Statistical Downscaling of Rainfall Under Climate Change in Krishna River Sub-basin of Andhra Pradesh, India Using Artificial Neural Network (ANN)

K.V.R. Satya Sai*, S. Krishnaiah** and A. Manjunath***

*Jawaharlal Nehru Technological University Anantapur, Anantapuramu, Andhra Pradesh, India
**Department of Civil Engineering, Jawaharlal Nehru Technological University Anantapur, Anantapuramu, Andhra Pradesh, India
***Civil Engineering Department, MVSR Engineering College, Nadergul, Hyderabad, India
†Corresponding author: Dr. A. Manjunath, manjunath.aluru@gmail.com

ABSTRACT

Due to the very coarse spatial resolution of the different global circulation model (GCM), we cannot use them in their natural form to study the various impacts of climate change. For matching this spatial inequality between the GCMs output (predictor) and historical precipitation data (predictands), we need to establish a relation between them which is known as downscaling. In the present study, we tried to examine the efficiency of the Artificial Neural Network (ANN) with Principal Component Analysis (PCA) for downscaling the rainfall for 3 districts of Andhra Pradesh of India. Firstly, for all the regions, the downscaling was performed by using ANN. Then seasonal and annual analysis was performed based on the R^2 and RMSE. The results show that the ANN worked adequately based on the statistical parameters. The study uses the Canadian Earth System Model (CanESM2) of the IPCC Fifth Assessment Report, re-analysis from the National Centre for Environmental Prediction (NCEP) as GCM model, and observed rainfall data as the observed rainfall. The analysis was performed for the three RCPs scenario, RCP 2.6, 4.5 and 8.5. Finally, the ANN model is applied to downscale the precipitation.

INTRODUCTION

Krishna River is one of the important rivers in AP and Telangana. Krishna basin is the second-largest eastward draining interstate river in peninsular India. It starts its journey near Jor village (near Mahabaleswar) in the Western Ghats at an elevation of about 1337 m above MSL and ends its journey into the Bay of Bengal near Hamsala Deevi near Vijayawada. The Krishna basin covers an area of nearly 8% of the total geographical area of the country. The Krishna river basin is located in the south part of Andhra Pradesh and Telangana. The basin spreads over parts of administrative limits of Mahabubnagar, Kurnool, Krishna, Nalgonda, Guntur, and Prakasam districts. It flows for a distance of 305 km in Maharashtra, 483 km in Karnataka and 612 km in Andhra Pradesh and Telangana combined before finally out falling into the Bay of Bengal.

The present study area (Fig. 1) is from Nagarjuna Sagar dam to Hamsala Deevi passing through Krishna, Guntur, and parts of Prakasam districts. The climate of the river basin varies from humid to sub-humid and arid conditions. Most precipitation occurs in the southwest monsoon season (June to September). In other seasons all tributaries and master

stream are also ephemeral streams. It is necessary to study climate change impacts on hydrological and water resources of the study area (Krishna River sub-basin containing 3 districts namely Krishna, Guntur and Prakasam Districts).

If we see the historical data, we can visualize that the frequency of heavy rainfall events is increasing year by year. As per the report of the Intergovernmental Panel on Climate Change (IPCC) 2013-AR5, it is predicted that the increase in temperature from the year 1990 to 2100 will be approximately 1.7°C to 4.9°C (Chen et al. 2012; Crawford et al. 2007). The study consists of the following steps: The Intergovernmental Panel on Climate Change (IPCC) in their fifth assessment report AR5, reported that the change in local precipitation and temperature due to climate change may cause the increase of hazards like droughts and floods and their severity (Cardona 2012; Randall et al. 2007) particularly at continental and larger scales. Confidence in these estimates is higher for some climate variables (e.g., temperature). General Circulation Model (GCM) models are capable to predict the expected change in climatic conditions for the future (Chen et al. 2012; Rajan 2014). Since the GCM model events are produced on a very massive grid
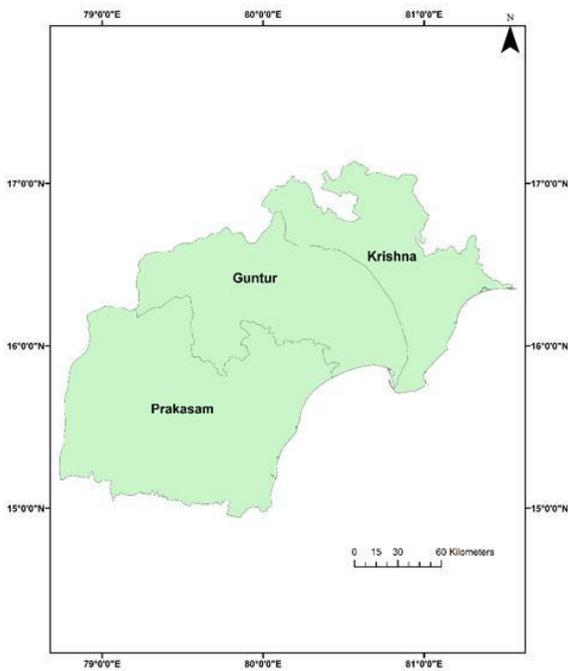
Fig. 1: Study area with longitudes and latitudes.

