Hybrid Deep Learning Approach in Predictability of Monthly Rainfall Using Endogenous property and Global Climatic Indices

Presented by

Mandala Venkata Surendra
17CE33003

Under Supervision of
prof . Rajib Maity

Department of Civil Engineering
IIT Kharagpur

Introduction

- Rainfall prediction
  - ➢ Rainfall prediction is clearly of great importance for India
  - ➢ The spatio-temporal distribution of precipitation is getting modified as an impact of changing climate
- One would like to make
  - ➢ long term prediction, i.e. predict rainfall a few weeks or months in advance
  - ➢ short term prediction, i.e. predict rainfall over different locations a few days in advance

Problem Statement

- Among various spatio-temporal scales, seasonal or monthly prediction of rainfall over a Homogenous Monsoon regions(HMRs) of India is one of the most important tasks
- How to capture its endogenous properties of hydrological time series and association of large-scale Indices to Indian summer monsoon rainfall with respect to their temporal evolution?
- Is the variation of rainfall solved using advanced algorithms better than the existing prediction methods?

Objective

- **Specific objective**
  - ➢ Capture the endogenous properties and association of different combinations of large scale climatic Indices with the HMRs monthly rainfall series.
  - ➢ Extract the hidden sequential information in the HMRs monthly rainfall series and of large-scale Indices.
  - ➢ Develop monthly rainfall prediction model each for different HMR with a goal for improved prediction performance.

- **General objective**
  - ➢ Implementing of rainfall prediction system using hybrid DL, a combination of one-dimensional Convolutional Neural Network (Conv1D) and Multi-Layer Perceptron (MLP) for monthly rainfall prediction of different monsoon regions of India.

Methodology

- The methodologies used in this study are :
  - ➢ Literature survey
  - ➢ Propose the system
  - ➢ Design and Implement
  - ➢ Testing

Contribution of this study

- To enhance the prediction mechanisms by developing different model according to monsoon region of India and their association with different indices.
- To help for selecting and implementing of appropriate prediction algorithms for spatio-temporal distribution of precipitation in India.

Literature Review

- Artificial Neural Network based forecasting of consecutive rainfalls.
  - ➢ In 2018, Lee et al.:“Application of artificial neural networks to rainfall forecasting in the geum river basin, korea”
- DL based deep network algorithm for forecasting next day precipitation using environmental factors.
  - ➢ In June 2017,Zhang et al.: “A deep-learning based precipitation forecasting approach using multiple environmental factors”

Data Description

- The dataset consists of the monthly rainfall for the period 1901-2015 for each subdivision in India and also contains data of 5 different index values for period 1901-2015
- Now these different contiguous subdivisions are grouped into five Homogeneous monsoon regions based on rainfall characteristics and association with global/ regional circulation parameters
- It is done in order to assess the region-wise variation in association with the large-scale climatic indices and to improve the prediction performance at regional scale

Data description......

- The dataset contains 19 attributes (individual months, annual, and combinations of three consecutive months)
- Five HMRs are:
    a) Central North East(CNE) region
    b) Peninsular(PE) region
    c) Northwest(NW) region
    d) West Central(WC) region
    e) Northeast(NE) region
- Different large-scale indices used are: EMI index, PDO index, NAO index, ENSO(NINO3.4) index and EQUINOO index.

Data description......

