Super Resolution Reconstruction of E3SM Data Using a FSRCNN

Linsey Sara Passarella[1], Salil Mahajan[2], Anikesh Pal[3], and Matthew Ross Norman[2]

[1]University of Tennessee
[2]Oak Ridge National Laboratory (DOE)
[3]Indian Institute of Technology Kanpur

November 26, 2022

Abstract

We present a first application of a fast super resolution convolutional neural network (FSRCNN) approach for downscaling climate simulations. Unlike other SR approaches, FSRCNN uses the same input feature dimensions as the low resolution input. This allows it to have smaller convolution layers, avoiding over-smoothing, and reduced computational costs. We further adapt FSRCNN to feature additional convolution layers after the deconvolution layer, we term FSRCNN-ESM. We use high-resolution (0.25°) monthly averaged model output of five surface variables over a part of North America from the US Department of Energy's Energy Exascale Earth System Model's control simulation. These high-resolution and corresponding coarsened low-resolution (1°) pairs of images are used to train the FSRCNN-ESM and evaluate its use as a downscaling approach. We find that FSRCNN-ESM outperforms FSRCNN and other methods in reconstructing high resolution images producing finer spatial scale features with better accuracy for surface temperature, surface radiative fluxes and precipitation.

Super Resolution Reconstruction of E3SM Data Using a FSRCNN

Linsey S. Passarella[1], Salil Mahajan[1], Anikesh Pal[2], Matthew R. Norman[1]

[1]Computational Earth Sciences, Oak Ridge National Laboratory, Oak Ridge, TN, USA
[2]Department of Mechanical Engineering, Indian Institute of Technology, Kanpur 208016, India

Key Points:

- We present a fast super resolution convolutional neural network (FSRCNN) based approach for downscaling gridded earth system model data.
- FSRCNN-ESM's reconstruction of high resolution spatial patterns improves upon both traditional and machine learning downscaling methods.
- FSRCNN-ESM is computationally less expensive to train over other machine learning downscaling methods.

Corresponding author: Linsey S Passarella, `lpassare@vols.utk.edu`

Abstract

We present a first application of a fast super resolution convolutional neural network (FSRCNN) based approach for downscaling earth system model (ESM) simulations. Unlike other SR approaches, FSRCNN uses the same input feature dimensions as the low resolution input. This allows it to have smaller convolution layers, avoiding over-smoothing, and reduced computational costs. We further adapt FSRCNN to feature additional convolution layers after the deconvolution layer, we term FSRCNN-ESM. We use high-resolution ($\sim$0.25°) monthly averaged model output of five surface variables over a part of North America from the US Department of Energy's Energy Exascale Earth System Model's control simulation. These high-resolution and corresponding coarsened low-resolution ($\sim$1°) pairs of images are used to train the FSRCNN-ESM and evaluate its use as a downscaling approach. We find that FSRCNN-ESM outperforms FSRCNN and other super-resolution methods in reconstructing high resolution images producing finer spatial scale features with better accuracy for surface temperature, surface radiative fluxes and precipitation.

Plain Language Summary

High resolution global climate data is computationally expensive to run but crucial for assessing climate change effects at local and regional scales. Here, we adapt a new deep learning technique, called fast super-resolution convolutional neural network, to remap climate data from low resolution to high resolution grids. This approach is faster and more accurate for statistical downscaling climate data compared to other prevalent methods.

1 Introduction

Accurate and reliable climate data is critical for assessing the risk of climate change to our society's well-being. Increases in temperature, sea-level, and extreme weather events can render many aspects of our society vulnerable including our health, natural resources and energy-systems (Nicholls & Cazenave, 2010; Trenberth, 2012). Local and regional climate future projection data is the most crucial for planning and mitigating these risks, but is often the least reliable (Schmidt, 2010). Current Earth System Models (ESMs) used for simulating Earth's past climate and future projections are often computationally limited to coarse horizontal resolutions, generally between 1° to 3° (Vandal et al., 2017). These low resolution models fail to accurately simulate important physical processes such as precipitation extremes (Kharin et al., 2007). Recent advances in computing resources have allowed for global ESMs to be run at higher resolutions (∼0.25°) for longer time periods and have shown to improve the simulations of regional mean climate as well as extremes (Wehner et al., 2010a; Mahajan et al., 2015). However, these high resolution models remain prohibitively expensive.

A computationally less expensive approach to derive high resolution climate data over a region of interest is to map data from low resolution global model simulations to high resolution grids using dynamical or statistical downscaling techniques. Dynamical downscaling involves running high resolution regional dynamical models to extrapolate large scale boundary conditions obtained from a coarser global ESM to finer resolutions on regional scales. Statistical downscaling (SD) aims to map coarse resolution data to high resolution projections using statistical methods like linear regression. Recent studies have shown that machine learning techniques, like neural networks (Vu et al., 2016; Fistikoglu & Okkan, 2011) and support vector machines (Ghosh, 2010), for SD significantly outperform other traditional SD methods. In this study, we use one such computer vision approach called super-resolution (SR), which generates a high resolution image from its low resolution equivalent. SR techniques attempt to generalize across images and have been shown to learn local scale patterns more efficiently than other downscaling methods (Vandal et al., 2017).

One pioneering work in SR deep learning is a SR convolutional neural network (SRCNN). A convolutional neural network is a type of artificial neural network that convolves a kernel with the data to extract features for further use in the overall neural network architecture (LeCun et al., 1989, 2015; Goodfellow et al., 2016). The SRCNN was originally proposed by Dong, Loy, He, and Tang (2014) and was shown to achieve significantly better performance over other traditional and state of the art SR methods. Vandal et al. (2018) demonstrated the usefulness of using a stacked SRCNN, called DeepSD, to downscale ensemble ESMs and showed that it outperforms other methods including bias correction spatial disaggregation (BCSD), artificial neural networks (ANN), Lasso and support vector machines (SVM). Several recent studies have used similar super resolution approaches to downscale ESMs to higher resolutions and demonstrated significant skill. These include a super resolution general adversarial network (SRGAN) (Stengel et al., 2020) to downscale wind and solar radiation fields to 50x resolutions and a Laplacian pyramid super-resolution network, termed, ResLap, that uses a residual dense block to allow hierarchical feature extraction from the convolutional layers (Cheng et al., 2020).