system (300 to 750 km) (Tan et al. 2017). Due to this large grid which is courser in nature the results obtained are not precise enough to be used directly to study the variation of different hydrological impacts at a local scale (Huang et al. 2011). General Circulation Model (GCM) models are capable to predict the expected change in climatic conditions for the future. Since the GCM model results are created on a very larger grid-scale (200 to 650 km) (Shukla et al. 2016). Due to this large grid which is courser in nature the results obtained are not precisely sufficient to be applied straight to study the change of different hydrological influences at the local scale (Nourani et al. 2019; Xu et al. 2020) little information is available regarding the downscaling using machine learning methods, specifically at hydrological basin scale. To overcome these scale parameters, we used downscaling which is capable to fill the rift among the local scaled climatic inputs and global scaled climatic parameters (Mahmood and Babel 2014, Wilby et al. 2002, Wilby & Dawson 2015). In the 21st century, climate change is considered to be one of the greatest environmental threats to the world, and the changes in climate extremes are estimated to have greater negative impacts on human society and the natural environment than the changes in mean climate. This study presents the projections of future changes in extreme temperature events under A2 and B2 SRES scenarios using the statistical downscaling model (SDSM). So, we can say that projection across different scales which is also known as downscaling is a procedure

that relates local and regional-scale climate variables to the larger-scale atmospheric components.

DOWNSCALING USING ANN MODEL

In the present study, the Artificial Neural Networks (ANNs) along with PCA has applied to downscale the precipitation in different three regions of Andhra Pradesh. In the present study, we have applied Feed Forward Back Propagation (FFBP) algorithm to develop a multi-layer perceptron (MLP) ANN model for downscaling of precipitation in the study area. The detailed process of the general downscaling process and the downscaling by ANN are shown in Fig. 2 and Fig. 4 respectively. Available precipitation data is used for predictands and CanESM2 model is used as GCM for predictor (Fig.7). While a random selection of data is used in the ANN model, 70 % of the data is used for the model and the remaining 30 % is used for testing and validation of the model. The performance evaluation of the ANN model is tested by R^2 and RMSE value.

The Multi-Layer Perceptron (MLP)

An artificial neural network (ANN) is a knowledge processing method based on the different data which has a comparable similar function like neurons of the human brain (Hannan et al. 2010; Kumar et al. 2012).

MLPs are the commonly accepted and the simplest type of ANN model (Joshi 2016). The MLPs are used to derive the relationship between different input and outputs (Ghosh and Misra 2010; Pervez and Henebry 2014). Multi-layer perceptron is feed-forward networks that comprise one or more hidden layers as shown in Fig. 3. The MLP used in the present study comprised a three-layer framework consisting of an input layer, a hidden layer, and an output layer. The
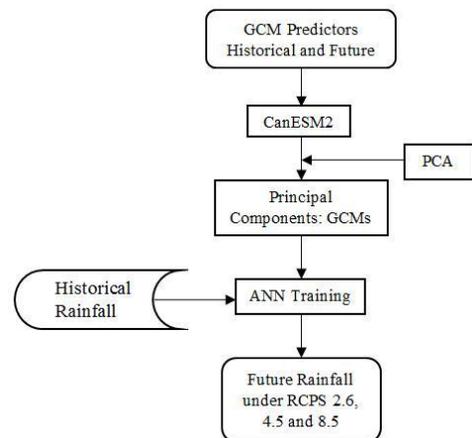

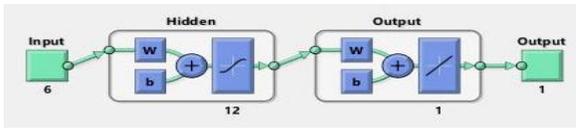Fig. 2: The flowchart for downscaling using ANN.

Fig. 3: The basic model of the ANN (Multi-layer Perceptron)
used in the study.

Levenberg-Marquardt (LM) algorithm (Luo et al. 2013) is an effective learning strategy for multilayer feed-forward networks (Kusunoki and Arakawa 2015; Onyutha et al. 2016). This method is a revised variant of the classic Newton approach for obtaining an optimum result for any optimization problem. This method uses an approach to the Hessian matrix as presented in Eq. (1).

$$x_{k+1} = x_k - [J^T J + \mu I]^{\wedge -1} J^T e \qquad \ldots(1)$$

Where, x is different weights of the neural network, J is the performance criteria Jacobian matrix and μ and e are learning process related parameter and residual error vector, respectively.

