Numerical Weather Model Super-Resolution

Alok Singh Terrafuse, Inc Berkeley, CA 94709 alok@berkeley.edu Brian White UNC Chapel Hill Chapel Hill, NC 27514 bwhite@unc.edu Adrian Albert Terrafuse, Inc Berkeley, CA 94709 toni@terrafuse.ai

Abstract

Numerical simulation of weather is constrained due to the high computational cost of integrating the coupled PDEs that govern atmospheric motion. Even the finest-scale numerical weather prediction models cannot model the scales that dictate weather in urban areas and regions with high topographic complexity, like mountains. Thus, several statistical methods have been developed in the climate community to upsample numerical model output to finer resolutions. This is conceptually similar to image super-resolution (SR) [1] and in this work we report the results of applying SR methods to this problem. We compare several methods and find ESRGAN [2] to give high-fidelity qualitative recovery but poorer performance on metrics such as MSE. However, the high frequency power spectrum is captured remarkably well by ESRGAN, virtually identical to the real data, while other method's fidelity drops significantly at high frequency. We use this observation to modify our approach to optimize the power spectrum directly in our loss function, and call this technique PSD-Net. We achieve better performance across all metrics, along with increased stability and faster training time.

1 Motivation

Global climate models are limited to ~100 km resolution, while numerical weather prediction models that produce daily forecasts and severe weather warnings are limited to ~3 km. However, accurate assessment of climate and extreme weather impacts near human populations would benefit substantially from finer resolution. While several methods have been developed for upsampling climate models output to finer resolutions, they consist for the most part of complex interpolation methods (see e.g. [3]). In this paper we explore deep learning methods to upsample weather model output.

Deep learning approaches have only recently started to receive attention in the earth sciences community [4]. Over traditional numeric-based approaches, they could address some key issues in climate modeling:

- 1. A pipeline trained end-to-end that automatically learns optimal filters and transformations between inputs (i.e., remote-sensing, in-situ, and simulation data) and their relationship to the parameters of interest (e.g., wind intensity, precipitation), can drastically accelerate the creation of practical ad-hoc relationships between observational datasets and models.
- 2. An end-to-end differentiable model will allow for the exploration of climate model sensitivities that lead to bias including the influence of the meteorological forcing dataset. Slight perturbations in precipitation phase, intensity, and/or location, short/longwave radiation, wind speed and direction, humidity, and, temperature, could be used to understand the downstream implications on variables that are difficult to measure or model directly.
- 3. Generative models can be run in parallel, and do not necessarily require iterative schemes to model data, allowing them to run quickly even on low-grade consumer hardware.

33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.

2 Methods and Data



Figure 1: Example wind speeds from the WRF model output, shown every 4 hours over 1 day.



Figure 2: Output of each model and the input that produced it. Each row is one image, and we can see some finer detail in the zoomed view, highlighted in the red box.

2.1 Dataset

We use 15 years of wind velocity fields from a numerical simulation of the WRF (Weather Research and Forecasting) model over Southern California (see [5] for details — we use region d04 from 2001-2015), gathered hourly (Figure 1), for a total of about 60,000 data points, represented as grids. In the rest of this work, we shall call these grids "images".

Each image is of size 153×153 , where each pixel represents the average wind speed over a 1.5 km \times 1.5 km region. They are stored as a 2-D array of 32-bit floats, linearly scaled to [0, 1]. For ease of upsampling, we clip a single row and column to resize our images to 152×152 . To generate input data, we downsample our images by a factor of 4 in each dimension.

The dataset is shuffled and split into training and validation sets, with 5% (3,000 images) held out for validation.

Figure 3: Architectures and PSD plot



(b) Note that ESRGAN tracks the true data far more closely than the other models.

2.2 Method

2.2.1 Bicubic Upsampling

Bicubic interpolation is useful as a baseline to compare against. There are no trainable parameters, so we simply upsample our validation set.

2.2.2 SR-CNN

SR-CNN [6] is a classic deep learning model used for image SR. As preprocessing, a low-resolution image is first upsampled to the desired size by another method like bicubic interpolation. The upsampled image is then passed to a CNN, essentially sharpening it. (See Figure 3a for a high-level overview.)

T_{α} L1	~ 1	. г)	1+0
тарі	ег	: r	cesu	IIIS

Model	PSNR	MSE	MAE	KL
ESRGAN	32.74	5.3e-4	0.0148	0.008
SR-CNN	36.06	2.4e-4	0.0091	0.015
Bicubic	35.52	2.7e-4	0.0097	0.006
PSD-Net	39.3	1.1e-4	0.0066	0.005

We train for 200 epochs with a batch size of 128, the Adam optimizer, and a learning rate of 0.001. For the upsampling preprocessing step, we use bicubic upsampling.