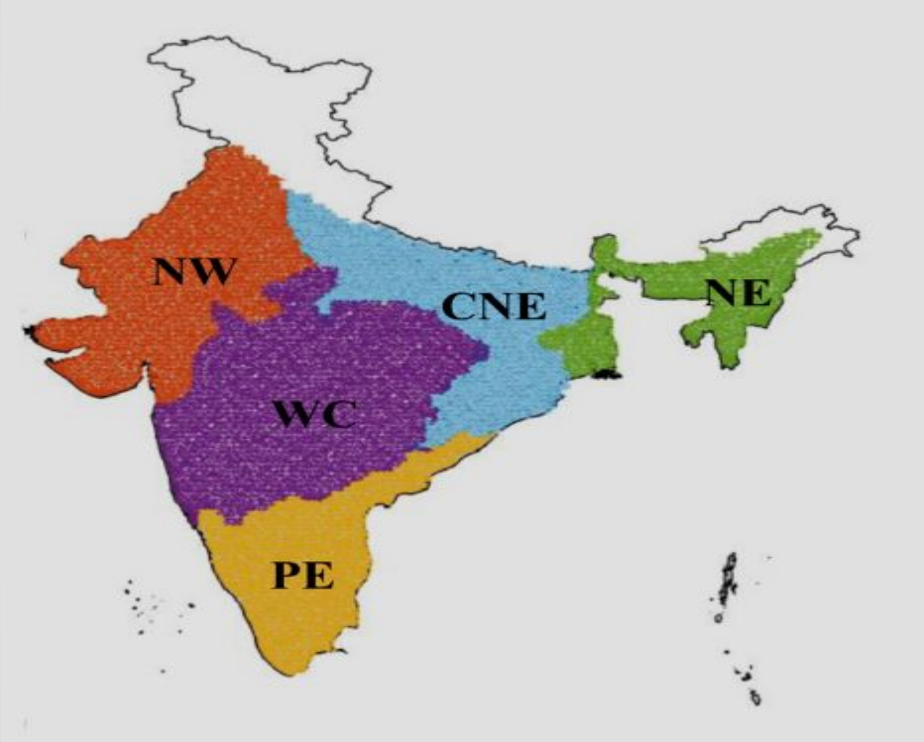


FIGURE 1: Map showing HMRs in India. These regions are the grouped contiguous sub-divisions based on the rainfall characteristics and association with global/regional circulation parameters to form the HMRs

Data description……

- The data from 1951–2000 were used to train the models, and the data from 2001–2015 were used for testing
- Previous months rainfall data and different combination Index values are used as input to predict the consecutive month rainfall
- Two different lags, i.e., 3 and 6 months are considered
- Lag indicates the gap (number of months) between input i.e. past data and the starting month of prediction.

Data preprocessing

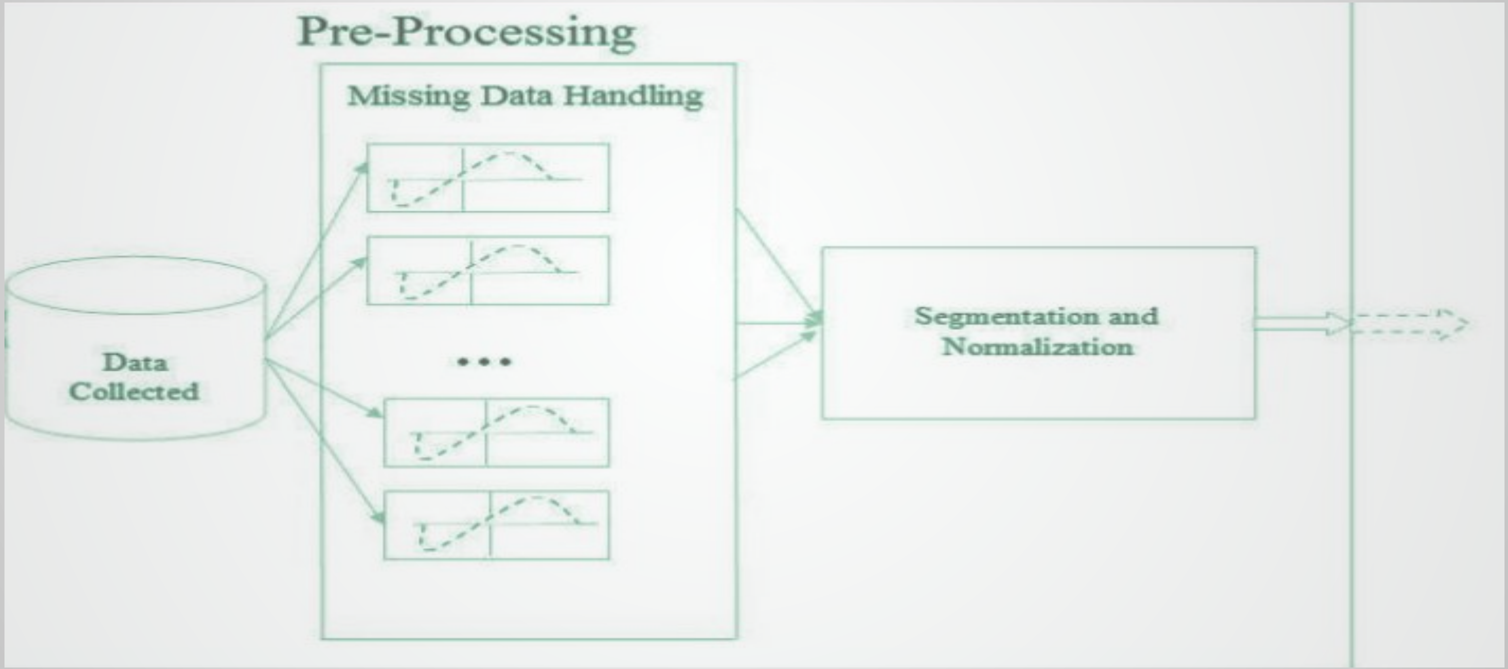


Figure 2: Block diagram of pre processing phase

preprocessing….

