Regional analysis of climate projections using Bias Corrected spatial disaggregated super-resolution convolutional neural networks

Akash Awasthi

PhD Student, University of Houston

Abstract

Climate change is very crucial for ecological systems and society. But Global climate models run at coarse spatial resolution which is difficult to do regional analysis. Regional-scale projections can be obtained by a technique called statistical downscaling which uses past data to find out the high resolution and low-resolution mapping. There are many methods for statistical downscaling of climate data: 1) Conventional methods 2) Deep learning architecture. Some of the existing works like DeepSd downscaled High-resolution climate projections but in such cases, Global climate model (GCM) data suffers from concept drift, change of mapping between input and label over time. So applying these deep learning models is not a good idea for statistical downscaling**.** In my study, I have developed new approach of downscaling which outperforms other deep learning architectures like super-resolution convolutional neural network (SRCNN), Long short term memory network (LSTM) in terms of accuracy and reliability. These existing models focus on minimizing the root mean square error (RMSE) and do not take care of the tails or extremes. Therefore the objective function of these models should be changed other than root mean square error (RMSE). My proposed model focuses on both means and extremes. I provide a comparison between proposed and other existing deep learning models in downscaling daily precipitation and temperature from 1.25 to 0.25 resolution over India. I have downscaled 6 Global climate model (GCM) models in my comparative study.

Introduction

Climate change causes very dangerous effects on society which leads to extreme precipitation and temperature events. Natural resources are very much sensible to these extreme events which may cause drought and flood etc. Earth system models simulate climate change. These physics-based models can predict the atmospheric variables on a very large scale of about 125x125 KM grid [4]. But for regional analysis of these variables, I need to downscale the GCM data into the resolution of 25x25 [3]. Downscaling is basically of two types statistical and dynamic downscaling [1]. Dynamic downscaling is physics-based models that run on a regional scale and these are computationally expensive [2]. In contrast, Statistical downscaling finds the relationship between observed small scale variables and larger scale variables using artificial neural networks or support vector machines, linear regression. But these methods do not care about the spatial correlations and other existing deep learning models like SRCNN and LSTM do not perform well for statistical downscaling due to concept drift in GCM data. SRCNN is used in computer vision for signal image super resolution. It tries to minimize the mean difference (RMSE) between high resolution and low-resolution images. BCSD is a state of the art technique for Statistical downscaling which reproduces statistical distribution by doing quantile mapping between GCM data and observation from each individual grid point. My proposed work, BCSRCNN, a combination of BCSD and SRCNN to perform statistical downscaling doesn't limit itself to minimize the means but it also captures the extremes. I have downscaled the low-resolution climate projections into high-resolution climate projections over India for 6 global climate models (CESM-CAM5, NOR-ESM, MIROC, MPI, BNU-ESM, GFDL). In my proposed model first BCSD is applied and then the weights of auto encoded SRCNN for high-resolution output prediction. BCSD cares about the statistics of the data and SRCNN cares about the spatial correlations and distribution of errors. Data used to train and validate downscaling methods include observed precipitation data (high resolution) and GCM. I have used different GCMs (CESM-CAM5, NOR-ESM, MIROC, MPI, BNU-ESM, GFDL) from 1920-2005 as coarse resolution Input with resolution 1.25 and observation Data as High-Resolution Labels with Resolution 0.25 has been used for training.

Related Work:

From Ahmed et al. study, it has been noticed that statistical downscaling and dynamic downscaling perform equally with a negligible difference over a small region from GCM, which encourages us to opt for statistical downscaling over dynamical downscaling [6-8]. But Statistical downscaling approaches are developed based on the assumption that the statistical relationship between GCM and observation will remain the same in future predicted data [sachindra pap]. Generally, statistical downscaling approaches have been divided into three major categories: weather classification, weather generators and regression-based approaches [9-19]. In my study as i am focusing on regression-based approaches.

Many regression-based approaches have been widely used in statistical downscaling which includes Automated regression-based statistical downscaling (which classify the wet and non-wet days first and later apply regression techniques) [20-24], Linear regression and stepwise regression model (they will estimate predictand by using an optimized linear combination of predictors) [25-27], Support vector machines and Relevance Vector Machine (In the SVM and RVM algorithms, I use kernel functions to map non-linear problems into linear problems in high dimensional space) [28-35], Bayesian model averaging [36-42], LSTM [70,71], DeepSD [tj's pap] (which tries to captures spatial correlations by using convolution neural network and elevation as bias. But in his study, he has taken input as downscaled observation instead of a GCM output. Due to this, his model performs good, as it does not suffer from concept drift. I will talk more elaborately about concept drift in my later discussion). BCSD (it will try to do quantile mapping, which performs quite better in spite of its simplicity) [53-59].

From the past literature it has been found that irrespective of the machine learning and deep learning models which has been used, they perform well in simulating average (Means) and underestimates the tails and the standard deviation [60-62]. These downscaling models overfits the trend of lower percentiles and underfits the trend of higher percentiles [68]. But all-natural calamities related to climate are considered as extremes; which occur at higher percentiles. Even though in past studies, machine learning has been applied to statistical downscaling, those studies lack good evaluation of models that were developed. Because the majority of the studies used only RMSE as their metric; but mean will reside in lower percentiles and these models' overfit lower percentiles, RMSE is not a good enough metric to evaluate the models [63-67].

