1}$	$38.5/98.08/3.2e^{-1}$
SR-CNN	NONE	$56.0/99.94/1.5e^{-2}$	$47.5/99.58/3.1e^{-2}$	$40.7/\mathbf{98.56}/6.0e^{-2}$
SR-CNN	G-H	$\mathbf{57.0}/\mathbf{99.95}/1.6e^{-6}$	$\mathbf{47.9}/\mathbf{99.60}/1.1e^{-5}$	$\mathbf{40.8}/98.55/4.4e^{-6}$
SR-CNN	CL (OURS)	$56.3/99.94/\mathbf{8.6e^{-7}}$	$\mathbf{47.9}/\mathbf{99.60}/\mathbf{1.3e^{-6}}$	$40.5/98.48/\mathbf{1.2e^{-6}}$
SR-CNN	SMCL (OURS)	$\mathbf{57.0}/\mathbf{99.95}/8.7e^{-7}$	$\mathbf{47.9}/\mathbf{99.60}/1.4e^{-6}$	$40.7/98.54/2.2e^{-6}$

We can see in Table 1 that for almost all cases applying constraining methodologies does not only enforce mass conservation, but also improves both PSNR and SSIM scores. Due to numerical reasons, the mass conservation scores vary across constraining methodologies and are not exactly zero. Overall the predicted values have a mean value of around $20\frac{kg}{m^2}$, therefore the violations are relatively small among all constraining methodologies. For predictive accuracy the G-H operator and our softmax constraining layer perform slightly better than the simple constraining layer. In the supplementary material results for more architecture are shown, here it shows that for the most complex model, a simultaneous time and spatial-SR network, our SMCL methods achieves significantly better scores than the G-H operator (see Table 4).
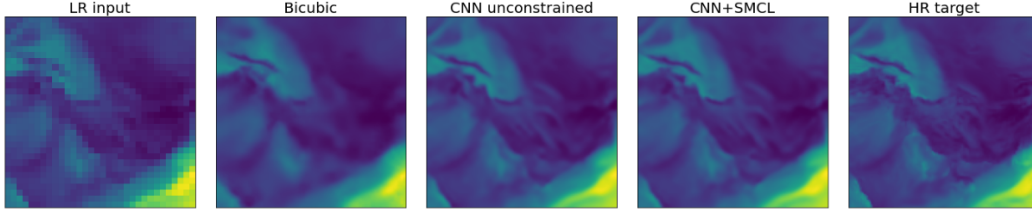

Figure 2: One example image chosen randomly from the test set. Shown here is the LR input, the SMCL-constrained and unconstrained predictions and the HR image as a reference.

5 Conclusion

This work shows a novel methodology to incorporate physics-based constraints into neural network architectures for climate downscaling. Our constrained models not only guarantee to satisfy conservation laws like mass conservation or their outputs but also increase predictive performance across different architectures.

Next steps include extending the evaluation to different SR climate datasets as well as applying the constraining layers to other climate or scientific ML tasks. One possible future direction could be transferring the constraining methodologies to super-resolution task in general and advancing their performance using our methods.

References

T. Beucler, M. Pritchard, S. Rasp, J. Ott, P. Baldi, and P. Gentine. Enforcing analytic constraints in neural networks emulating physical systems. *Phys. Rev. Lett.*, 126:098302, Mar 2021. doi: 10.1103/PhysRevLett.126.098302. URL https://link.aps.org/doi/10.1103/PhysRevLett.126.098302.

C. Dong, C. C. Loy, K. He, and X. Tang. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2):295–307, 2016. doi: 10.1109/TPAMI.2015.2439281.

A. Geiss and J. C. Hardin. Strict enforcement of conservation laws and invertibility in cnn-based super resolution for scientific datasets, 2020. URL https://arxiv.org/abs/2011.05586.

A. Geiss, S. Silva, and J. Hardin. Downscaling atmospheric chemistry simulations with physically consistent deep learning. *Geoscientific Model Development Discussions*, 2022:1–26, 2022. doi: 10.5194/gmd-2022-76. URL `https://gmd.copernicus.org/preprints/gmd-2022-76/`.

P. Harder, D. Watson-Parris, D. Strassel, N. Gauger, P. Stier, and J. Keuper. Physics-informed learning of aerosol microphysics. *arXiv preprint arXiv:2109.10593*, 2021.

C. M. Jiang, S. Esmaeilzadeh, K. Azizzadenesheli, K. Kashinath, M. Mustafa, H. A. Tchelepi, P. Marcus, Prabhat, and A. Anandkumar. Meshfreeflownet: A physics-constrained deep continuous space-time super-resolution framework. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, SC '20. IEEE Press, 2020. ISBN 9781728199986.

C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690, 2017.

J. Leinonen, D. Nerini, and A. Berne. Stochastic super-resolution for downscaling time-evolving atmospheric fields with a generative adversarial network. *IEEE Transactions on Geoscience and Remote Sensing*, 59(9):7211–7223, 2021. doi: 10.1109/TGRS.2020.3032790.