One common feature of these SR approaches is a pre-processing step where the low resolution images are pre-interpolated to the desired high-resolution output image size (say, using bilinear interpolation) before running the network. Dong, Loy, and Tang (2016) developed a method, termed as fast SRCNN (FSRCNN), that alleviates this pre-processing and replaces it with a deconvolution layer at the end, facilitating mapping directly at the resolution of the low resolution image. Since the computational complexity of a CNN is proportional to the input image size, this lowers the computational cost of the network significantly - almost by a factor of n^2 compared to a SRCNN, where n is the downscaling factor. Further, the smaller input image size in FSRCNN implies that narrower filters can capture the same information, thus allowing for more filters for greater feature extractions while also lowering computational cost. This use of smaller kernel sizes to improve CNN models was proved in Simonyan and Zisserman (2014). The FSRCNN has been shown to improve image reconstruction skill significantly compared to a general SRCNN for high downscaling factors, with the convolution-deconvolution structure reducing edge smoothing and loss of detail and improving feature reconstruction. Further, the same FSRCNN network can be used for different upscaling factors with only the deconvolution layer needing further tuning.

Here, we present a first attempt to apply FSRCNN for ESM downscaling and find that it is generally more skillful than DeepSD. Further, we improve upon the FSRCNN by adapting it to use additional SRCNN-like convolutional layers after the deconvolution step. By adding these additional convolutional layers, we are able to extract more information and finer spatial details in the high resolution images. We refer to this new adapted FSRCNN architecture as FSRCNN-ESM. Following previous validation studies (Stengel et al., 2020) of the application of super resolution approaches to climate data, we also reconstruct high resolution data from a coarsened version of the same data. We evaluate the reconstruction results using an objective evaluation metrics like the mean bias error, and find FSRCNN-ESM to be a promising downscaling method with superior skills as compared to both DeepSD and FSRCNN. Section 2 describes the ESM data used as well as the improved FSRCNN network architecture in more detail. We present the results of our objective and subjective evaluations in Section 3 and summarize our work in Section 4.

2 Methods

2.1 Data Collection

For this study we use monthly output of a 30-year segment of the 1950-control simulation with the global high resolution (0.25°) configuration of the Energy Exascale Earth Systems Model (E3SM) (E3SM Project, 2018). Model data was obtained from the Earth System Grid Federation (ESGF) (Cinquini et al., 2014). It should be noted that E3SM data is bilinearly interpolated from its native non-orthogonal cubed-sphere grid to an equivalent regular 0.25°x0.25° longitude-latitude grid. We call this model data, E3SM-HR. This high resolution data is interpolated to a 1°x1° grid using a bicubic method to obtain the corresponding low resolution input images, which lose the fine scale features present in the high resolution data. Our goal is to reconstruct the high resolution images back from these coarsened data using FSRCNN. When using a computer vision approach to gridded E3SM data, we can think of each grid point as a pixel in an image. For this study, we look at a subset of E3SM data corresponding to North America. The low-resolution images as a result are 60×60 pixels, while high-resolution images are four times larger across both dimensions and have a size of 240×240 pixels each.

We chose five variables to test the FSRCNN - surface temperature (TS), shortwave heat flux (FSNS), longwave heat flux (FLNS), precipitation convective rate (PRECC) and the large scale precipitation rate (PRECL). This results in 360 images for one variable, or 1,800 images in total. We use all the variables together, each normalized using the min-max scaler, in a single one-channel network when training so our algorithm can learn how to extract multiple spatial features. The addition of multiple variables in one network enhanced our reconstruction. For example, when training just PRECC on a single network, the average mean square error (MSE) for testing was $7.58e{-}7$. When using all variables together, the MSE on the testing dataset for PRECC is $2.79e{-}11$ (see table 1). It is common in computer vision to learn many different classifications of images in a single algorithm to improve feature reconstruction. Before training the model, we split the last 3 years of data into a testing set, corresponding to 180 images or 36 images for a single variable.

We also explore the use of elevation as a second input channel to our methods. Several studies have shown the addition of elevation as an input is important for enhancing the reconstruction quality of precipitation data (Vandal et al., 2017; Liu et al., 2020).

2.2 Deep Learning

2.2.1 SRCNN and DeepSD

The SRCNN is trained to directly learn practical mappings between low resolution and high resolution images with little pre- and post-processing (Dong et al., 2014). The low resolution image must be interpolated to the desired output size before training using a bicubic method. The SRCNN consists of three operations: patch extraction and representation, non-linear mapping and reconstruction (Dong et al., 2014). The goal of an SRCNN is to take the low resolution image Y and generate a high resolution image G that is close to the ground truth image I.

Layer 1 of the SRCNN consists of 64 filters of 9x9 kernels, layer 2 has 32 filters of 1x1 kernels and the output layer has a single filter with a 5x5 kernel, the same as described in Dong et al. (2014). We trained this SRCNN on 100 epochs with an adam loss optimizer, sigmoid activation function and an initial learning rate of 0.001 using tensorflow. The choice of a sigmoid activation function allows for back propagation to return an output value between 0 and 1 which is useful in this context, since we normalize the images before training with a min-max scaler.