Methodology

Data Pre-processing

Data for a single day at the coarse resolution (GCM) of 1.25 ◦ is an "image" of size 25x27. Precipitation and elevation are used as input channels while precipitation is the sole output. Images are obtained at each resolution through downsampling using bicubic interpolation. For instance down-sampling 1.25^{o} to 0.25^{o} increases image size from 25x27 to 129x135 similar to the resolution of observed data. This interpolated image is given as input to all models. Data pre-processing is same across all the methods

Methods

- Super-Resolution Convolutional Neural Networks (SRCNN)
- Long-short term memory network (LSTM)
- Convolutional Long short term memory network (ConvLSTM)
- Auto Encoded Bias corrected Super-Resolution Convolutional Neural Networks(AE BCSRCNN)
- Bias Corrected Super-Resolution Convolutional Neural Networks (BCSRCNN)

Super-Resolution Convolutional Neural Networks (SRCNN)

CNN's are good at dealing with spatially related data. SR convolutional neural networks (SRCNNs) is a special type of Deep Neural Network. SRCNN is used to learn the functional mapping between LR images and HR images [5]. SRCNN involves three main operations:

Patch Extraction

Nonlinear mapping

Reconstruction

In my work, I have used three-layered SRCNN which takes two-channel images as input. One channel is the Low-resolution precipitation data for India and other is High-resolution Elevation. Layer one is formulated as follows

$$F(x) = max\,(W * X + B_0)$$

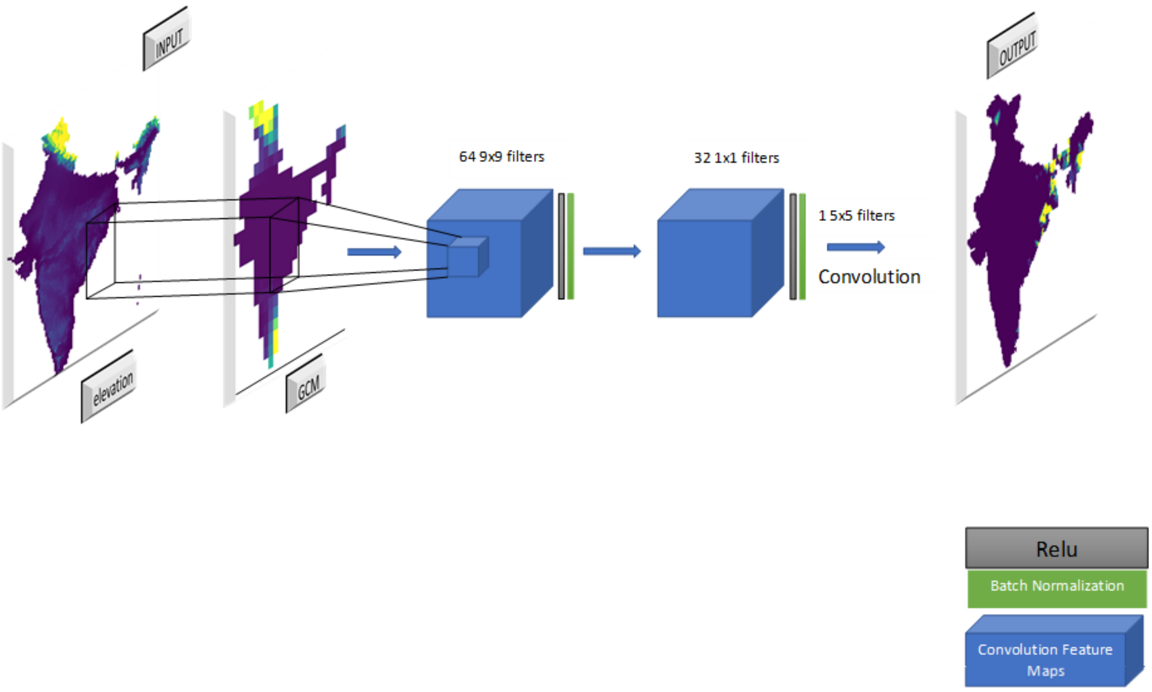

Fig 1 SRCNN Architecture

As shown in **fig 1** layer one involves convolution operation between kernel and input image followed by nonlinear mapping where W is filters and B is biased. W consists of 64 filters of size 9 x 9. Each filter of size n x n slides over the image and works as of the patch extraction layer. I have used relu activation function for nonlinear mapping. I have used padding with a replication method which preserves the size of an image similar to the input image after the convolution operation. Layer 2 is formulated as same as layer one but it takes input from layer one. Layer one output feature maps are fed as input to layer two which respectively performs convolution with 32 filters of size 1x1 and Relu operation. The output of layer two is fed as input to layer three. It performs convolution operation with 1 filter of size 5x5. End to end mapping involves learning of the parameters W and B of each layer. A Mean square loss function is used as an objective function which is defined as

$$l = argmin\ (\Theta) \sum_{i=1}^{n} \quad ||F\ (Xi\ ;\ \Theta)\ -\ Yi\ ||\ ^2$$

Long Short term memory network

LSTM is a very special type of recurrent neural network[70]. It is good at dealing with temporally related data. LSTM introduces a special so-called memory cell, which acts as an accumulator to learn long term dependency in a time-series. The cell is self-connected and copied its own real-valued state. Memory cell contains three gates input gate, output gate and forget gate. These gates indicate how much of the information should be passed to the next state and how much should be forgotten. Therefore LSTM preserves the long term dependency without vanishing gradient. The formulation of the LSTM cell is as follows:

$$x_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}\, c_{t-1} + b_i)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}\, c_{t-1} + b_f)$$

$$c_t = f_t\, c_{t-1} + i_1\ tanh(W_{hc}h_{t-1} + W_{hc}\, h_{t-1} + b_c)$$

$$o_t = \sigma(W_{x0}x_t + W_{ho}h_{t-1} + W_{co}\, c_t + b_o)$$

$$h_t = o_t tanh(c_t)$$

Here f is forget gate, I is input gate and o is output gate, c is cell memory, h is the previous state.