Principal Component Analysis (PCA)

Principal component analysis, also recognized as the Karhunen Loeve transform, is one of the commonly accepted techniques for reducing dimensionality (Beaumont 2012; Yang et al. 2017). PCA is used to modify a set of correlated M-dimensional predictors (as shown in Table 1) into another set of N-dimensional uncorrelated vectors called principal components (PCs) by using a linear combination (shown in Table 2). While transformation it is necessarily maintained that maximum information captured by the original data set is saved in the first few dimensions of the new set as presented in Table 3. In the present study, PCA was done to decrease the dimensionality of the predictors from 26 to 7, which contains the 99.98% information of the original data as shown in Table 1. These new datasets are used as input to the ANN downscaling model. So, we can say that to lessen the input dimension to the downscaling method and to pick the best set of predictors that hold all basic climate data we used PCA i.e. in PCA, the main aim is to find a set of N principal components which:

1. Is much less than the original set of M variables,

2. Having descriptions for nearly all of the total specimen variance.

MATERIALS AND METHODS

Study Area and Data Used in the Downscaling Process

The Krishna basin climate is dominated by the southwest monsoon, which provides most of the rainfall or precipitation for the basin. Maximum flow in the river occurs during
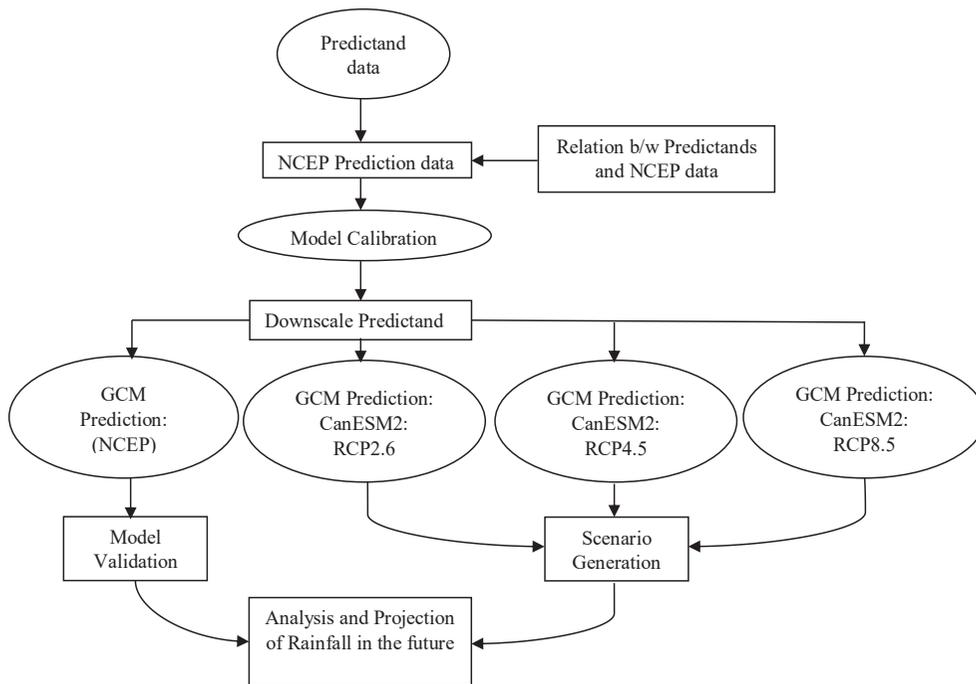

Fig. 4: The flowchart of the downscaling process using GCM Predictor.

the months of August-November and the low flow season is from April to May (at Vijayawada). The South-Central part of the basin is truly arid. On average, annual rainfall in the Krishna basin is 784 mm. The southwest monsoon sets in by the middle of June and withdraws by the middle of October. The annual rainfall received during the monsoon months is about 90%. Except for the monsoon months, the climate of the catchment remains dry. From the climatological observations, it is seen that the mean daily maximum temperature in the basin varies from 27°C to 40°C and the mean daily minimum temperature varies from 20°C to 27°C. The relative humidity in the basin ranges from 17 to 92%. Mean relative humidity is high during the monsoon period and comparatively low during the post-monsoon period. In summer, the weather is dry and the humidity is low. The catchment is influenced by south-west winds during the

Table 1: The score of different predictors after PC.

PCs	Score	Total Summation of first 7 PCs
PC1	89.99004	89.99004
PC2	09.56467	99.35471
PC3	0.352653	99.80737
PC4	0.035374	99.74274
PC5	0.022946	99.97569
PC6	0.013162	99.96585
PC7	0.006646	99.98555
PC8	0.003629	99.98913
PC9	0.002732	99.99186
PC10	0.001984	99.99384
PC11	0.001901	99.99574
PC12	0.001244	99.99699
PC13	0.001054	99.99804
PC14	0.000668	99.99871
PC15	0.000487	99.99919
PC16	0.000318	99.99951
PC17	0.000190	99.99971
PC18	0.000137	99.99984
PC19	8.43E-05	99.99992
PC20	6.74E-05	99.99999
PC21	7.79E-06	100.0000
PC22	1.12E-11	100.0000
PC23	2.48E-13	100.0000
PC24	1.56E-13	100.0000
PC25	1.15E-13	100.0000
PC26	5.83E-14	100.0000

monsoon season. In the post-monsoon season, they blow from north-west to north. In the winter season, the winds blow from the north-west and south-west directions. In the Krishna basin, wind speed varies from 4.0 to 21.7 kmph.