Vandal et al. (2017) uses a stacked SRCNN apporach called DeepSD and here we test that method against the FSRCNN-ESM. DeepSD uses elevation as a second input channel for the SRCNNs to train on. The first SRCNN is used to interpolate the images from a 1° resolution to 0.5°. The estimated 0.5° resolution images are then passed to the next SRCNN and interpolated to the final 0.25° images. Here, we only use 2 stacked SRCNN, as opposed to 3 in the original paper, since we are only downscaling by a factor of 4. It is important to note that Vandal et al. (2017) used the DeepSD method to downscale one variable, precipitation, and here we are using it to downscale five.

2.2.2 General FSRCNN

The basic FSRCNN method was created to accelerate the SRCNN process and the redesign involved three features: (1) smaller convolutional kernels but more feature maps,

(2) an added deconvolutional layer, and (3) the input feature dimension is the same as the original low resolution image (Luo et al., 2019). Because the FSRCNN takes the original 60×60 image as the input and does not have to interpolate it to a 240×240 image, it learns 16 times fewer weights than the SRCNN and as a result is much faster when training. The final reconstruction to the HR image then requires a deconvolutional layer at the end to remap the data from the low resolution reconstruction steps to a higher resolution grid. We note that the deconvolutional layer is not the same as an unpooling plus convolutional layer, sometimes known as convolutional interpolation, which in its purest form resizes an image by copying pixels as many times as needed to achieve the desired image size before passing through a convolutional layer. A deconvolutional layer, also known as a transpose convolutional layer, instead pads the image with zeros to desired size before upsampling the image using learned kernels. Dong et al. (2016) found that replacing the deconvolutional layer with an unpooling layer resulted in a significant drop in reconstruction quality.

The FSRCNN can be broken down into five parts: feature extraction, shrinking, nonlinear mapping, extension, and deconvolution (Dong et al., 2016). The feature extraction step in Dong et al. (2016) consists of a 5 by 5 filter size with d number of filters. The number of filters can be thought of as the number of desired learned features in the low resolution image. The shrinking step is a 1 by 1 filter with s number of filters, here $s < d$, that acts as a way to condense the number of features found in step 1. The nonlinear mapping step maps the features in step 2 nonlinearly to a new set of features. It uses multiple layers of nonlinear mapping with a filter size of 3 by 3. By selecting smaller convolutional kernels but more feature maps (large d), the FSRCNN learns more non-linear features in the data and creates better SR reconstruction results compared to the SRCNN. The FSRCNN then moves on to the expansion layer, which acts like an inverse of the shrinking layer, to generate a larger number of feature maps to improve high resolution reconstruction. Finally, the FSRCNN uses the deconvolution step to achieve the final high resolution image.

2.2.3 FSRCNN-ESM Architecture

Here, we expand upon the basic FSRCNN method to maximize accuracy for image reconstruction for E3SM data. We include additional convolutional layers after the deconvolution step in the FSRCNN - an added patch extraction step consisting of 64 kernels, and a nonlinear mapping step with 32 kernals, similar to those in the SRCNN - both of which are applied to the high resolution data generated after the deconvolution step as shown in Fig. 1. Fig. 1 shows the SRCNN, original FSRCNN and the new configuration. We refer to this new network as FSRCNN-ESM. We found that these additions to the network further improve image reconstruction in our data, as determined by a loss function. The loss is calculated by using a pixel-wise MSE using the following equation:

$$MSE(I, G) = \frac{1}{N} \sum_{i,j} (I_{i,j} - G_{i,j})^2 (1)$$

where I is the original high resolution image and G is the generated high resolution image and i and j denote the location of the pixel. The training loss decreased by 50% with the added feature extraction step after the deconvolutional layer in the FSRCNN-ESM compared to the original FSRCNN architecture. Similarly, the average reconstruction loss in MSE for the FSRCNN was 0.000568 compared to the average MSE for the FSRCNN-ESM at 0.000261.

We also evaluate the impact of the addition of elevation as a second input on reconstruction quality. Table 1 shows the MSE for variables in the test set for the FSRCNN, FSRCNN plus elevation, FSRCNN-E3SM and FSRCNN-E3SM plus elevation. Elevation improved the simple FSRCNN method, but we did not see the same improvement for the FSRCNN-ESM. Therefore, we add elevation as an input channel to the FSRCNN but not our FSRCNN-ESM in our further evaluations with the testing data.

2.3 Evaluation Metrics

We evaluate the mean absolute error (MAE) for each reconstruction method on each variable the testing dataset. Before computing the MAE, we scale the variables back to their original values using min-max scaler. We define the MBE as follows:

$$MAE = \frac{1}{N} \sum_{1}^{N} |P_i - A_i| \qquad (2)$$

where P_i is the predicted image i and A_i is the actual truth image. The MAE for each variable is expressed in the variable's units. Here N is 36, or the number of images per variable in the testing dataset.

We also evaluate the skill of each method's ability to reconstruct high resolution images by computing the L_1 and L_2 error for each sample as follows:

$$L_1 = \frac{\sum_{i,j} |G_{i,j} - I_{1,j}|}{\sum_{i,j} |I_{1,j}|}, L_2 = \frac{\sum_{i,j} (G_{i,j} - I_{1,j})^2}{\sum_{i,j} (I_{1,j})^2} \tag{3}$$

We compute both the L_1 and L_2 errors and the MBE across each variable on the held out testing dataset.

3 Results

We compare the FSRCNN-ESM method as applied to E3SM data with the DeepSD method (Dong et al., 2014) and with a basic bicubic interpolation using the above stated metrics for determining image reconstruction quality: MSE, MAE, L1- and L2-error. We also compare the computational training time of the FSRCNN and FSRCNN-E3SM against DeepSD.