2.2.3 ESRGAN

ESRGAN is an conditional GAN designed for image SR. We pass the generator G a batch of lowresolution images, which are upsampled by G and then passed to the discriminator D, as is standard for GANs. There is also a "pixel loss", which compares the MSE between the SR and ground truth images.

We follow the training procedure described in [2]. We use a batch size of 16 images and train for 200 epochs.

3 Evaluation

Table 1 gives an overview of final performance on the validation set. PSNR (peak signal to noise ration), MSE and MAE (mean absolute error) are averaged over all images in the validation set. "KL" represents the KL divergence between the empirical distributions of the generated images and the ground truth.

Notice that ESRGAN performs worst on these metrics. But the generated images tell quite a different story. As we see in Figure 2, ESRGAN generates clearer images than most other methods. Zooming in on the highlighted red box reveals that the images generated by ESRGAN are sharper.

PSNR and MSE have been noted as poor indicators of image quality, as they fail to capture the underlying dynamics of images well. A key metric that illustrates the spatial resolution and higher moments of the data distribution is the power spectral density, shown in Figure 3b. The PSD plot reveals the power of ESRGAN to capture the high frequency information present in the wind field. In fact, ESRGAN's spectrum is so close to that of the true data that they are nearly indistinguishable, whereas SR-CNN and bicubic upsampling fall off significantly at higher frequencies. This is perhaps not surprising as the upsampling and SR-CNN are fundamentally methods of interpolation, whereas ESRGAN is learning the data distribution at all scales. This suggests that ESRGAN is able to capture the true data distribution.

If we remove the adversarial loss from ESRGAN and just optimize the MSE of the generated pixels, we see performance similar to SR-CNN. This suggests that the adversarial loss only helps match the higher frequencies of the data, and hurts otherwise.

4 PSD-Net

We can use these observations to create a better model. The adversarial loss optimizes PSD indirectly, but PSD is a differentiable function. Formally, $PSD(x) := |F(x)|^2$, where F is the 2D real-to-complex Discrete Fourier Transform.

This lets us replace the adversarial loss term with PSD to directly optimize the spectra, in a fully-supervised manner. Concretely, we take the generator G from ESRGAN, remove the discriminator completely, and train using a loss of the form Loss(l,h) := MSE(G(l),h) + MSE(PSD(G(l)), PSD(h)), where l and h are a batch of low and high-resolution images, G is our generator. We then backpropagate to optimize both MSE and PSD directly. We call this model PSD-Net, and it gives the best metrics of all (Table 1), along with visually superior images (see Figure 2) and a closely matching PSD plot (Figure 3b).

PSD-Net trains considerably faster than ESRGAN, as it is stable with larger batch sizes (we used a batch size of 128), and does not require training a discriminator.

5 Future work

We plan to see if PSD-based losses can be used in other areas of physical science to achieve similar results. With respect to the wind problem, our technique does a good job reproducing single images. However, we do not currently deal with a sequence of images over time or space. For example, capturing the effects of winds over a larger surrounding region, e.g. from a coarse climate model, would help in regional climate prediction. In addition, being able to capture a sequential time series would also be beneficial for capturing synoptic weather patterns such as fronts or tropical cyclones. Both will be the goals of future work.

We also plan to incorporate additional variables such as temperature and pressure, and to see if models based on attention mechanisms can improve accuracy.

References

- T. Vandal, E. Kodra, S. Ganguly, A. R. Michaelis, R. R. Nemani, and A. R. Ganguly, "Deepsd: Generating high resolution climate change projections through single image super-resolution," *CoRR*, vol. abs/1703.03126, 2017.
- [2] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. Change Loy, "Esrgan: Enhanced super-resolution generative adversarial networks," *ECCV18*, September 2018.
- [3] G. Bürger, S. Sobie, A. Cannon, A. Werner, and T. Murdock, "Downscaling extremes an intercomparison of multiple methods for future climate," *Journal of Climate*, vol. 26, 2013.

- [4] M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, and Prabhat, "Deep learning and process understanding for data-driven earth system science," *Nature*, vol. 566, no. 7743, pp. 195–204, 2019.
- [5] P. Vahmani and A. D. Jones, "Water conservation benefits of urban heat mitigation," *Nature Communications*, vol. 8, no. 1, p. 1072, 2017.
- [6] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *CoRR*, vol. abs/1501.00092, 2015.