We have used shapefiles and rainfall data for our analysis. Fig. 5 shows the area for which shapefiles were digitized from various topo sheets and mosaicked using QGIS3.

The study area comprised of Krishna, Guntur and Prakasam Districts, which lies in the southern part of the Krishna River basin. The rainfall data is procured from Indian Meteorological Department, Hyderabad and http://imdpune.gov.in/ndc_new/ stations.html.

RAINFALL PATTERN

Theissen Polygons method: This method is applied to calculate average precipitation in different villages of the district.

The working principle of the Thiessen Polygon method is explained in Eq. (2) and shown in Fig. 6.

$$P_{avg} := \frac{\sum_{i=1}^{n} (A_i \cdot P_i)}{\sum_{i=1}^{n} A_i} \quad ..(2)$$

Where P_{avg} is the areal average precipitation over the watershed, A_i is the area of polygon i and P_i is the precipitation for polygon i. The integer n is the number of polygons and gauges.
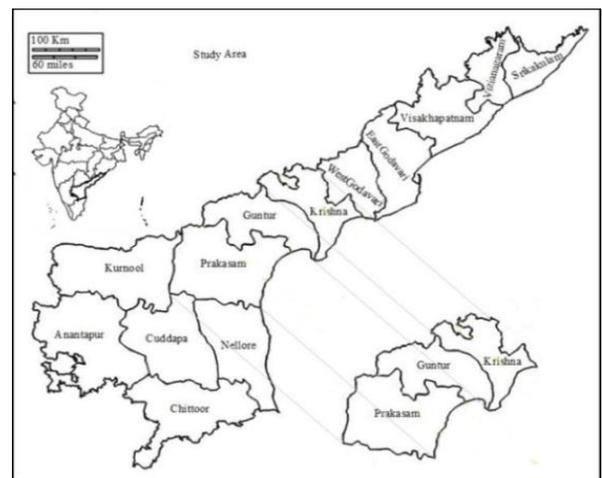

Fig. 5: Study area showing Krishna, Guntur and Prakasam districts.

Table 2: Correlation between 9 predictors before PCA.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
PC1	1	0.527673	0.255703	0.672074	0.685847	0.188658	0.797942	0.665339	0.673378
PC2	0.527673	1	0.696146	0.885125	0.479155	0.014698	0.487015	0.930171	0.850616
PC3	0.255703	0.696146	1	0.665569	0.242076	0.123365	0.217838	0.673246	0.722237
PC4	0.672074	0.885125	0.665569	1	0.552003	0.02717	0.578878	0.907174	0.859587
PC5	0.685847	0.479155	0.242076	0.552003	1	0.137442	0.867429	0.604516	0.640432
PC6	0.188658	0.014698	0.123365	0.02717	0.137442	1	0.182404	0.024236	0.006538
PC7	0.797942	0.487015	0.217838	0.578878	0.867429	0.182404	1	0.640683	0.669715
PC8	0.665339	0.930171	0.673246	0.907174	0.604516	0.024236	0.640683	1	0.960464
PC9	0.673378	0.850616	0.722237	0.859587	0.640432	0.006538	0.669715	0.960464	1

Table 3: Correlation between 9 predictors after PCA.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
PC1	1	8.44E-34	3.73E-31	4.75E-33	3.33E-32	5.10E-34	2.12E-34	6.80E-33	1.38E-33
PC2	8.44E-34	1	8.16E-32	4.96E-35	2.08E-32	1.50E-31	5.75E-32	1.71E-32	4.86E-33
PC3	3.73E-31	8.16E-32	1	1.62E-32	1.33E-31	5.61E-32	1.37E-33	6.20E-33	8.34E-34
PC4	4.75E-33	4.96E-35	1.62E-32	1	1.57E-31	4.26E-33	5.04E-33	1.56E-33	3.21E-32
PC5	3.33E-32	2.08E-32	1.33E-31	1.57E-31	1	4.87E-31	1.31E-31	5.69E-33	3.46E-32
PC6	5.10E-34	1.50E-31	5.61E-32	4.26E-33	4.87E-31	1	7.50E-34	1.78E-32	1.51E-32
PC7	2.12E-34	5.75E-32	1.37E-33	5.04E-33	1.31E-31	7.50E-34	1	2.08E-33	2.79E-32
PC8	6.80E-33	1.71E-32	6.20E-33	1.56E-33	5.69E-33	1.78E-32	2.08E-33	1	5.96E-32
PC9	1.38E-33	4.86E-33	8.34E-34	3.21E-32	3.46E-32	1.51E-32	2.79E-32	5.96E-32	1

Table 4: List of 26 NCEP/GCM predictors.