❑ Correlation among the variables

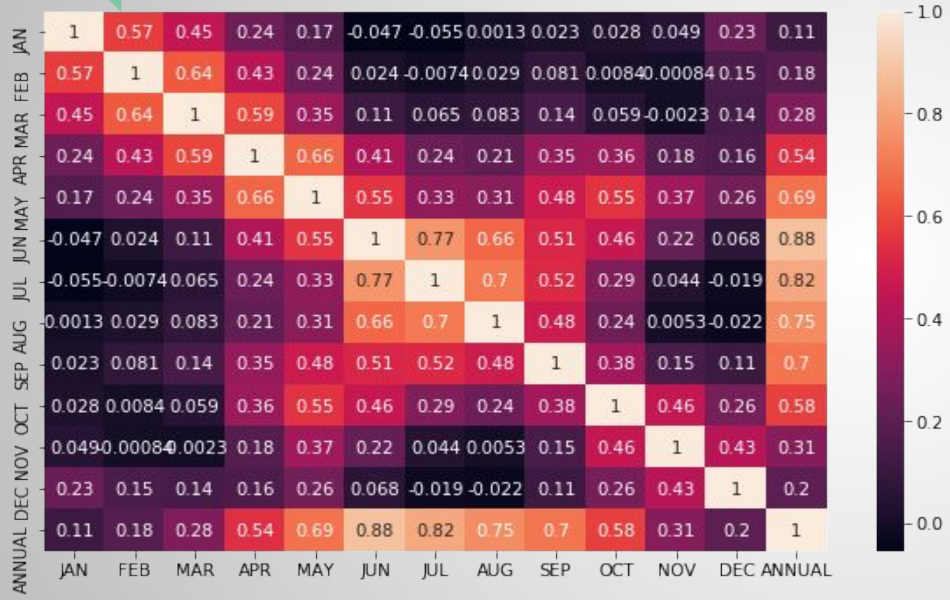


Figure 3. Heat map showing the correlation coefficients between rainfall in different months and the same with annual rainfall

Figure 4. Heat map showing the correlation coefficients between rainfall in different seasons and the same with annual rainfall

Model Architecture

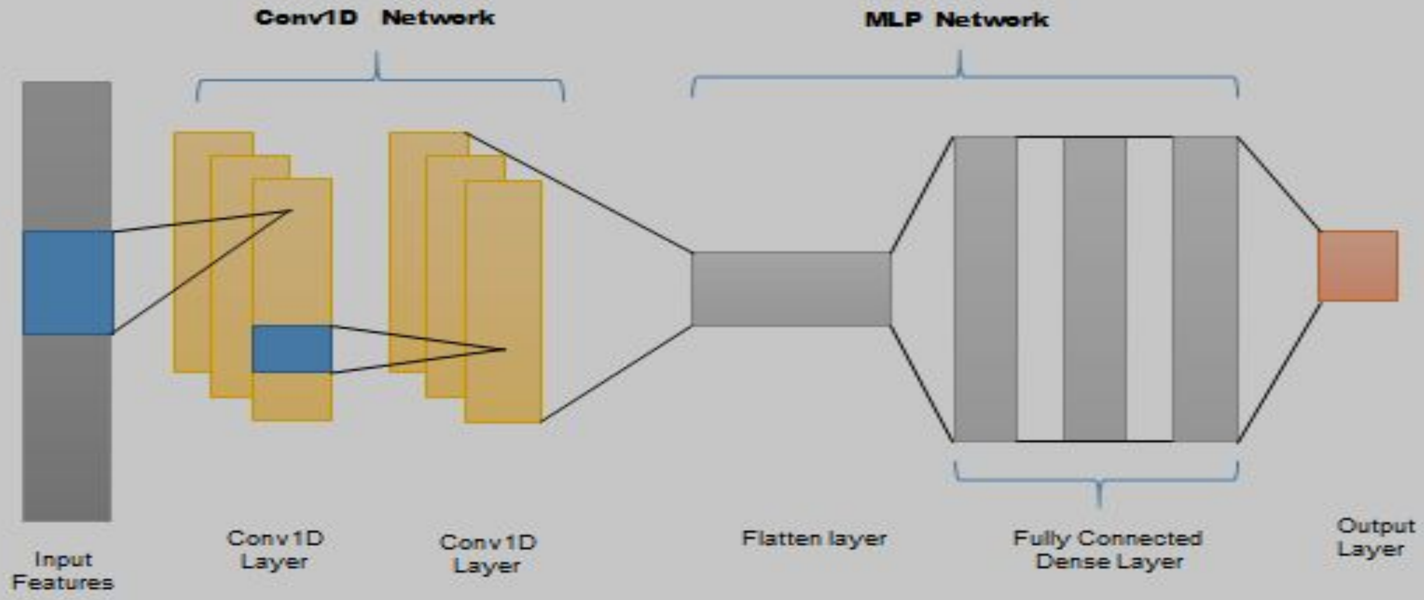


Figure 5: A schematic representation of the proposed hybrid Conv1D-MLP architecture.