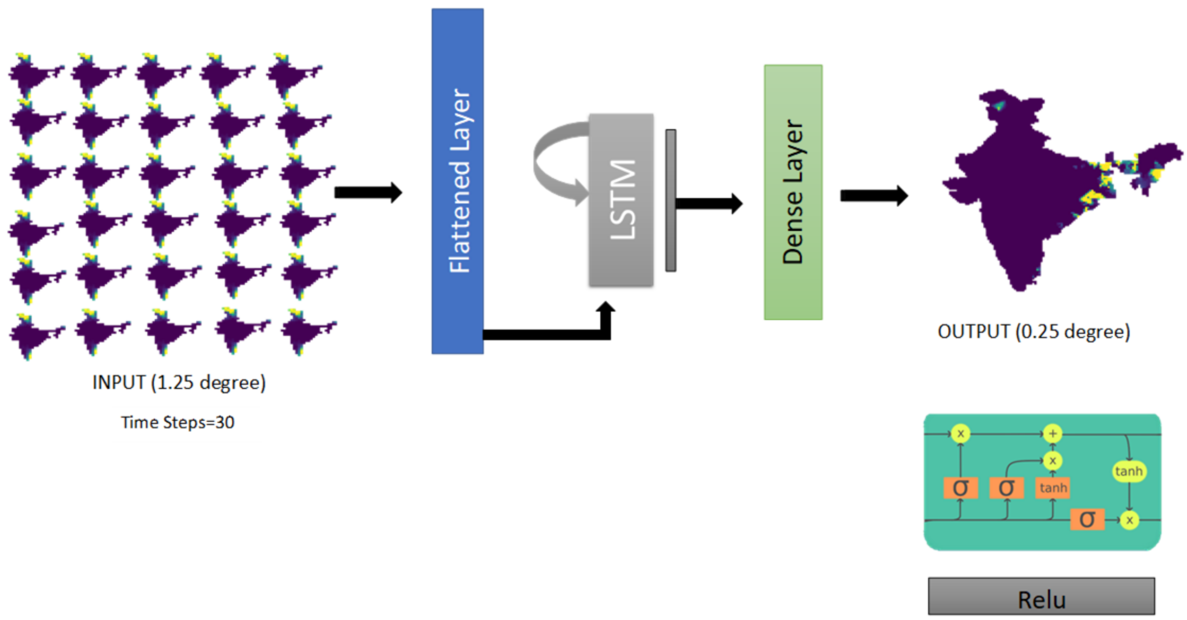

Fig 2 LSTM Architecture

As shown in fig 2, I have flattened the image and given as input to the LSTM. Initially cell memory (C) and hidden state (h) is initialized with 0. LR images from the past 30 days are used to predict the 30th-day high-resolution image. 30th day output of the LSTM is fed to dense layer which gives the output of dimension equal to label vector dimension

Convolution LSTM (ConvLSTM)

Long short term memory network is a special type of Recurrent network that preserves the temporal correlations and deals with long term dependencies. But in my case, I am giving a flattened image as input to LSTM which loses the special correlations. CNN preserves the special correlations. Therefore, the combination of LSTM and CNN deals with both special and temporal dependencies. ConvLSTM replaces the multiplication with convolution operation.

$$x_t = \sigma(W_{xi} * x_t + W_{hi} * h_{t-1} + W_{ci} * c_{t-1} + b_i)$$

$$f_t = \sigma(W_{xf} * x_t + W_{hf} * h_{t-1} + W_{cf} * c_{t-1} + b_f)$$

$$c_t = f_t c_{t-1} + i_1 \; tanh(W_{hc} * h_{t-1} + W_{hc} * h_{t-1} + b_c)$$

$$o_t = \sigma(W_{x0} * x_t + W_{ho} * h_{t-1} + W_{co} * c_t + b_o)$$

$$h_t = o_t tanh(c_t)$$

Here f is forget gate, I is input gate and o is output gate, c is cell memory, h is the previous state,* is convolution operation.