S.No.	Predictors (Atmospheric Variables)	Origin (i.e. NCEP or GCM)	Time Window
1.	Mean Sea Level Pressure	CanESM2	1961-1990
2.	1000hPa Wind Speed		(baseline climate)
3.	1000hPa Zonal Velocity		2000-2099
4.	1000hPa Meridional Velocity		
5.	1000hPa Vorticity		
6.	1000hPa Wind Direction		
7.	1000hPa Divergence		
8.	500hPa Wind Speed		
9.	500hPa Zonal Velocity		
10.	500hPa Meridional Velocity		
11.	500hPa Vorticity		
12.	500hPa Geopotential		
13.	500hPa Wind Direction		
14.	500hPa Divergence		
15.	850hPa Wind Speed		
16.	850hPa Zonal Velocity		
17.	850hPa Meridional Velocity		
18.	850hPa Vorticity		
19.	850hPa Geopotential		
20.	850hPa Wind Direction		
21.	850hPa Divergence		
22.	Specific Humidity at 500hPa		
23.	Specific Humidity at 850hPa		
24.	Specific Near Surface Humidity		
25.	Mean Temperature at 2m		
26.	Total Precipitation		

Climate Change Downscaling

Downscaling, or projection across different scales, is a procedure that links local and regional-scale climate variables to the larger scale atmospheric components (Zhang et al. 2016).

Downscaling joins the gap between large and local scale climatic data. The interpretation across scales is based on the assumption that similar atmospheric models produce similar climatic conditions. Now the basic question is that why we need downscaling?

In a better way, we can answer this question that: GCM (General Circulation Models) outputs are of insufficient spatial and temporal resolution, causing an insufficient representation of orography and land surface characteristics, when interoperated may cause lack of some of the features which may have important impacts on the local climate. To overcome these faults, we have to find a way which fulfils the gap by connecting the information that the climate modeling society can currently provide and that needed by different researchers. To overcome this scale mismatch two approaches have been suggested:

1. Develop finer resolution regional climate models that are driven by boundary conditions simulated by global GCMs at coarser scales i.e. RCMs. But they are computationally costly and time taking (Feyissa et al. 2018; Zhang et al. 2016) which have been shown to benefit water resources management and prediction, especially at the basin scale. In this study, the Soil and Water Assessment Tool (SWAT).

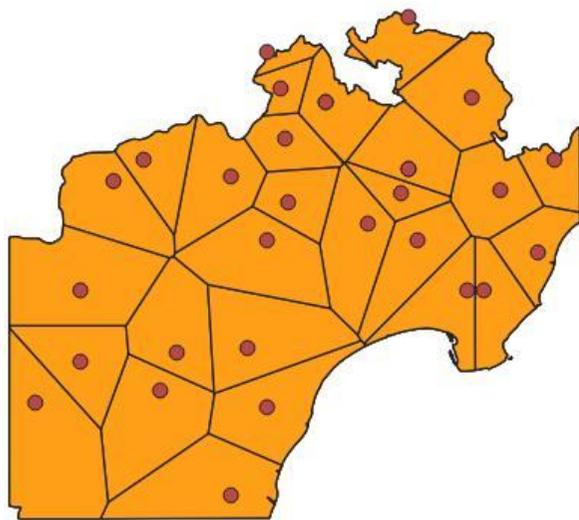2. Derive statistical models from the observed relationship

Fig. 6: Theissen Polygons method.

between the large-scale atmospheric fields and local variables and that is achieved by Statistical Downscaling (Chen et al. 2012).

Selection of GCM Model

Global Climate Change Model: CanESM2 Predictors: CMIP5 Experiments: The second generation Canadian Earth System Model (CanESM2):

CanESM2 is the fourth generation coupled global climate model developed by the Canadian Centre for Climate Modelling and Analysis (CCCma) of Environment and Climate Change Canada. CanESM2 represents the Canadian contribution to the IPCC Fifth Assessment Report (AR5). This CanESM2 model is a combination of CanCM4 model and the Canadian Terrestrial Ecosystem Model (CTEM), which based on the terrestrial carbon cycle (Feyissa et al. 2018) which helps us to determine consequences earlier and prepare for necessary adaptation measures. However, it is difficult to apply the raw data of GCMs at a local scale, such as the urban scale, without downscaling due to coarse resolution. This study, therefore, statistically downscaled daily maximum temperature, minimum temperature, and precipitation in 30-year intervals from the second generation of the Earth System Model (CanESM2). The CTEM model explains the land-atmosphere carbon transaction phenomena. CanESM2 consists of three Scenarios: RCP2.6, RCP4.5, and RCP8.5. The Representative Concentration Pathways (RCPs) are four greenhouse gas concentration (not emissions) trajectories selected by the IPCC for its Fifth Assessment Report (AR5) (Kim et al. 2016; Zhang et al. 2016) several downscaling methods have been developed. These methods predominantly focus on a single meteorological series at specific sites. Spatial and temporal correlation of the precipitation and temperature fields is important for hydrologic applications. This research uses a nearest neighbor-genetic algorithm (NN-GA). The four RCPs, RCP2.6, RCP4.5, RCP6.0, and RCP8.5, are described after a reasonable range of radiative forcing values projected in the year 2100. The different RCP Scenarios studied in this study is RCP2.6, RCP4.5, and RCP8.5. RCPs represent a broad area of possible problems related to Climate Change like the effect of greenhouse gases, air pollutants, and their emissions, and different land-use scenario. RCP8.5 considers the highest and RCP2.6 considers the lowest scenarios of greenhouse gases that have been recently reviewed by the study based on climatic research.