3.1 Reconstruction Evaluation

Figure 2 show samples of reconstruction of high resolution images over a part of Northern America from low-resolution images using bicubic interpolation, DeepSD, FSRCNN and FSRCNN-ESM approaches. We randomly pick one sample of a summer (June) month from 3 years of reconstruction test data for each of the five variables to visually illustrate the reconstruction quality. The inset plot shows a zoomed in portion to better visualize some of the finer spatial details of some region, generally over the Rocky Mountains and over Northern Andes - regions that show strong gradients in the high resolution images owing to the topography there. The plots (first column) show the loss of fine-scale features as the high resolution images (last column) are interpolated to low-resolution images. To restate, the goal is then to reconstruct the high resolution image from these low resolution images. It is clear that a bicubic interpolation to downscale performs poorly. The DeepSD generally improves over the bicubic interpolation, but still lacks the finer-scale details noted in the high resolution images - this is apparent in most of the images for all variables. For example, the DeepSD is not able to discern the strong gradients over the Northern Andes clearly for convective precipitation (PRECC) (Figure 2, second and fourth rows). Similarly, it is not able to capture the fine scale features of surface temperature (TS), net surface shortwave radiation (FSNS), net surface longwave radiation flux (FLNS) and large-scale precipitation (PRECL) over North America in the summer (Figure 2, first, third and fifth rows). It is clear visually that FSRCNN-ESM better generates finer details in high resolution data when compared to the DeepSD and bicubic interpolation, which both tend to over-smooth the images. For the above stated examples, the FSRCNN-ESM is able to isolate the strong gradients in PRECC over the Northern Andes and in TS, FSNS, FLNS and PRECL over North America in the summer.

To quantify the reconstruction skill of these methods, we use the MBE metric over the testing dataset. Table 2 summarizes the results of our analysis for the three methods. We calculate the MBE for each variable image over each month in the testing dataset. Consistent with the visual illustrations (Figure 2), the FSRCNN-ESM generally outperforms the other methods in terms of this metric across most variables and months, with the exception of FSNS in which the FSRCNN had a higher reconsturction skill on average. Dong et al. (2016) noted that a large fraction of the increase in reconstruction results came from replacing the bicubic interpolation step in the SRCNN with the deconvolution layer in FSRCNN when they partitioned their error metric into different steps. We see this here with the increase in skill from DeepSD to both the FSRCNN and FSRCNN-ESM. It is interesting to note that the bicubic interpolation skill is comparable to DeepSD for our data, and exceeds it for most variables. Most methods tend to underestimate precipitation variables. Precipitation occurs intermittently and is a highly non-linear process and results from multi-scale, multi-phase physical processes creating large spatial heterogeneity. The above suggest that image reconstruction becomes more difficult as spatial heterogeneity increases.

We also use the L_1 and L_2 error metrics to quantify reconstruction skill of the various methods on our held out testing dataset. Figure 4 shows the histogram of the L_1 and L_2 errors for FSRCNN-ESM (a,e), the original FSRCNN (b,f), DeepSD (c,g) and bicubic (d,h) respectively.The majority of our FSRCNN-ESM predicted samples (Figure 4a) have a L_1 error less than 10% and an L_2 error less than 1% suggesting the effectiveness of our FSRCNN-ESM for interpolating E3SM data to high resolution grids. Our FSRCNN-ESM also generally achieves better reconstruction skill compared to the other two methods based on both the metrics, consistent with the findings using the MBE metric.

3.2 Computational Performance

Figure 4j shows the cumulative CPU computational time required to train FSRCNN-ESM (blue), DeepSD (red) and FSRCNN (green) on 100 epochs. FSRCNN-ESM completed training 600% faster than the DeepSD, comparable to the decrease in training time with the FSRCNN, while still maintaining similar loss values (Figure 3i). This is largely due to the fact that the DeepSD has a pre-processing step of interpolating low resolution images to the desired output size before training, while the FSRCNN-ESM uses the low resolution image size. As stated earlier this allows using narrower filters on the smaller images, reducing the total number of parameters and computations. Once trained, however, the DeepSD and FSRCNN-ESM take roughly the same amount of time, about 25 seconds, to downscale all 1,800 images.

Dong et al. (2016) showed that an FSRCNN with similar number of steps as a SRCNN produced a speed-up of about 40x when upscaling by a factor of three. This was largely (about 30x) due to the narrower filters in the FSRCNN that led to the large difference between the number of trainable parameters in the nonlinear mapping step of SRCNN and that in the corresponding three steps (combination of shrinking, non-linear mapping and expanding steps) in FSRCNN. The use of the low-res image as input in the FSRCNN contributed to the remaining reduction in computational cost of network training. The computation of FSRCNN-ESM only provides a speed up of 6x over the DeepSD for a upscaling factor of four. The increase in cost of FSRCNN-ESM as compared to the base FSRCNN is due to the addition of new patch extraction, nonlinear mapping and reconstruction steps after the deconvolution step in FSRCNN-ESM. The combination of these new steps is equivalent to a full SRCNN in itself that uses the full HR image size as an input, but with narrower filters. The use of narrower filters in these steps then still allows for a faster training of the overall FSRCNN-ESM network as compared to the DeepSD.

4 Summary and Discussion

We apply a novel super-resolution based approach for downscaling ESM data that uses a modified version of the FSRCNN method. We find that this FSRCNN-ESM is able to map low resolution climate images to a four times higher resolution with a better skill than DeepSD, FSRCNN and bicubic interpolation; for surface radiative fluxes and large scale and convective precipitation; while remaining computationally inexpensive to train. Our FSRCNN-ESM is also able to downscale images in a single-step process and without need for access to GPUs for training. The FSRCNN-ESM as a result is a more approchable method of downscaling using machine learning.