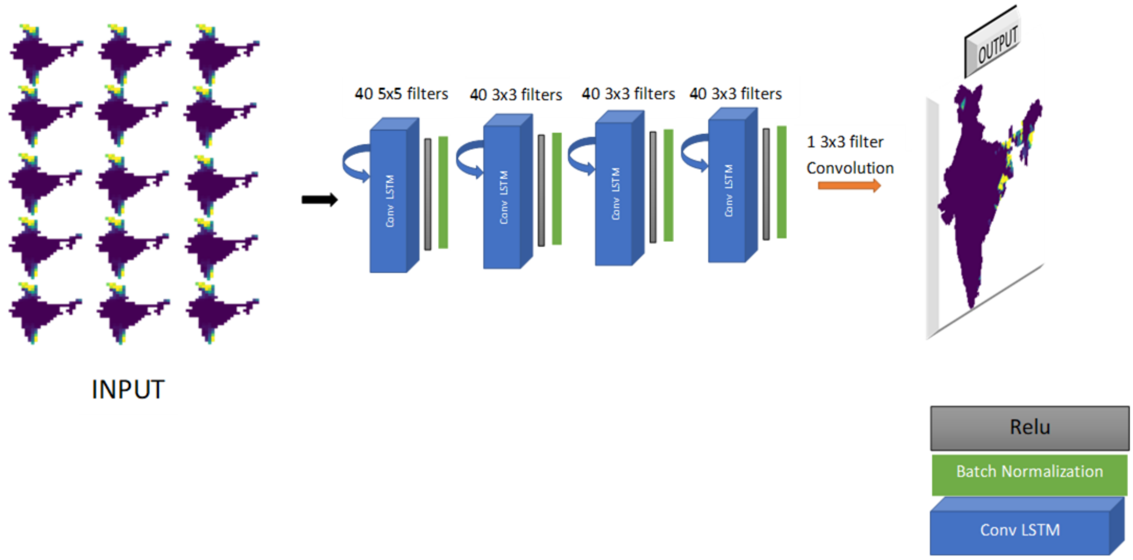

Fig 3 ConvLSTM Architecture

It will take input as a video (continuous frames of images). ConvLSTM2D will take input as 15 days of precipitation data and learns spatial and temporal relation and predict 15th day high-resolution image. I have used a generator to yield input. As shown in **fig 3** I have used 4 ConvLSTM2D layers and 1 conv2d layer, where layer 1 has 40 filters of size 5x5 followed by Relu activation and the rest of the layers have 40 filters of size 3x3 followed by Relu activation. The output from the last ConvLSTM2D is given as input to the conv2d which has 1 filter of size 3x3.

Concept drift

In my study, I have used GCM daily precipitation as input to the SRCNN. But raw GCM and observed data have no daily to daily correlation. Mapping between Raw GCM and observed data is changing with respect to time, this is called concept drift.

Auto encoded SRCNN

Data for a single day at the highest resolution, 0.25 ◦ , covering CONUS is an “image” of size 128X134. Images are obtained at each resolution through upsampling using a bicubic interpolation. For instance, up-sampling to 1.25 ◦ decreases the image size from 128x134 to 25x27. This interpolated image in an LR image (1.25) and given as input to the SRCNN model. Here, It is a mapping from Label to label (Y ->Y) which removes Concept drift

BCSRCNN:

Due to concept drift deep learning models are not able to find the mapping; as the relation between GCM and observed is changing with respect to time. So, I have first bias-corrected the GCM data. Bias correction will take care of a statistical relationship but it doesn't account for spatial relations. As CNN's will account for spatial relations, I have trained an autoencoder that takes input as extrapolated observed data (1.25^o) and maps to high resolution observed data (0.25^o) as discussed in Auto encoded SRCNN. Now, the weights which are obtained from Auto encoded SRCNN will account for spatial relation, so I have applied these weights to the bias-corrected data. For bias correction, I have used BCSD technique.

Comparison

I have used many methods for statistical downscaling like DeepSD, LSTM AEBCSRCNN, BCSRCNN, DeepSD is a stacked convolutional neural network that uses three successive CNNs to downscale from to. It does not perform well in terms of extremes but it gives good RMSE. Climate data is temporal-spatial data so we have used LSTM because LSTMs is good to handle the long term temporal dependencies but these also do not perform well in terms of extremes. AEBCSRCNN is a skated model of BCSD and AESRCNN. It performs well in terms of means and extremes. I have used the transfer learning and applied the trained weights of interpolated and raw observation on BCSD output.

BCSRNN is skated model of SRCNN and BCSD which uses the BCSD output as input to it and applies the trained weights on SRCNN to predict the high resolution GCM data. It performs well in terms of both means and extremes. The following table describes the validation RMSE for each model.

Model	RMSE (mm/day)
SRCNN	5.3
LSTM	5.03
convLSTM	5.7
AUBCSRCNN	2.5
BCSRCNN	1.1

Table 1 Comparison RMSE between BCSRCNN and all other models

Mean differences: I have calculated the mean difference between the observation and output of each model for 15 years of data. Mean difference between observations and each model output has been calculated over time dimension. As shown in fig 4 all Deep learning models have good performance in terms of means. These models try to minimize the mean error

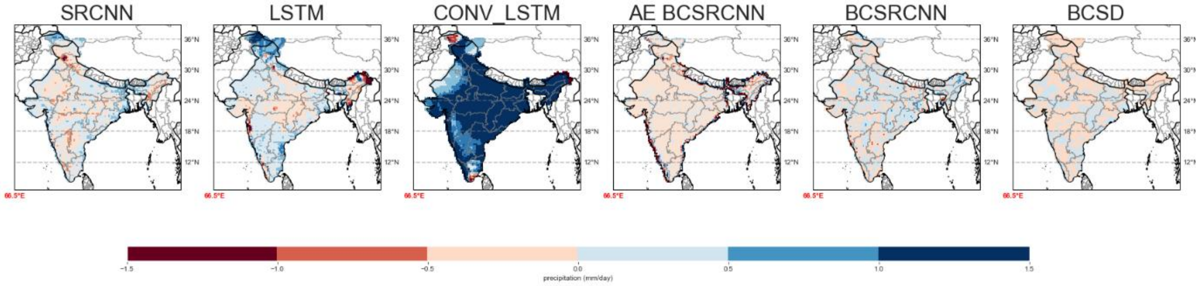

Fig 4 Mean Differences

Percentiles: I have calculated the 25,50,75,90,99 percentiles for observations and each model output and taken the difference between each percentile of observation and model's output. As shown in fig 5,6,7,8,9,10. All the models perform well enough for lower percentiles but SRCNN and LSTM don't perform well for higher percentile due to the presence of extreme values. Therefore these do not capture extremes. but BCSRCNN and AUBCSRCNN perform well even for higher percentiles.