Structure of Predictor/Large Scale Dataset Files

A series of the graded daily value of long-term datasets are extracted into a single column text file per grid cell (box). This grid is uniform along the longitude with a horizontal

resolution of 2.8125° and nearly uniform along the latitude of roughly 2.8125°, which is available at the Canadian Climate service centre site. The predictors correlated with each grid cell are designed by the identical folder named *BOX_iiiX_jjY*, where iii represents longitudinal and jj represents the latitudinal index. The predictors organized this way can be used as input in statistical downscaling models.

Selection and Screening of Predictor Variables for Downscaling Process in Case of Precipitation

Choosing a predictor is a major measure in the downscaling method since finally, it reflects the main output i.e. nature of the generated scenarios. In the SDSM, it is a cyclic process that lasts until we get the optimized objective function (Chu et al. 2010; Mahmood and Babel 2014) evaporation, and precipitation in Haihe River basin, China. The data used for evaluation were large-scale atmospheric data encompassing daily NCEP/NCAR reanalysis data and the daily mean climate model results for scenarios A2 and B2 of the HadCM3 model. Selected as climate variables for downscaling were measured daily mean air temperature, pan evaporation, and precipitation data (1961-2000). The selection of large-scale predictors is a two-step process: firstly, we performed a correlation analysis between the NCEP reanalysis historical data with the past precipitation data to screen all the 26 predictor variables (NCEP Re-Analysis) for predictand data. Then the predictors having the highest correlation are selected for further processing. The metrological stations used for downscaling is listed in Table 5.

ANN performs two ways of model calibration based on the characteristics of climate data. They are known as conditional and unconditional methods. A conditional process is established for the precipitation and evaporation data analysis as they are based on the local scale predictors i.e. it is assumed that there is an indirect connection between the data and predictors. In case of an unconditional process, which is used for downscaling of temperature data with considering that there is a direct link to the predictor's i.e. large-scale datasets. During the calibration process, the NCEP-Re-analysis data set is used in accordance with the specified year period for each predictand (as shown in Table 6). The historical data of predictands in this study i.e., precipitation, are divided into two segments: the first segment is used for calibration of the model and the second part of the dataset is used for validation as an independent dataset.

Model Calibration and Validation

In ANN there are two ways to optimize the model output, one is Ordinary Least Square (OLS) method and the other one is Dual Simplex (DS). In the present study, OLS is used because it is faster than DS (Chu et al. 2010; Wilby and Dawson 2015). The Root Mean Square Error (RMSE) and Coefficients of Correlation (R^2) were used to compare the performance of historical and simulated data of the model during the calibration and validation period. The model was calibrated for the period from 1961-1995 is used as the base period and simulate the daily rainfall for the period of 1996-2005 with the help of NCEP and CanESM2 predictors which is used for the validation of the model. The description of R^2 and RMSE of rainfall data are given in Table 7.

WORKING OF ANN MODEL

Data Used in the ANN Model

The rainfall data obtained from IMD are shown in Fig.8 (a-c) for all the three districts of Andhra Pradesh. In case of any missing rainfall data in IMD database, the linear interpolation method is used to find the missing data. As discussed earlier that PCA is applied to select the predictor to train the ANN model. The first seven principal components i.e. predictors are used in the analysis which covers the 99.77% of the predictor data property. In the ANN model the selection of data is random and 70% of data is used for training purpose and the remaining 30% is used for testing and validation. Multi-layer perceptron is feed-forward networks that are based on the Levenberg-Marquardt (LM) algorithm that has been
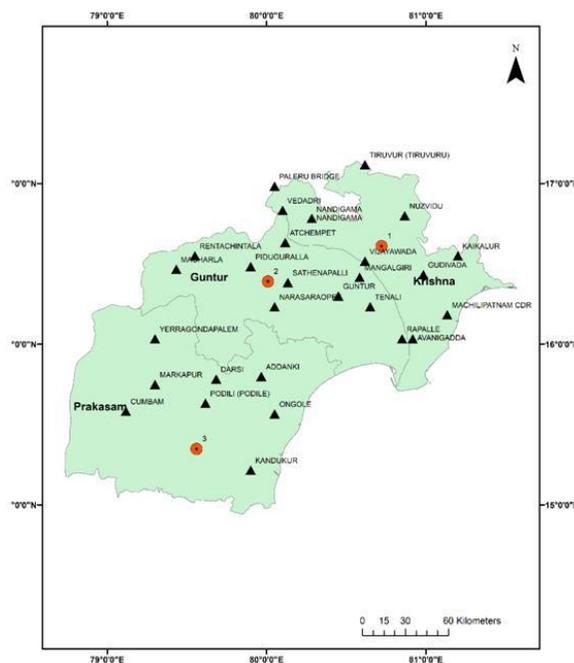

Fig. 7: Rain gauge stations (Black Triangles) and GCM Points (Round Points) in the study area.

used in the ANN model. A total of twelve hidden layers are used in the model along with the six inputs as shown in Fig. 3. The nine hidden layer combination give the best Correlation Coefficient in comparison with the other combination when we are using more or less than nine hidden layers. In this study, three statistical indicators were used to measure the efficiency of the artificial neural network models developed. The assessment indicators include coefficient determination (R^2), Nash Coefficient (E) and Root Mean Square Error (RMSE), as the same is used in the evaluation of the SDSM model.