Algorithm : CNN Training Procedure

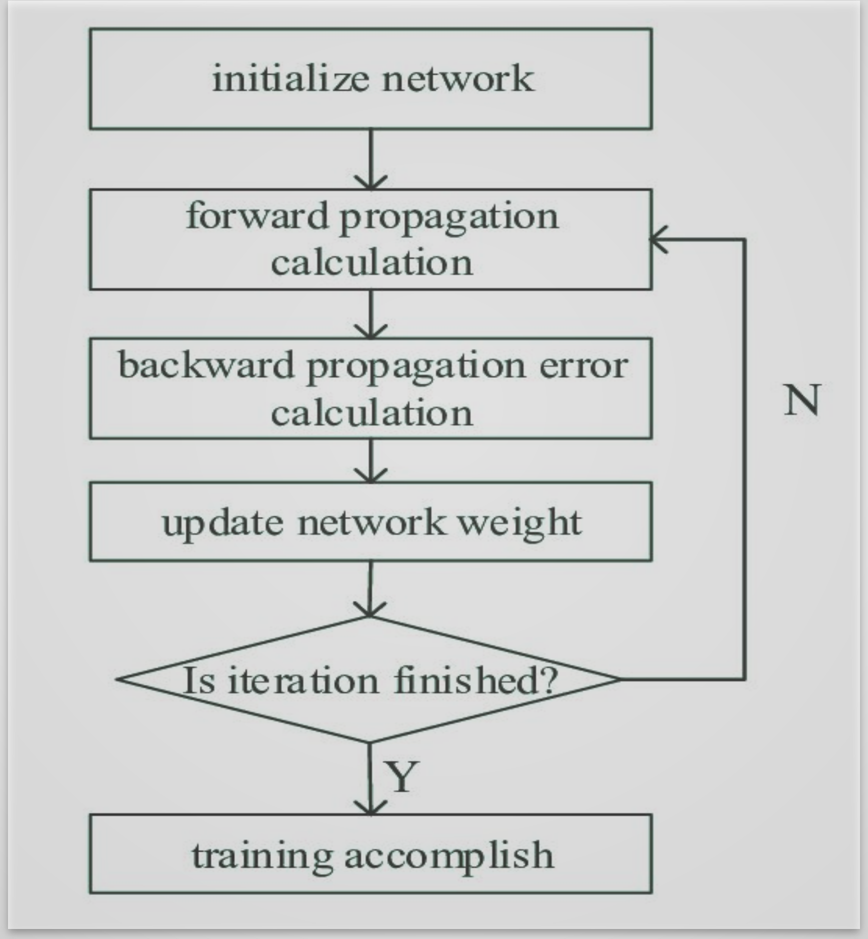
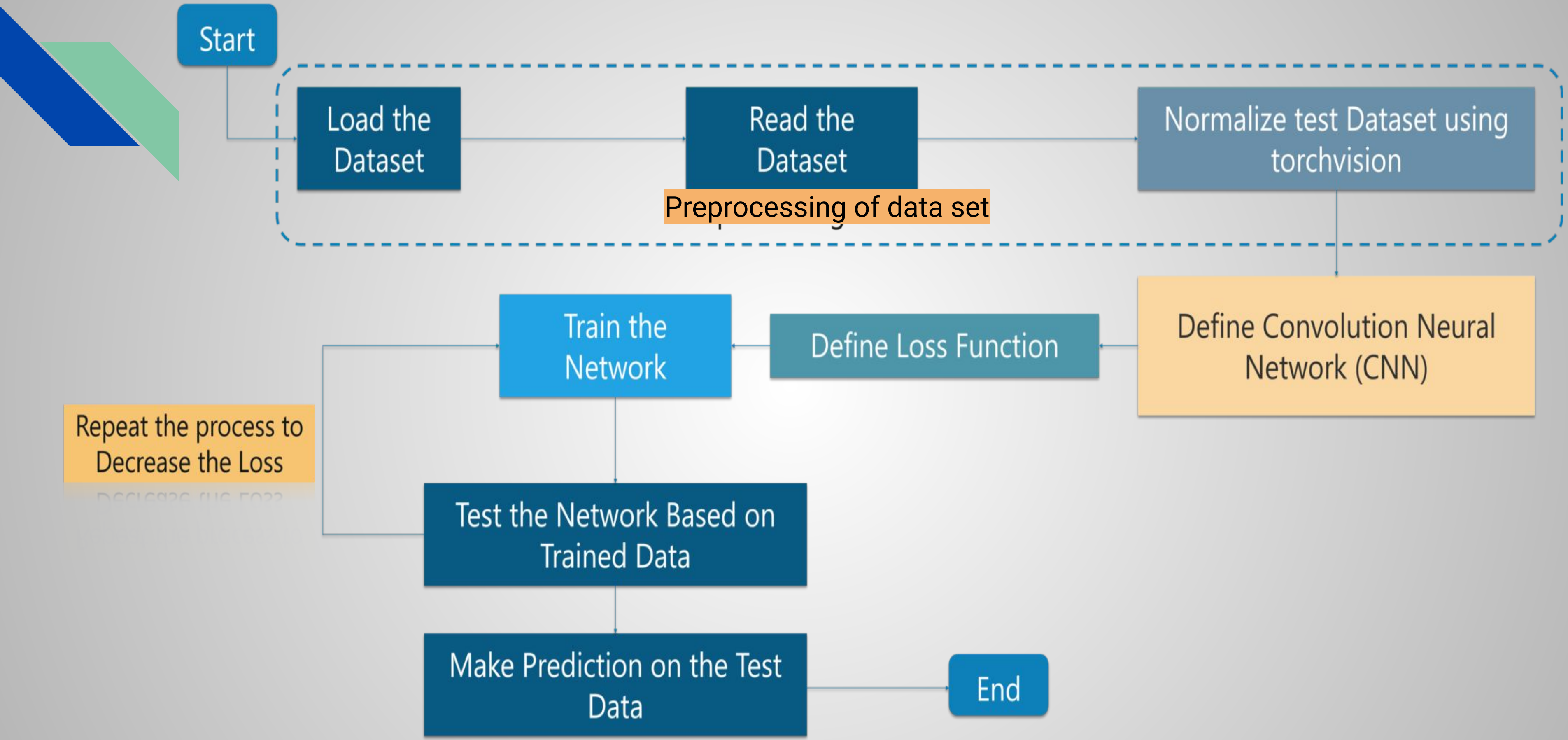


