Statistical Downscaling of Rainfall Projections using Deep CNNs

Videsh Suman, UG

Guides: Prof. Subimal Ghosh (Civil Engg.) & Prof. Amit Sethi (Electrical Engg.) | IIT Bombay

Abstract

- GCMs are tools, designed to simulate time series of global climate variables for future. Though the spatial scales, on which these operate, are very coarse and not desirable for any hydrologic process of interest.
- Downscaling models [dynamical (physics based) and statistical (data driven)] are developed to address the limitations of GCMs by projecting HR climate variables from coarse scale GCM simulations.
- SD techniques are very fast in obtaining projections, although, they demand observed data availability over longer time scale.
- With the advent of parallel computing frameworks and huge data storage devices, there has been a meteoric boom in the application of deep neural networks.

Problem Definition

- The problem being attempted in this research involves using CNNs for obtaining future rainfall projections at HR (0.25°) over Indian landmass.
- The reanalysis predictors serve as the LR (2.5°) input, which later can be replaced with corresponding bias corrected GCM predictors.

Motivation

- The HR climate projections can be used to obtain future patterns of extreme events pertaining to different climate variables like temperature, rainfall.
- The rainfall projections can play crucial role for a country like India in formulating different strategies regarding water-food-energy nexus, disaster mitigation planning, etc.
- The recent advances of deep learning has helped in solving complex computational problems in fields like Computer Vision (CV), Natural Language Processing (NLP). The volumes of climate data available, serve as an apt motivation to apply similar networks for such impactful problems too.

Literature

Salvi et al.[2013]

- Model Pipeline: • quantile based remapping for bias correction • PCA for dimensionality reduction • division of the landmass in meteorological homogeneous zones • clustering to obtain daily rainfall states for each zone • decision trees for classifying each zone into a state • kernel regression to obtain the projected daily rainfall for every grid point.
- Results: • good match to observed data in terms of statistical properties • for future projections over 21st century, the results indicated spatial non-uniformity for changes in mean rainfall.

Vandal et al.[2017]

- Analogy: The climate projections (with multiple predictor channels for a region) have been considered analogous to images (with RGB colours). Using the analogy, statistical downscaling can be related to image super-resolution, where one aims to learn a low to high-resolution image mapping.
- DeepSD Model: Retrieving 1/8° projections from 1° projections for historical data through a framework of 3 stacked SRCNNs with auxiliary input of elevation topography at each upscaling.
- Results: State-of-the-art results when compared to BCSD and other regression based frameworks.

What is Convolution

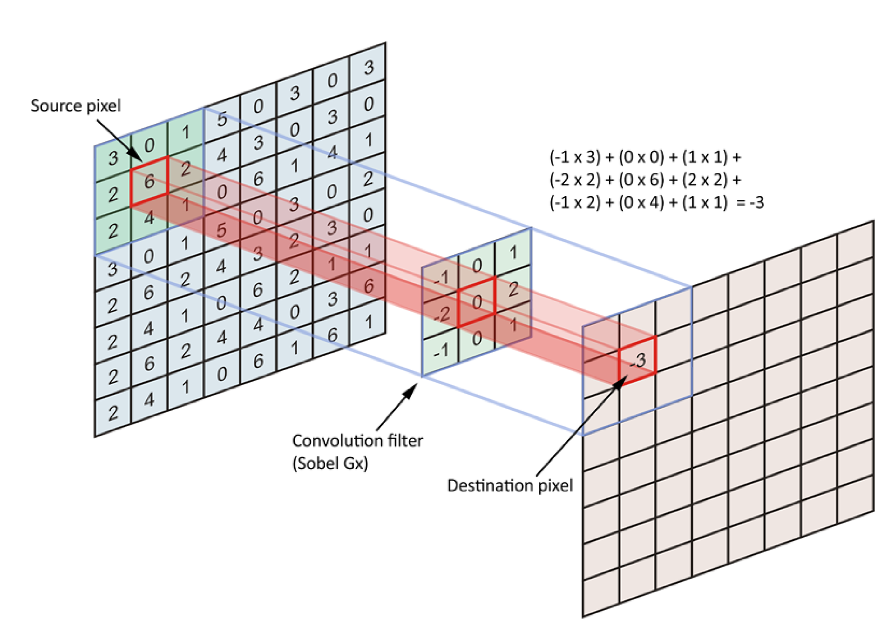

Figure 1: (Source: Internet) A kernel is applied across the source, and an element-wise product between each element of the kernel and the source is calculated at each location and summed to obtain the output value in the corresponding position of the destination, called a feature map.

Data

- Reanalysis is from NCEP/NCAR at 2.5° resolution which serves as the input to the model.
- Host GCM: CESM simulations at 1.25° resolution from 1850 to 2005 (past phase) and from 2006 to 2100 (future phase).
- Observed daily rainfall from APHRODITE is available at 0.25° resolution serve as the high-resolution ground truth for the model. This data covers all dates from 1951 to 2007.

Method

- One deep convolutional network that takes the input of a stack of 10 predictors (low-resolution) of 2-D grids with latitudes $5° - 40°$ and longitudes $65° - 100°$, returning the grid of observed rainfall in high-resolution (0.25°). This is essentially the single image super-resolution problem.
- Some features of this network: • transposed convolutions • dilated convolutions • Dropout regularization • Adam optimizer • optimization with warm learning rate restarts.