(a) Root Mean Square Error (RMSE)

The Root Mean Square Deviation (RMSD) or Root Mean Square Error (RMSE) is commonly used to measure the deviations between sample or population values predicted by a model and the observed values (Xiao et al. 2015).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model,i})^2}{n}} \qquad ...(4)$$

Where X_{obs} is observed values and X_{model} is modelled values at time/place i.

(b) Coefficient of Determination (R^2)

The coefficient of Determination (R^2): Shows the intensity

Table 5: Metrological stations used for downscaling.

Station Name	Longitude (degrees)	Latitude (degrees)	Period (year)
Krishna	80.7214 E	16.6100 N	1961-2012
Guntur	80.0088 E	16.3906 N	1961-2012
Prakasam	79.5603 E	15.3485 N	1961-2012

and control of a linear relationship between two variables (Kazmi et al. 2014). The correlation is +1 in the case of a perfect increasing linear relationship, and -1 in the case of a decreasing linear relationship.

$$R^2 = \frac{[\sum_{i=1}^{n}(O_i - \bar{O}_t) * (P_i - \bar{P}_t)]^2}{[\sum_{i=1}^{n}(O_i - \bar{O}_t) * \sum_{i=1}^{n}(P_i - \bar{P}_t)]} \qquad ...(5)$$

Where, O_i and P_i is observed and simulated value, n is the total number of test data $\bar{P}_t, \bar{O}_t$ is the mean value.

(c) Nash-Sutcliffe Coefficient (E)

The Nash-Sutcliffe Model Efficiency Coefficient (E) is commonly used to assess the predictive power of hydrological discharge models (Chen et al. 2012; D. N. Moriasi et al. 2007). It is defined as:

$$E = 1 - \frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model})^2}{\sum_{i=1}^{n}(X_{obs,i} - \bar{X}_{obs})^2} \qquad ...(6)$$

Table 6: Data used for calibration and validation of ANN.

Station Name	Longitude (degree)	Latitude (degree)	Data available Period (year)	Calibration Period	Validation Period
Krishna	80.7214 E	16.6100 N	1961-2012	1961-1995	1996-2005
Guntur	80.0088 E	16.3906 N	1961-2012	1961-1995	1996-2005
Prakasam	79.5603 E	15.3485 N	1961-2012	1961-1995	1996-2005

Table 7: R^2 and RMSE value during calibration and validation of the model.

Station Name	Calibration			Validation		
	R^2	RMSE	Nash Coefficient, E	R^2	RMSE	Nash Coefficient, E
Krishna	0.82	3.88	0.80	0.83	3.31	the 0.82
Guntur	0.77	3.29	0.78	0.69	3.34	0.81
Prakasam	0.81	2.91	0.71	0.86	1.99	0.70

Table 8: The parameters selected for the ANN downscaling model.

ANN Network Type	Parameters	Name
MLP (Multi-Layer Perceptron: feed-forward networks)	Number of layers:	03
	Neurons:	
	Inputs:	05
	Hidden:	12
	Output:	01
	Number of iteration:	Sigmoid Losing linear
	Activation function:	Logsig linear
	Activation function in output layer:	Levenberg-Marquardt

Where X_{obs} is observed values and X_{model} is modelled values at time/place i. Nash-Sutcliffe efficiencies can range from $-\infty$ to 1. An efficiency of 1 ($E = 1$) corresponds to a perfect match between model and observations.

Sensitivity Analysis shows that the ANN model combined with PCA for the selection of predictors is more suitable than the SDSM model as shown in Table 4 and Table 8. Now for the Scenario Analysis for the three RCPs 2.6, 4.5 and 8.5, the ANN model is used.

Scenario Analysis

In the present study, for predicting the effect of Climate Change on precipitation trends CanESM2 GCM scenario, i.e. RCP2.6, RCP4.5, and RCP8.5 is used. The period from 1961-2000 is selected as the base period to visualize the changing pattern of rainfall. The selection of the base period is based on the literature review available and these 40 years of data is sufficient to assess the transformation in climate. So, the prediction of future rainfall is based on

Table 9: Projected future changes of mean precipitation in the 3 districts of Andhra Pradesh.