Table 1: Configurations of proposed hybrid Conv1D-MLP model

Layer no.	Layer	Type	Parameters of layers			
			Activation func.	Kernel Size	No.of filters	Neurons
1	Conv1D	Convolution Layer	ReLU	1	64	-
2	Conv1D	Convolution Layer	ReLU	2	128	-
3	Flatten	Flatten Layer	-	-	-	-
4	dense	Fully connected Layer	ReLU	-	-	128
5	dense	Fully connected Layer	ReLU	-	-	64
6	dense	Fully connected Layer	ReLU	-	-	32
7	dense	Fully connected Layer (output layer)	Linear	-	-	1

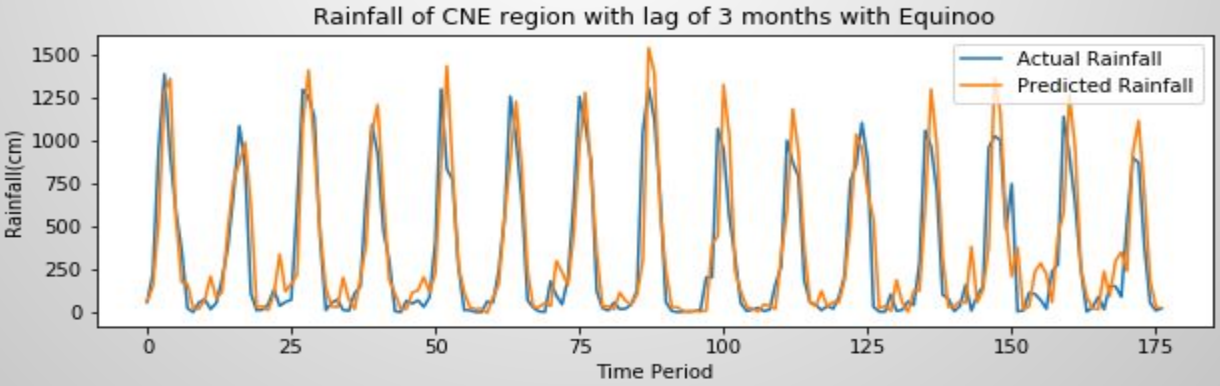

Start
Load the Dataset
Read the Dataset
Normalize test Dataset using torchvision
Preprocessing of data set
Define Convolution Neural Network (CNN)
Define Loss Function
Train the Network
Repeat the process to Decrease the Loss
Test the Network Based on Trained Data
Make Prediction on the Test Data
End

Results and Discussion

- Different models are trained based on the different HMRs with their respective data and observed the predicted rainfall for all the months from 2001-2015.

- Below are the some of the results for 3 months and 6 months lag period with different combination of climatic Indices.