This study focused on reconstructing monthly averaged images. In the future, we will explore the application of FSRCNN-ESM for downscaling higher temporal resolution data like daily and sub-daily data, and target extreme events. Also, FSRCNN-ESM does not consider the concurrency of images of the different variables - the network is agnostic of the presence of data for other variables. In the future, we plan to use multichannel approaches (Vandal et al., 2018; Stengel et al., 2020; Cheng et al., 2020) using all the concurrent variables simultaneously in the network for reconstruction allowing the network access to more correlated data, which may improve reconstruction skill. We also plan to evaluate the FSRCNN-ESM against other recent machine learning approaches that have been used for downscaling, for example, SRGAN (Stengel et al., 2020), and ResLap (Cheng et al., 2020), all of which have been shown to perform better than the SRCNN or other approaches for different variables, regions and scaling factors.This study is the first to use FSRCNN for ESM data and we applied the original version of FSRCNN here, and with some improvements (FSRCNN-ESM), to demonstrate its utility. We plan to explore the applicability of the latest advances in FSRCNN and SRCNN to ESM downscaling in the future. Some advanced applications include using GANs as a latent blank to improve image reconstruction on an SRCNN (Chan et al., 2021), using a student-teacher supervised learning approach to training (L. Wang & Yoon, 2022), applying skip connections to alleviate the vanishing gradient problem (Zou et al., 2021), and the use of a multi-path residual to improve efficiency in the SRCNN (Q. Wang et al., 2021).

The resolution of finer scales in the high resolution model and scale-agnostic nature of current sub-grid scale physical parameterizations used in climate models imply that a low-resolution model simulation is not statistically equivalent to coarsened data from a high-resolution model. For example, the simulation of precipitation extremes is found to be stronger in high resolution simulations than the low resolution simulations, even after conservative mapping of high-resolution simulation data to the low-resolution grid (Wehner et al., 2010b, 2014; Mahajan et al., 2015). In order to apply FSRCNN-ESM directly to low-resolution model output to generate high resolution images at the skill of a high resolution simulation, we would thus require a bias-correction step. We plan to explore traditional bias-correcting methods as well as utilize machine learning approaches for it. We would also explore the use of nudged simulations or regionally refined models to generate equivalent pairs of low-resolution and high-resolution model simulations that could be used for training a FSRCNN-ESM network.

Acknowledgments

The E3SMv1 data used in this study is freely available through the Earth System Grid Federation (ESGF) distributed archives via `https://doi.org/10.1029/2018MS001603` and is available through the ESGF interface `https://esgf-node.llnl.gov/projects/e3sm/` (E3SM Project, 2018).

This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231. This research was supported as part of the Energy Exascale Earth System Model (E3SM) project, funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research.

References

Chan, K. C., Wang, X., Xu, X., Gu, J., & Loy, C. C. (2021). Glean: Generative latent bank for large-factor image super-resolution. In *Proceedings of the ieee/cvf conference on computer vision and pattern recognition* (pp. 14245–14254).

Cheng, J., Kuang, Q., Shen, C., Liu, J., Tan, X., & Liu, W. (2020). Reslap: Generating high-resolution climate prediction through image super-resolution. *IEEE Access*, *8*, 39623–39634.

Cinquini, L., Crichton, D., Mattmann, C., Harney, J., Shipman, G., Wang, F., ... others (2014). The earth system grid federation: An open infrastructure for access to distributed geospatial data. *Future Generation Computer Systems*, *36*, 400–417.

Dong, C., Loy, C. C., He, K., & Tang, X. (2014). Learning a deep convolutional network for image super-resolution. In *European conference on computer vision* (pp. 184–199).

Dong, C., Loy, C. C., & Tang, X. (2016). Accelerating the super-resolution convolutional neural network. In *European conference on computer vision* (pp. 391–407).

E3SM Project. (2018, April). *Energy Exascale Earth System Model (E3SM).* [Computer Software] `https://dx.doi.org/10.11578/E3SM/dc.20180418.36`. Retrieved from `https://dx.doi.org/10.11578/E3SM/dc.20180418.36` doi: 10.11578/E3SM/dc.20180418.36

Fistikoglu, O., & Okkan, U. (2011). Statistical downscaling of monthly precipitation using ncep/ncar reanalysis data for tahtali river basin in turkey. *Journal of Hydrologic Engineering*, *16*(2), 157–164.

Ghosh, S. (2010). Svm-pgsl coupled approach for statistical downscaling to predict rainfall from gcm output. *Journal of Geophysical Research: Atmospheres*, *115*(D22).

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning.* MIT press.

Kharin, V. V., Zwiers, F. W., Zhang, X., & Hegerl, G. C. (2007). Changes in temperature and precipitation extremes in the ipcc ensemble of global coupled model simulations. *Journal of Climate*, *20*(8), 1419–1444.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, *521*(7553), 436–444.

LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W., & Jackel, L. (1989). Handwritten digit recognition with a back-propagation network. *Advances in neural information processing systems*, *2*.

Liu, Y., Ganguly, A. R., & Dy, J. (2020). Climate downscaling using ynet: A deep convolutional network with skip connections and fusion. In *Proceedings of the 26th acm sigkdd international conference on knowledge discovery & data mining* (pp. 3145–3153).

Luo, Z., Yu, J., & Liu, Z. (2019). The super-resolution reconstruction of sar image based on the improved fsrcnn. *The Journal of Engineering*, *2019*(19), 5975–5978.

Mahajan, S., Evans, K. J., Branstetter, M., Anantharaj, V., & Leifeld, J. K. (2015). Fidelity of precipitation extremes in high resolution global climate simulations. *Procedia Computer Science*, *51*, 2178–2187.

Nicholls, R. J., & Cazenave, A. (2010). Sea-level rise and its impact on coastal zones. *science*, *328*(5985), 1517–1520.