Z. Liu, R. A. Yeh, X. Tang, Y. Liu, and A. Agarwala. Video frame synthesis using deep voxel flow. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 4473–4481, 2017. doi: 10.1109/ICCV.2017.478.

A. Lugmayr, M. Danelljan, L. Van Gool, and R. Timofte. Srflow: Learning the super-resolution space with normalizing flow. In *ECCV*, 2020.

D. Maraun and M. Widmann. *Statistical Downscaling and Bias Correction for Climate Research*. Cambridge University Press, 2018. doi: 10.1017/9781107588783.

K. Stengel, A. Glaws, D. Hettinger, and R. N. King. Adversarial super-resolution of climatological wind and solar data. *Proceedings of the National Academy of Sciences*, 117(29):16805–16815, 2020. doi: 10.1073/pnas.1918964117. URL `https://www.pnas.org/doi/abs/10.1073/pnas.1918964117`.

T. Vandal, E. Kodra, S. Ganguly, A. Michaelis, R. Nemani, and A. R. Ganguly. Deepsd: Generating high resolution climate change projections through single image super-resolution. KDD '17, page 1663–1672, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450348874. doi: 10.1145/3097983.3098004. URL `https://doi.org/10.1145/3097983.3098004`.

X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, C. C. Loy, Y. Qiao, and X. Tang. Esrgan: Enhanced super-resolution generative adversarial networks, 2018. URL `https://arxiv.org/abs/1809.00219`.

F. Yang, H. Yang, J. Fu, H. Lu, and B. Guo. Learning texture transformer network for image super-resolution. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5790–5799, 2020. doi: 10.1109/CVPR42600.2020.00583.

Supplemetary Material

SR-GAN   Our conditional GAN architecture uses the above introduced CNN architecture as the generator network. The discriminator simply consists of convolutional layers with a stride of 2 to decrease the dimensionality in each step and ReLU activations. It is trained as a classifier to distinguish SR images from real HR images using a binary cross entropy loss. The conditional GAN takes as input both Gaussian noise as well as the LR data and then generates an SR output. It is trained with a combination of an MSE loss and the adversarial loss given by the discriminator, like a standard SR GAN, e.g. Ledig et al. [2017].

SR-ConvGRU We apply the architecture presented by Leinonen et al. [2021], which uses ConvGRU layers to address the spatio-temporal nature of super-resolving a time-series of climate data.

SR-VoxelConvGRU To increase the temporal resolution of our data we employ the Deep Voxel Flow method Liu et al. [2017], a deep learning architecture for video frame interpolation combining optical flow methods with neural networks. We introduce a new architecture by stacking the Deep Voxel Flow model and the ConvGRU network (VoxelConvGRU): First, we increase the temporal resolution resulting in a higher-frequency time-series of LR images on which we then apply the ConvGRU architecture to increase the spatial resolution. The combined neural networks are then trained end-to-end.

Table 2: Scores for different constraining methods applied to a SR GAN architecture, calculated over 10,000 test samples. Best scores are highlighted in bold, second best in blue. The metrics shown are PSNR/SSIM/Mass violation.

MODEL	CONSTRAINT	METRICS
BICUBIC	-	$44.3/99.29/1.7e^{-1}$
	NONE	$46.5/99.50/4.6e^{-2}$
	G-H	$\mathbf{47.4}/\mathbf{99.56}/2.3e^{-6}$
SR-GAN	CL (OURS)	$46.9/99.53/1.4e^{-6}$
	SMCL (OURS)	$47.3/\mathbf{99.56}/\mathbf{1.3e^{-6}}$

Table 3: Scores for different constraining methods applied to the SR ConvGRU architecture, calculated over 10,000 test samples for 4 times upsampling. Best scores are highlighted in bold, second best in blue. The metrics shown are PSNR/SSIM/Mass violation.

MODEL	CONSTRAINT	METRICS
BICUBIC	-	$44.2/99.28/1.7e^{-1}$
	NONE	$44.8/99.37/1.8e^{-1}$
	G-H	$\mathbf{48.3}/\mathbf{99.64}/5.5e^{-5}$
SR-CONVGRU	CL (OURS)	$44.7/99.30/\mathbf{1.3e^{-6}}$
	SMCL (OURS)	$48.2/99.63/\mathbf{1.3e^{-6}}$

Table 4: Scores for different constraining methods applied to our VoxelConvGRU SR architecture, calculated over 10,000 test samples for 4 times upsampling spatially and 2 times temporally. Best scores are highlighted in bold, second best in blue.

MODEL	CONSTRAINT	METRICS
BICUBIC+INTERP.	-	$42.9/99.18/3.1e^{-1}$
	NONE	$45.1/99.40/1.5e^{-1}$
	G-H	$44.7/99.38/1.7e^{-6}$
SR-VOXELCONVGRU	CL (OURS)	$44.0/99.16/\mathbf{8.7e^{-7}}$
	SMCL (OURS)	$\mathbf{47.9}/\mathbf{99.62}/\mathbf{8.7e^{-7}}$