- **3 months Lag**



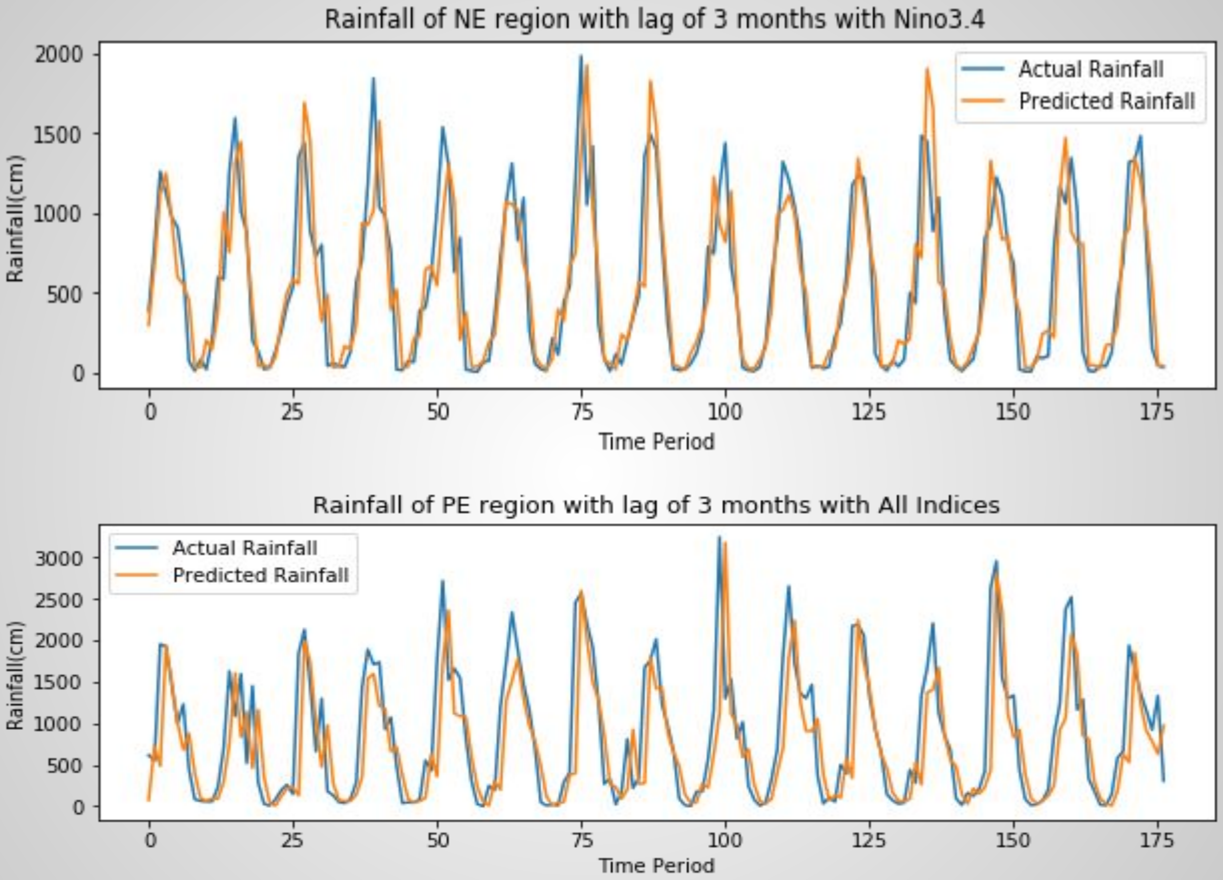


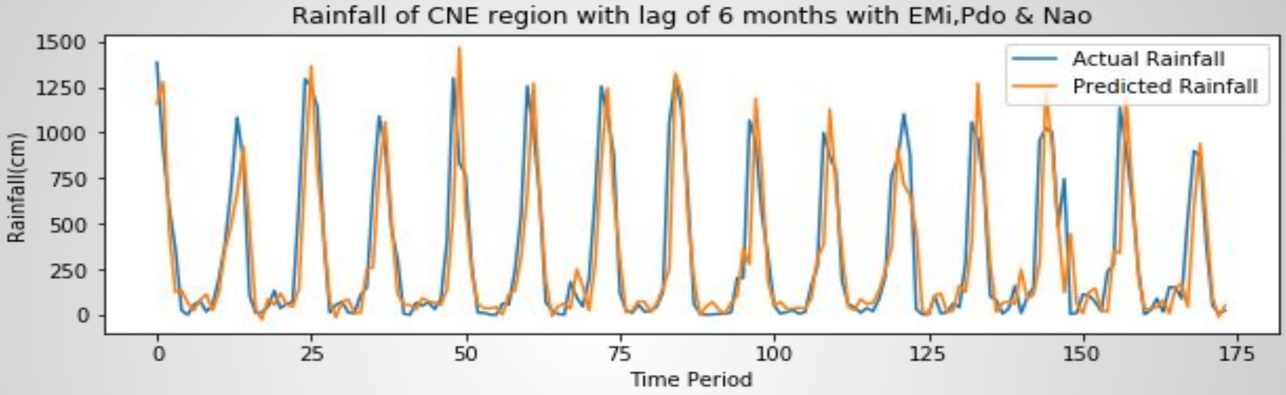
Figure 5. Comparison plots between actual and predicted monthly using inputs from 3 previous months with different combination of Indices for different HMRs . X-axis shows the months from the year 2001 to 2015.

TABLE 2: Performance statistics viz. r, RMSE and NSE for different combinations of Indices for different HMRs considering 3 months lag period.