Schmidt, G. (2010). The real holes in climate science. *Nature*, *463*, 21.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

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

Trenberth, K. E. (2012). Framing the way to relate climate extremes to climate change. *Climatic change*, *115*(2), 283–290.

Vandal, T., Kodra, E., Ganguly, S., Michaelis, A., Nemani, R., & 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* (pp. 1663–1672).

Vandal, T., Kodra, E., Ganguly, S., Michaelis, A., Nemani, R., & Ganguly, A. R. (2018). Generating high resolution climate change projections through single image super-resolution: An abridged version. In *International joint conferences on artificial intelligence organization.*

Vu, M. T., Aribarg, T., Supratid, S., Raghavan, S. V., & Liong, S.-Y. (2016). Statistical downscaling rainfall using artificial neural network: significantly wetter

Table 1. The MSE for each variable in the test data set for the FSRCNN, FSRCNN plus elevation, FSRCNN-ESM and FSRCNN-ESM plus elevation.

	TS	FSNS	FSNT	PRECC	PRECL
FSRCNN	3.251e-5	4.377e-5	3.563e-5	3.299e-5	3.767 e-5
FSRCNN + elevation	**3.189e-5**	**4.285e-5**	**3.499e-5**	**3.246e-5**	**3.709e-5**
FSRCNN-ESM	**2.628e-5**	**3.582e-5**	**3.026e-5**	**2.796e-5**	**3.233e-5**
FSRCNN-ESM + elevation	3.336e-5	4.467e-5	3.841e-5	3.614e-4	4.167 e-5

bangkok? *Theoretical and applied climatology*, *126*(3-4), 453–467.

Wang, L., & Yoon, K.-J. (2022). Semi-supervised student-teacher learning for single image super-resolution. *Pattern Recognition*, *121*, 108206.

Wang, Q., Gao, Q., Wu, L., Sun, G., & Jiao, L. (2021). Adversarial multi-path residual network for image super-resolution. *IEEE Transactions on Image Processing*, *30*, 6648–6658.

Wehner, M. F., Reed, K. A., Li, F., Prabhat, Bacmeister, J., Chen, C.-T., … Jablonowski, C. (2014). The effect of horizontal resolution on simulation quality in the Community Atmospheric Model, CAM5.1. *Journal of Advances in Modeling Earth Systems*, *6*(4), 980–997. Retrieved from `http://dx.doi.org/10.1002/2013MS000276` doi: 10.1002/2013MS000276

Wehner, M. F., Smith, R. L., Bala, G., & Duffy, P. (2010a). The effect of horizontal resolution on simulation of very extreme us precipitation events in a global atmosphere model. *Climate dynamics*, *34*(2-3), 241–247.

Wehner, M. F., Smith, R. L., Bala, G., & Duffy, P. (2010b). The effect of horizontal resolution on simulation of very extreme us precipitation events in a global atmosphere model. *Climate Dynamics*, *34*(2-3). Retrieved from `http://dx.doi.org/10.1007/s00382-009-0656-y`

Zou, Y., Zhang, L., Liu, C., Wang, B., Hu, Y., & Chen, Q. (2021). Super-resolution reconstruction of infrared images based on a convolutional neural network with skip connections. *Optics and Lasers in Engineering*, *146*, 106717.