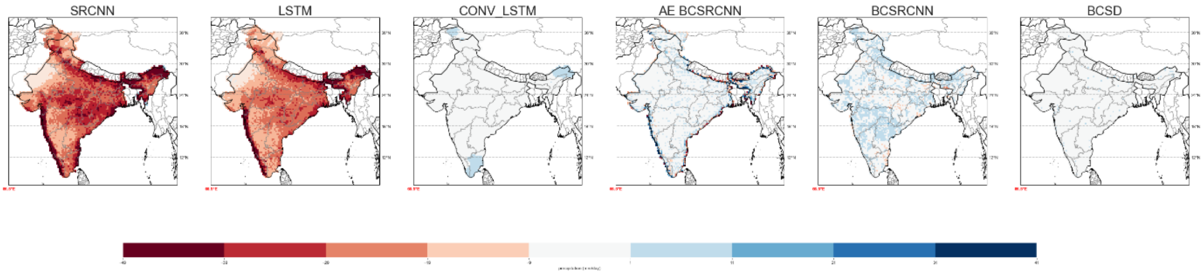

Fig 5 99th Percentile difference

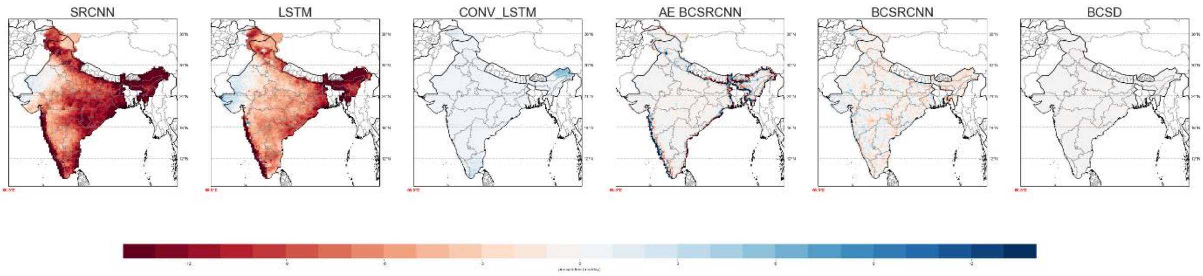

Fig 6 95th Percentile difference

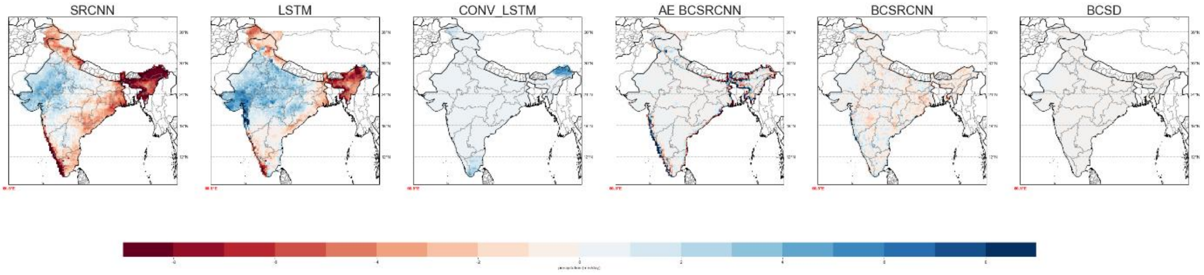

Fig 7   90 Percentile differences

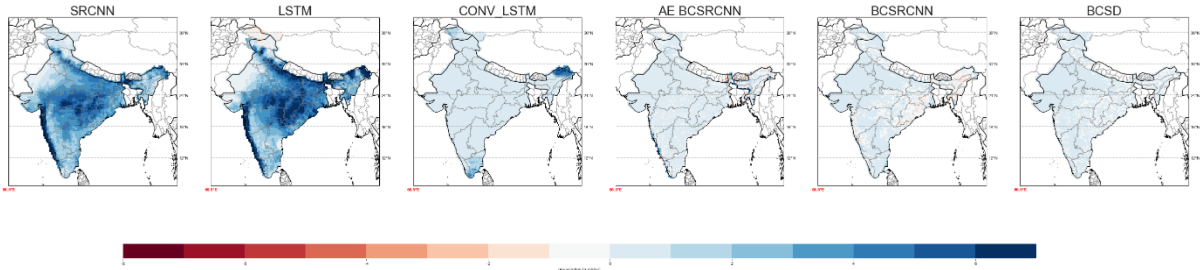

Fig 8   75 Percentile differences

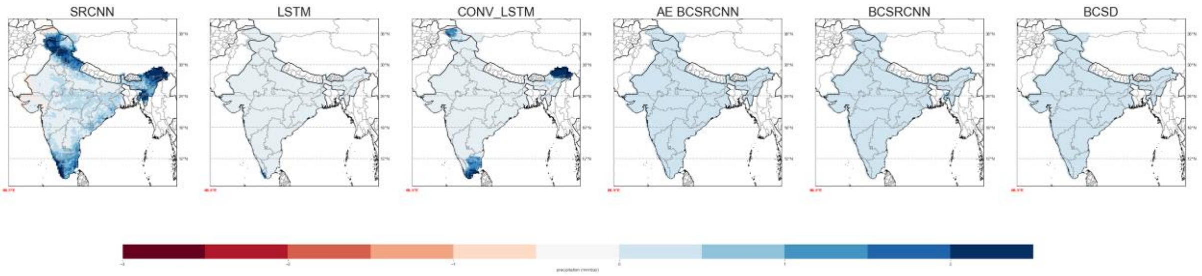

Fig 9   50 Percentile differences

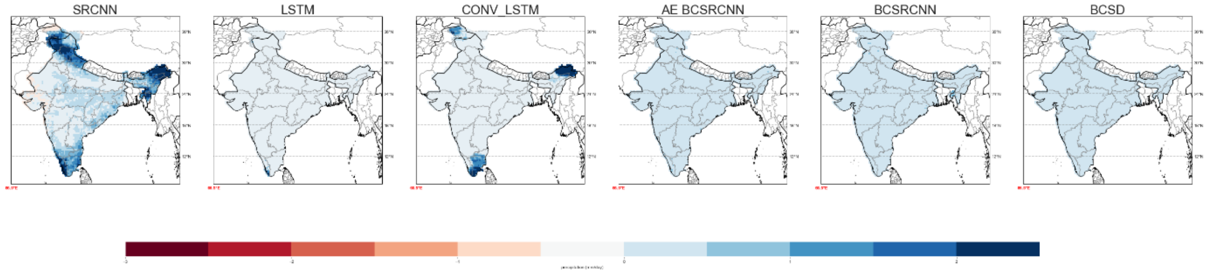

Fig 10   25 Percentile differences

*

Random day plot: I have randomly selected a day from each model's output and plotted along with Input (GCM) and label (observation). As shown in fig 11 , I can not do day to day mapping because each model gives a different outputs which is perceptually not similar to the label.

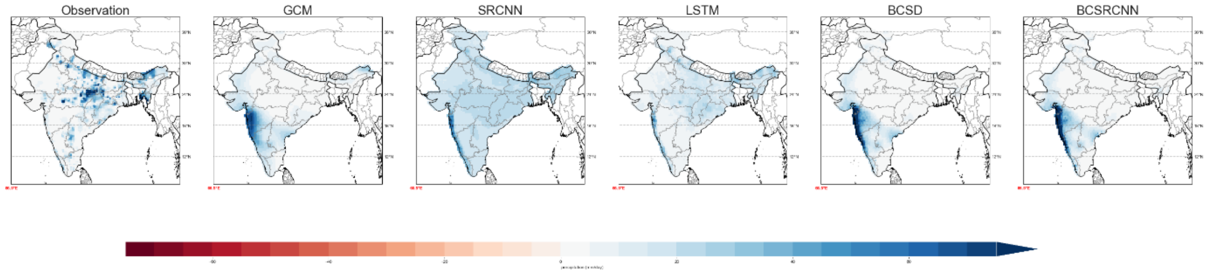

Fig 11 Random day Plot

References

1. Wilby, R. L., & Wigley, T. M. L. (1997). Downscaling general circulation model output: a review of methods and limitations. Progress in Physical Geography: Earth and Environment, 21(4), 530–548. https://doi.org/10.1177/030913339702100403

2. Wilby, R.L. and Fowler, H.J. (2010) Regional Climate Downscaling. In: Fung, C.F., Lopez, A. and New, M., Eds., Modelling the Impact of Climate Change on Water Resources, Wiley-Blackwell Publishing, Chichester, 34-85.

3. QuirinSchiermeier.2010. Therealholesinclimatescience. Nature News 463, 7279(2010),284–287.

4. Karl E Taylor, Ronald J Stouffer, and Gerald A Meehl. 2012. An overview of CMIP5 and the experiment design. Bulletin of the American Meteorological Society 93,4(2012),485

5. Dong, Chao, Chen Change Loy, Kaiming He, and Xiaoou Tang. "Image super-resolution using deep convolutional networks." IEEE transactions on pattern analysis and machine intelligence 38, no. 2 (2015): 295-307.

6. Okkan U, Fistikoglu O (2014) Evaluating climate change effects on runoff by statistical downscaling and hydrological model GR2M. Theor Appl Climatol 117:343-361.