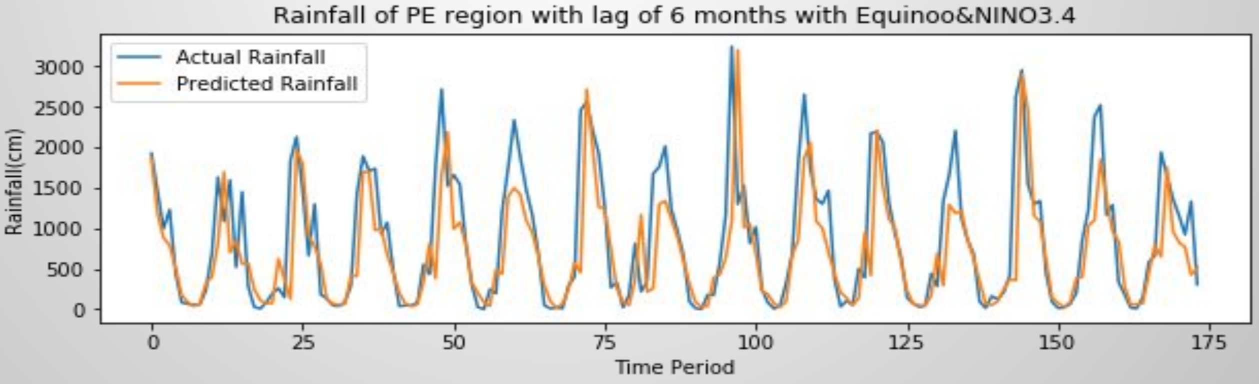
HMR	Index Combination	Performance Statistics: RMSE	Correlation coefficient(r)	NSE coefficient
CNE region	No Index	249.4	0.807	0.539
	EMI	246.19	0.813	0.56
	PDO	235.33	0.825	0.604
	NAO	223.90	0.838	0.62
	NINO3.4	211.26	0.857	0.679
	EQUINOO	208.83	0.872	0.744
	All Indices	244.03	0.812	0.528
PE region	No Index	588.03	0.707	0.250
	EMI	593.46	0.709	0.202
	PDO	581.08	0.713	0.294
	NAO	571.34	0.721	0.264
	NINO3.4	553.25	0.733	0.245
	EQUINOO	541.85	0.740	0.235
	All Indices	583.32	0.710	0.223

	No Index	322.66	0.692	0.05
	EMI	336.83	0.671	-0.167
	PDO	321.84	0.692	-0.018
NW region	NAO	309.86	0.710	0.068
	NINO3.4	300.96	0.728	0.135
	EQUINOO	289.61	0.751	0.095
	All Indices	328.57	0.682	-0.043
	No Index	724.77	0.744	0.392
	EMI	717.02	0.755	0.431
	PDO	703.04	0.771	0.536
WC region	NAO	677.34	0.778	0.528
	NINO3.4	670.85	0.781	0.534
	EQUINOO	663.64	0.785	0.545
	All Indices	717.27	0.750	0.377
	No Index	320.4	0.805	0.521
	EMI	310.03	0.812	0.585
	PDO	302.61	0.818	0.603
NE region	NAO	296.05	0.823	0.582
	NINO3.4	286.04	0.832	0.632
	EQUINOO	282.47	0.836	0.613
	All Indices	309.32	0.811	0.588

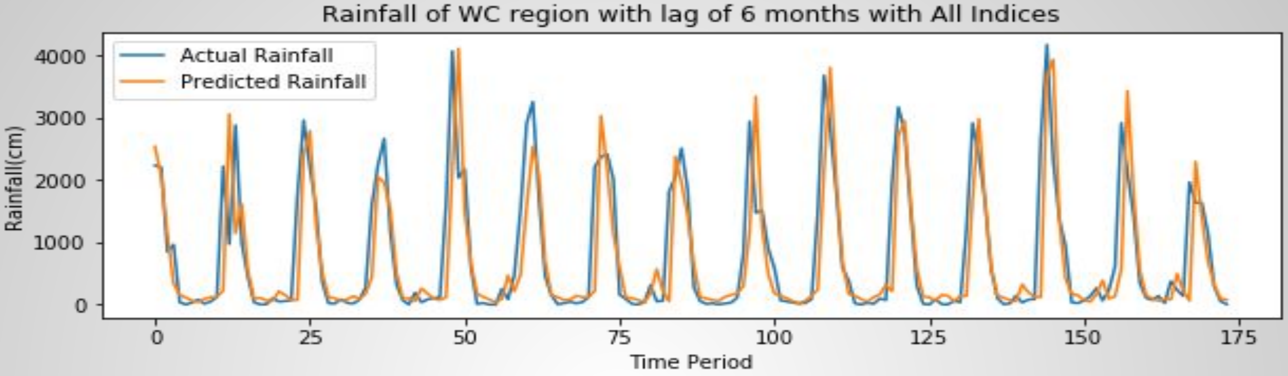
❑ 6 months lag



Rainfall of CNE region with lag of 6 months with EMi,Pdo & Nao
Actual Rainfall
Predicted Rainfall
Rainfall(cm)
1500
1250
1000
750
500
250
0
0
25
50
75
100
125
150
175
Time Period