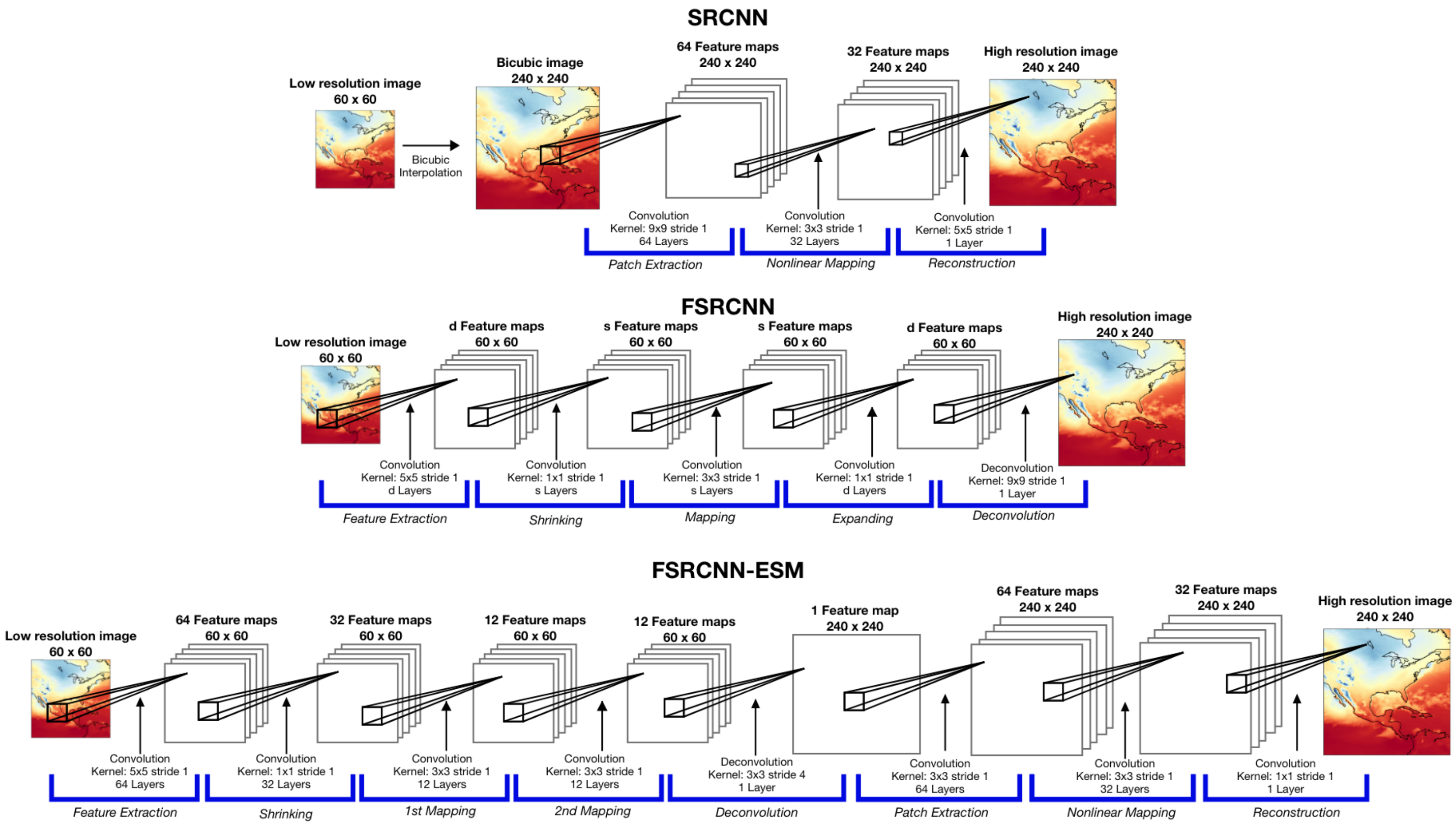
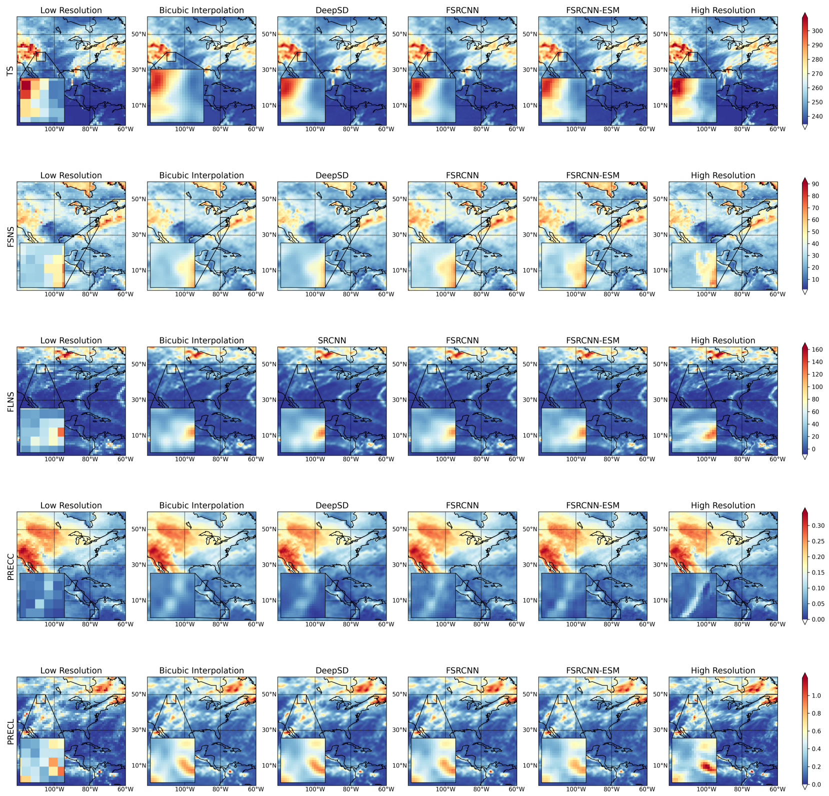


Figure 1. The network architecture of the SRCNN (24,513 trainable parameters), FSRCNN (15,681 trainable parameters) and our expanded FRSCNN-ESM (34,246 trainable parameters) used in this study. We add feature extraction and nonlinear mapping layers after the deconvolutional layer because it was shown to significantly enhance image reconstruction in our E3SM data.



463 **Figure 2.** Super resolution reconstruction results for a sample in January.

Table 2. The mean bias error (MBE) for surface temperature (TS K) , shortwave heat flux (FSNS w/m^2), longwave heat flux (FLNS w/m^2), precipitation convective rate (PRECC mm/day) and the large scale precipitation rate (PRECL mm/day). The PSNR for the variables is averaged across each month and presented here in terms of logarithmic decibel scale.

	Bicubic TS	DeepSD TS	FSRCNN TS	FSRCNN-ESM TS	Bicubic FSNS	DeepSD FSNS	FSRCNN FSNS	FSRCNN-ESM FSNS	
Jan	0.377	0.499	**0.013**	0.061	3.869	3.208	1.179	**1.128**	
Feb	0.409	0.511	0.106	**0.074**	3.116	2.527	0.949	**0.863**	
Mar	0.314	0.437	0.117	**0.046**	1.736	1.116	**0.070**	0.520	
Apr	0.294	0.417	0.069	**0.003**	0.318	**0.289**	0.620	0.788	
May	0.380	0.504	0.048	**0.011**	**0.319**	0.978	0.617	0.998	
Jun	0.432	0.558	0.063	**0.027**	**0.215**	0.847	0.479	0.499	
Jul	0.469	0.586	0.141	**0.031**	**0.058**	0.567	0.313	0.312	
Aug		0.430	0.547	0.125	**0.006**	0.606	**0.021**	0.238	0.056
Sep	0.394	0.514	0.048	**0.017**	1.304	0.676	0.523	**0.356**	
Oct	0.260	0.392	0.070	**0.023**	2.502	1.886	1.162	**0.926**	
Nov	0.188	0.335	0.047	**0.030**	3.392	2.715	1.029	**1.028**	
Dec	0.327	0.453	0.072	**0.046**	4.088	3.421	1.124	**1.096**	