7. Rashid MM, Beecham S, Chowdhury RK (2015) Statistical Downscaling of Rainfall: A Non-stationary and Multi-resolution Approach. Theor Appl Climatol 124:919–933.

8. Sachindra DA, Ng AWM, Muthukumaran S, Perera BJC (2016) Impact of climate change on urban heat island effect and extreme temperatures: a case-study. Q J Roy Meteor Soc 142:172–186

9. Hay LE, McCabe GJ, Wolock DM, Ayers MA (1991) Simulation of precipitation by weather type analysis. Water Resour Res 27(4):493–501

10. Bardossy A, Plate EJ (1992) Space-time model for daily rainfall using atmospheric circulation patterns. Water Resour Res 28(5):1247– 1259

11. Corte-Real J, Zhang X, Wang X (1995) Downscaling GCM information to regional scales: A non-parametric multivariate regression approach. Clim Dyn 11:413–424

12. Conway D, Jones PD (1998) The use of weather types and air flow indices for GCM downscaling. J Hydrol 212/213:348–361

13. Bardossy A, Duckstein L, Bogardi I (1995) Fuzzy rule-based classificationofatmosphericcirculationpatterns.IntJClimatol15:1087– 1097

14. Murphy JM (1999) An evaluation of statistical and dynamical techniques for downscaling local climate. J Clim 12(8):2256– 2284

15. von Storch H, Zorita E, Cubash U (1993) Downscaling of global climate change estimates to regional scales: an application to Iberian rainfall in wintertime. J Clim 6:1161–1171

16. Wilks DS (1999) Multisite downscaling of daily precipitation with a stochastic weather generator. Clim Res 11:125–136

17. Hughes JP, Guttorp P (1994) A class of stochastic models for relating synoptic atmospheric patterns to regional hydrologic phenomena. Water Resour Res 30(5):1535–1546

18. Hughes JP, Lettenmair DP, Guttorp P (1993) A stochastic approach for assessing the effect of changes in synoptic circulation patterns on gauge precipitation. Water Resour Res 29(10):3303–3315

19. RL Wilby, SP Charles, E Zorita, B Timbal, P Whetton, and LO Mearns. 2004. Guidelinesforuseofclimatescenariosdevelopedfromstatisticaldownscaling methods. (2004).