Rainfall of PE region with lag of 6 months with Equinoo&NINO3.4
Actual Rainfall
Predicted Rainfall
Rainfall(cm)
3000
2500
2000
1500
1000
500
0
0
25
50
75
100
125
150
175
Time Period



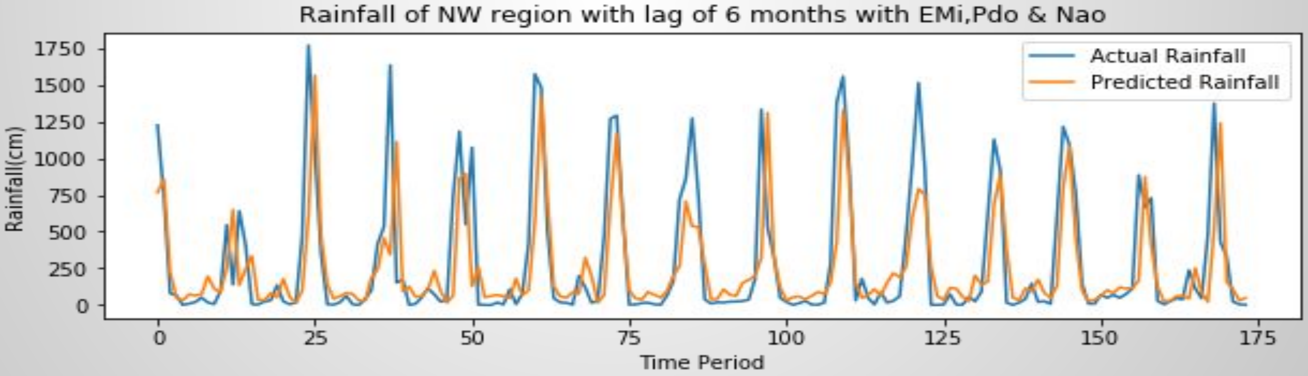


Figure 5. Comparison plots between actual and predicted monthly using inputs from 6 previous months with different combination of Indices for different HMRs . X-axis shows the months from the year 2001 to 2015.

TABLE 3: Performance statistics viz. r, RMSE and NSE for different combinations of Indices for different HMRs considering 6 months lag period.

HMR	Index Combination	Performance Statistics		
		RMSE	Correlation coefficient(r)	NSE coefficient
CNE region	No Index	234.44	0.837	0.575
	(EMI+PDO +NAO)	229.56	0.835	0.609
	(NINO3.4+ EQINOO)	235.97	0.827	0.576
	All Indices	237.94	0.820	0.546
PE region	No Index	551.03	0.750	0.315
	(EMI+PDO +NAO)	552.51	0.741	0.338
	(NINO3.4+ EQINOO)	555.58	0.746	0.228
	All Indices	549.28	0.753	0.336

NW region	No Index	329.35	0.687	0.012
	(EMI+PDO +NAO)	312.67	0.712	0.026
	(NINO3.4+ EQINOO)	328.37	0.688	0.061
	All Indices	317.57	0.704	0.065
WC region	No Index	695.89	0.776	0.507
	(EMI+PDO +NAO)	703.67	0.767	0.496
	(NINO3.4+ EQINOO)	703.47	0.769	0.468
	All Indices	690.82	0.771	0.493
NE region	No Index	275.28	0.852	0.658
	(EMI+PDO +NAO)	273.97	0.851	0.666
	(NINO3.4+ EQINOO)	275.67	0.855	0.651
	All Indices	270.73	0.854	0.685

- Increase in the performance with the increase in no. of inputs i.e better results with 6 month lag period.
- Central North East(CNE) and North East(NE) regions of India seems to produce better results compared to other regions with the proposed Model Architecture.
- Remaining regions seems to produce slightly bad results compared to CNE and NE regions in terms of RMSE and NSE coefficients.
- Their performance can be improved by slightly changing the different parameters of the architecture.
- Also considering the individual Index, EQUINOO produced better results in most of the cases i.e it has close association with the summer monsoon rainfall of India.
- Similarly, Combination of EMI,NAO and PDO indices also produced better results in terms of predictability.

Conclusion

- Hybrid Conv1D-MLP model has potential for monthly rainfall prediction of daily rainfall using past data and climatic indices as the input variables
- From performance it is concluded that deep learning has the potential to capture the non-linear relationship between the past data,association of large scale Indices with monthly rainfall variability.
- So the predictions obtained from the proposed hybrid DL model can be helpful in agriculture, flooding due to heavy rainfall etc..

THANK YOU!