	Bicubic FLNS	DeepSD FLNS	FSRCNN FLNS	FSRCNN-ESM FLNS	Bicubic PRECC	DeepSD PRECC	FSRCNN PRECC	FSRCNN-ESM PRECC
Jan	0.409	0.217	0.187	**0.164**	0.071	0.041	0.038	**0.033**
Feb	0.361	0.169	0.141	**0.115**	0.091	0.061	0.058	**0.053**
Mar	0.302	0.110	0.085	**0.057**	0.092	0.062	0.059	**0.054**
Apr	0.269	0.077	0.054	**0.024**	0.079	0.049	0.046	**0.041**
May	0.222	0.030	**0.009**	0.023	0.023	**0.006**	0.009	0.014
Jun	0.261	0.069	0.046	**0.015**	**0.016**	0.046	0.049	0.054
Jul	0.288	0.096	0.072	**0.042**	**0.033**	0.063	0.066	0.071
Aug	0.294	0.102	0.077	**0.048**	0.033	0.063	0.066	0.072
Sep	0.365	0.173	0.145	**0.120**	**0.041**	0.070	0.073	0.079
Oct	0.362	0.170	0.143	**0.117**	**0.031**	0.061	0.064	0.069
Nov	0.435	0.243	0.211	**0.190**	**0.013**	0.016	0.019	0.024
Dec	0.439	0.247	0.215	**0.194**	0.054	0.024	0.021	**0.016**

	Bicubic PRECL	DeepSD PRECL	FSRCNN PRECL	FSRCNN-ESM PRECL
Jan	0.106	0.062	0.057	**0.049**
Feb	0.137	0.093	0.088	**0.080**
Mar	0.138	0.094	0.089	**0.081**
Apr	0.119	0.074	0.069	**0.062**
May	0.034	0.009	0.014	0.022
Jun	**0.025**	0.069	0.074	0.082
Jul	**0.050**	0.094	0.099	0.107
Aug	**0.051**	0.095	0.100	0.107
Sep	**0.061**	0.106	0.110	0.118
Oct	**0.047**	0.091	0.096	0.104
Nov	**0.020**	0.024	0.029	0.036
Dec	0.081	0.036	0.032	**0.024**

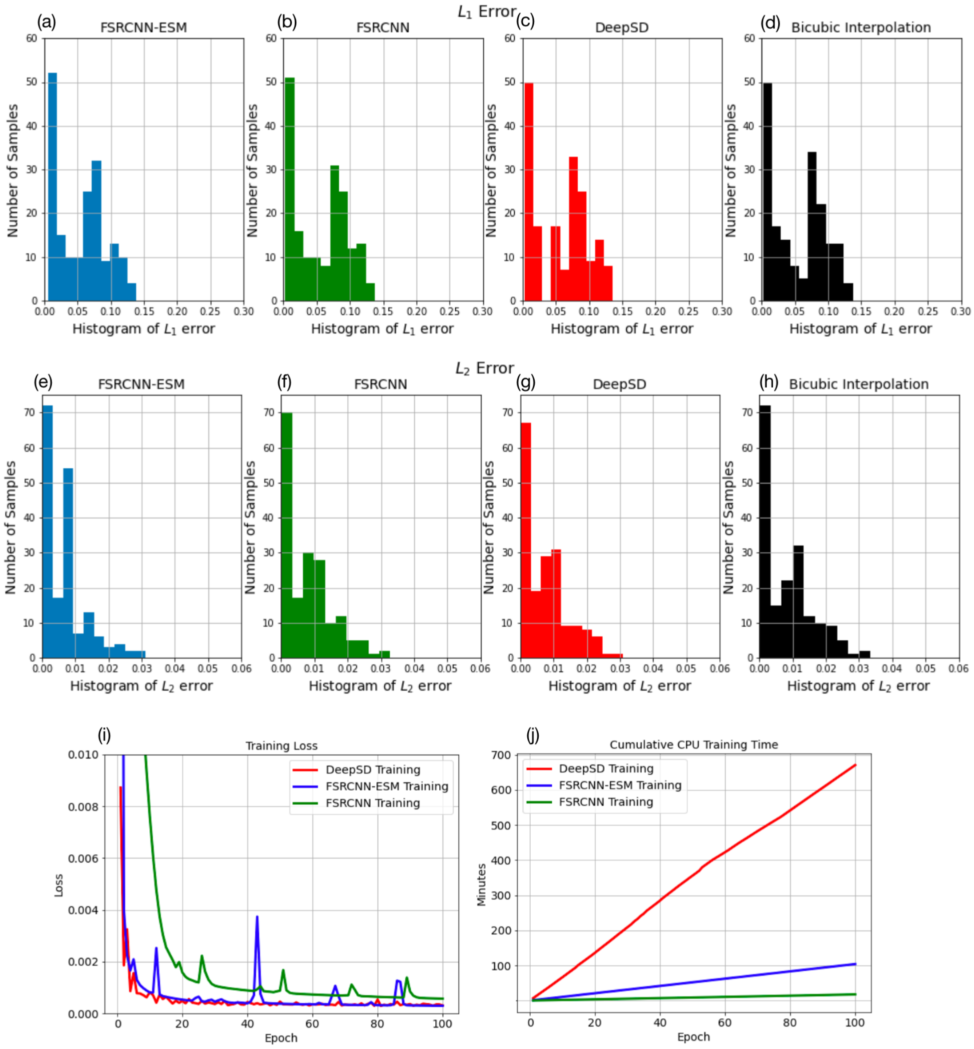


Figure 3. L_1 and L_2 error computed for all reconstruction methods on the held out testing dataset (a-h) and the cumulative training time in minutes (i) and the training loss (j) for the DeepSD, FSRCNN and FSRCNN-ESM over 100 epochs.