20. Mandal S, Srivastava R K, Simonovic S P (2016) Use of beta regression for statistical downscaling of precipitation in the Campbell River basin, British Columbia. J Hydrol 538:49–62

21. Kannan S, Ghosh S (2013) A nonparametric kernel regression model for downscaling multisite daily precipitation in the Mahanadi basin. Water Resour. Res. Res. 49:1360–1385

22. Schoof JT, Pryor S (2001) Downscaling temperature and precipitation: a comparison of regression-based methods and artificial neural networks. Int J Climatol 21(7):773–790

23. Masoud Hessami, Philippe Gachon, Taha BMJ Ouarda, and Andr´e St-Hilaire. 2008. Automated Regression-based statistical downscaling tool. Environmental Modelling & Software 23,6(2008),813–834.

24. Taylor, J. W., 2000: A quantile regression neural network approach to estimating the conditional density of multiperiod returns. J. Forecasting, 19, 299–311.

25. Kannan S, Ghosh S (2013) A nonparametric kernel regression model for downscaling multisite daily precipitation in the Mahanadi basin. Water Resour. Res. Res. 49:1360–1385

26. AlexJCannon.2011. Quantile regression neural networks: Implementation in R and application to precipitation downscaling. Computers & geosciences 37,9 (2011),1277–1284.

27. Haerter, J. O., Hagemann, S., Moseley, C., and Piani, C.: Climate model bias correction and the role of timescales, Hydrol. Earth Syst. Sci., 15, 1065–1079, https://doi.org/10.5194/hess-15-1065-2011, 2011.

28. Subimal Ghosh.2010.SVM PGSL coupled approach for statistical downscaling to predict rainfall from GCMoutput. Journal of Geophysical Research: Atmospheres 115,D22(2010).

29. Ghosh S, Mujumdar P (2008) Statistical downscaling of GCM simulations to streamflow using relevance vector machine. Adv Water Resour 31:132-146.

30. Anandhi A, Srinivas VV, Nanjundiah RS, Kumar DN (2008) Downscaling precipitation to river basin in India for IPCC SRES scenarios using support vector machine. Int J Climatol 28:401-420.

31. Anandhi A, Srinivas VV, Nanjundiah RS, Kumar DN (2009) Role of predictors in downscaling surface temperature to river basin in India for IPCC SRES scenarios using support vector machine. Int J Climatol 29:583–603

32. Sachindra DA, Huang F, Barton AF, Perera BJC (2013) Least square support vector and multi-linear regression for statistical downscaling general circulation model outputs to catchment streamflows. Int J Climatol 33:1087-1106

33. Goly A, Teegavarapu RSV, Mondal A (2014) Development and Evaluation of Statistical Downscaling Models for Monthly Precipitation. Earth Interact 18:1-28.

34. Okkan U, Serbes ZA, Samui P (2014) Relevance vector machines approach for long-term flow prediction. Neural Comput Appl 25:1393–1405. doi:10.1007/s00521-014-1626-9

35. Okkan U, Inan G (2014) Bayesian Learning and Relevance Vector Machines Approach for Downscaling of Monthly Precipitation. J Hydrol Eng 20:04014051

36. Vandal T, Kodra E, Dy J, Ganguly S, Nemani R, Ganguly AR (2018) Quantifying uncertainty in discrete-continuous and skewed data with Bayesian deep learning. In: Proceedings of the 24rd ACM SIGKDD international conference on knowledge discovery and data mining. ACM

37. Xianliang Zhang and Xiaodong Yan 2015. A new statistical precipitation downscaling method with Bayesian model averaging: a case study in China. Climate Dynamics, Vol. 45, 9--10 (2015), 2541--2555.

38. Xianliang Zhang and Xiaodong Yan 2015. A new statistical precipitation downscaling method with Bayesian model averaging: a case study in China. Climate Dynamics, Vol. 45, 9--10 (2015), 2541--2555.

39. AlonManorandSigalitBerkovic.2015. Bayesian Inference aided analog downscaling for near-surface winds in complex terrain. Atmospheric Research 164 (2015),27–36.

40. Joshi, D, St-Hilaire A, Ouarda, TBMJ, Daigle A (2015) Statistical downscaling of precipitation and temperature using sparse Bayesian learning, multiple linear regression and genetic programming frameworks. Can Water Resour J 40:392–408. doi:10.1080/07011784.2015.1089191

41. Okkan U, Inan G (2014) Bayesian Learning and Relevance Vector Machines Approach for Downscaling of Monthly Precipitation. J Hydrol Eng 20:04014051.

42. Tipping ME (2001) Sparse Bayesian learning and the relevance vector machine. J Mach Learn Res 1:211- 244.

43. Fowler HJ, Wilby RL (2010) Detecting changes in seasonal precipitation extremes using regional climate model projections: Implications for managing fluvial flood risk. Water Resour Res. 46:W03525.