Results

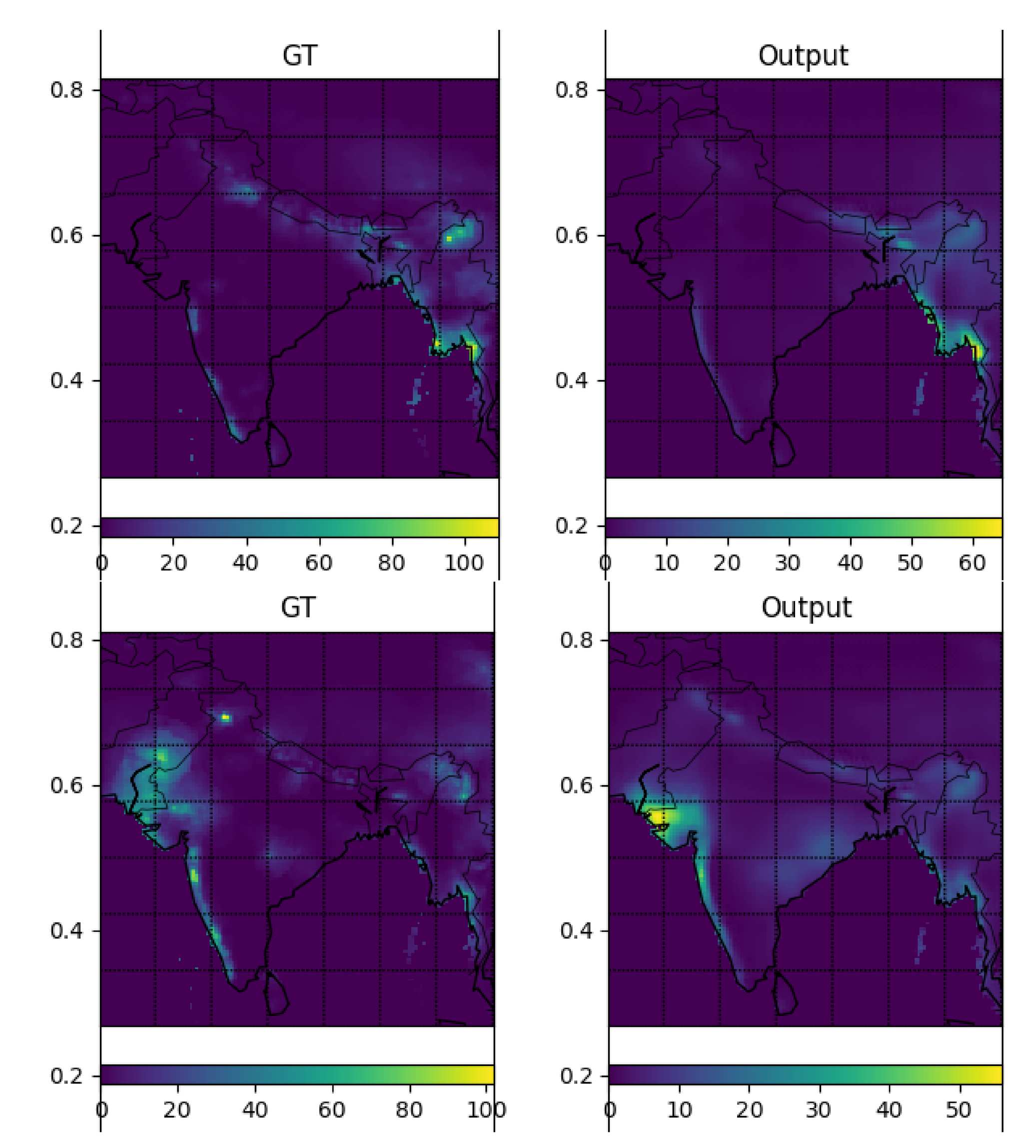

Figure 2: Randomly chosen predicted Output and GT projection pairs from testing results

Conclusion

- With only the reanalysis input data & no extra information, its very difficult to train a single model (however deep) that could be generalized on all locations of the observed area.
- The plots show that the model is unable to capture the variability in the range of values of rainfall across time and space throughout the landmass.
- The 10× super-resolution is a difficult problem, especially when the input and ground truth samples do not have a very good correlation.
- The network devised, consists of 1.2M trainable parameters which is quite large for an input of size 13870 training samples. This makes learning the super-resolution problem really difficult.

Future Work

- Write a shallower network with less parameters.
- Give the input in the form of zone-wise projections based on the meteorological homogenous zones presented in the study by Salvi et al.
- Use topographical features like elevation and land use/land cover from the NRSC project as relevant auxiliary information for the model.
- Obtain bias corrected GCM simulations for the required period and perform relevant experiments, comparing the overall downscaling quality with other baseline methods available.

References

- Thomas Vandal, Evan Kodra, Sangram Ganguly, Andrew Michaelis, Ramakrishna Nemani, Auroop Ganguly. **DeepSD**: Generating High Resolution Climate Change Projections through Single Image Super-Resolution. **ACM SIG-KDD 2017**.
- Kaustubh Salvi, Kannan S., Subimal Ghosh. High-resolution multisite daily rainfall projections in India with statistical downscaling for climate change impacts assessment. **Journal of Geophysical Research: Atmosphere, 2013**.