44. Okkan U, Fistikoglu O (2014) Evaluating climate change effects on runoff by statistical downscaling and hydrological model GR2M. Theor Appl Climatol 117:343-361.

45. Rashid MM, Beecham S, Chowdhury RK (2015) Statistical Downscaling of Rainfall: A Non-stationary and Multi-resolution Approach. Theor Appl Climatol 124:919–933.

46. Sachindra DA, Ng AWM, Muthukumaran S, Perera BJC (2016) Impact of climate change on urban heat island effect and extreme temperatures: a case-study. Q J Roy Meteor Soc 142:172–186

47. Karl TR, Wang WC, Schlesinger ME, Knight RW, Portman D (1990) A method of relating general circulation model simulated climate to observed local climate. Part I: seasonal statistics. J Clim 3:1053–1079

48. YangC, ChandlerRE, IshamVS, WheaterHS (2005) Spatial-temporal rainfall simulation using generalized linear models. Water Resor Res 41:W11415. https://doi.org/10.1029/2004WR003739

49. Wilby RL, Hey LE, Leavesly GH (1999) A comparison of downscaled and raw GCM output: implications for climate change scenarios in the San Juan River Basin, Colorado. J Hydrol 225:67–91

50. Wilks DS (1999) Multisite downscaling of daily precipitation with a stochastic weather generator. Clim Res 11:125–136

51. Vu MT, Aribarg T, Supratid S, Raghavan SV, Liong SY (2016) Statistical downscaling rainfall using artificial neural network: significantly wetter Bangkok? Theor Appl Climatol 126: 453–467

52. Schoof JT, Pryor S (2001) Downscaling temperature and precipitation: a comparison of regression-based methods and artificial neural networks. Int J Climatol 21(7):773–790

53. Bridget Thrasher, Edwin P Maurer, CMc Kellar, and PBDuffy. 2012. Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. Hydrology and Earth System Sciences 16,9(2012),3309–3314.

54. Andrew W Wood, Lai R Leung, V Sridhar, and DP Lettenmaier. 2004. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. Climatic change 62, 1-3 (2004), 189–216.

55. Thrasher, B., Maurer, E. P., McKellar, C., and Duffy, P. B.: Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping, Hydrol. Earth Syst. Sci., 16, 3309–3314, https://doi.org/10.5194/hess-16-3309-2012, 2012.

56. Haerter, J. O., Hagemann, S., Moseley, C., and Piani, C.: Climate model bias correction and the role of timescales, Hydrol. Earth Syst. Sci., 15, 1065–1079, https://doi.org/10.5194/hess-15-1065-2011, 2011.

57. Dobler, A. and Ahrens, B.: Precipitation by a regional climate model and bias correction in Europe and South Asia, Meteorol. Z., 17, 499–509, https://doi.org/10.1127/0941-2948/2008/0306, 2008.

58. Piani, C., Haerter, J. O., and Coppola, E.: Statistical bias correction for daily precipitation in regional climate models over Europe, Theor. Appl. Climatol., 99, 187, https://doi.org/10.1007/S00704-009-0134-9, 2010a.

59. Piani, C., Weedon, G. P., Best, M., Gomes, S., Viterbo, P., Hagemann, S., and Haerter, J. O.: Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models., J. Hydrol., 395, 199–215, https://doi.org/10.1016/j.jhydrol.2010.10.024, 2010b.

60. G. Burger, T. Q. Murdock, a. T. Werner, S. R. Sobie, and a. J. Cannon. 2012. Downscaling extremes -an inter comparison of multiple statistical methods for present climate. Journal of Climate 25, 12 (June2012), 4366–4388. http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-11-00408.1

61. EdwinPMaurerandHugoGHidalgo.2008. Utility of daily vs. monthly large-scale climate data: an inter comparison of two statistical downscaling methods. (2008).

62. ThomasVandal,Evan Kodra, and Auroop R Ganguly. 2017. Inter comparison of Machine Learning Methods for Statistical Downscaling: The Case of Daily and Extreme Precipitation. arXiv preprint arXiv:1702.04018 (2017).

63. Muhammad Zia Hashmi, Asaad Y Shamseldin, and Bruce W Melville. 2011.Comparison of SDSM and LARS-WG for simulation and downscaling of extreme precipitation events in a watershed. Stochastic Environmental Research and Risk Assessment 25,4(2011),475–484.

64. Campozano L, Tenelanda D, Sanchez E, Samaniego E, Feyen J (2016) Comparison of statistical downscaling methods for monthly total precipitation: Case study for the Paute River basin in Southern Ecuador. Adv Meteorol doi:10.1155/2016/6526341

65. Raje D, Mujumdar PP (2011) A comparison of three methods for downscaling daily precipitation in the Punjab region. Hydrol Process 25:3575–3589. doi:10.1002/hyp.8083

66. Burger G, Murdock TQ, Werner AT, Sobie SR, Cannon AJ (2011) Downscaling extremes - an intercomparison of multiple statistical methods for present climate. J Clim 25:4366–4388

67. Haylock M R, Cawley G C, Harpham C, Wilby R L, Goodess C M (2006) Downscaling heavy precipitation over the UK: a comparison of dynamical and statistical methods and their future scenarios. Int J Climatol 26:1397–1415

68. D.A. Sachindra, K. Ahmed, M.M. Rashid, S. Shahid, B.J.C. Perera , Statistical downscaling of precipitation using machine learning techniques downscaling with machine learning techniques. Atmospheric Research(2017), doi:10.1016/j.atmosres.2018.05.022

69. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comp 9(8):1735–1780