Scenario Variable	RCP2.6				RCP4.5				RCP8.5			
	2020s	2050s	2080s	2100s	2020s	2050s	2080s	2100s	2020s	2050s	2080s	2100s
Guntur ($\Delta P/P$ (%))	-1.69	-0.33	4.73	5.25	1.02	-1.78	7.07	8.59	5.82	5.38	9.27	19.51
Krishna ($\Delta P/P$ (%))	-0.22	1.25	-1.89	4.66	1.30	4.81	6.13	8.09	8.00	11.39	12.53	18.50
Prakasam ($\Delta P/P$ (%))	-1.08	2.70	1.53	5.05	1.05	2.80	5.17	10.7	4.32	9.18	10.85	15.99
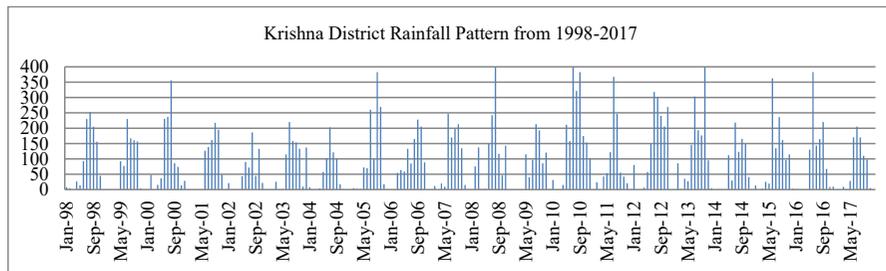
(a)　Krishna District Rainfall Pattern
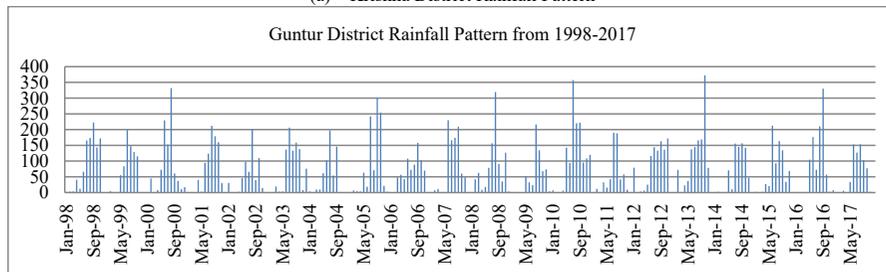
(b)　Guntur District Rainfall Pattern
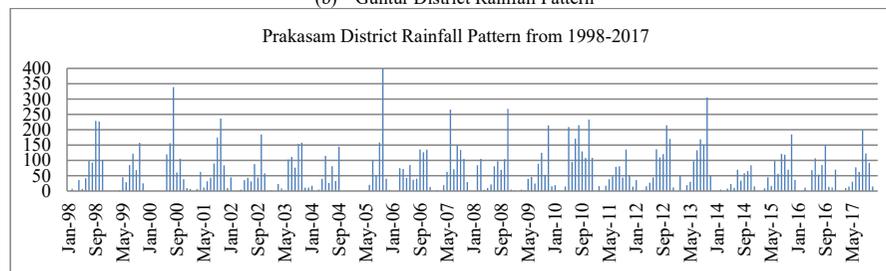
(c)　Prakasam District Rainfall Pattern

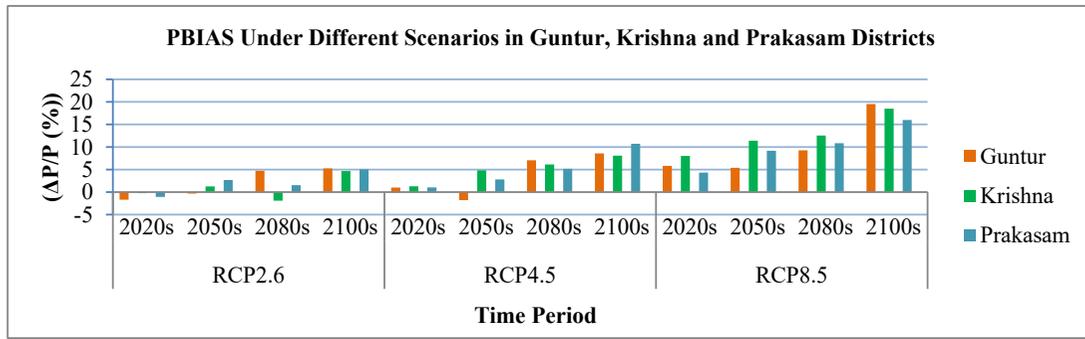Fig. 8(a-c): Mean monthly rainfall from three districts of Andhra Pradesh.

Fig. 9: Showing PBIAS under different scenarios.

the comparison of these two-time extents i.e., 1961-2000 and 2005-2100.

After calibration and validation of the ANN model, the model is used to downscale the large scale predictor variables derived from the RCP2.6, RCP4.5, and RCP8.5 scenarios of CanESM2, with daily precipitation simulated for the following periods: historical (1961-2000), the 2020s (2005-2021), 2050s (2022-2051), 2080s (2052-2080) and 2100s (2081-2100). As mentioned above, the historical simulation (1961-2000) acts as a reference for future projection and changes. Predicted changes in annual mean precipitation during future periods (the 2020s, 2050s, 2080s, and 2100s) in the three districts namely Guntur, Krishna, Prakasam of Andhra Pradesh are shown in Table 9, which shows a mixed pattern of positive or negative changes, with different trends in the 2020s and 2050s, and steady with the increases in the 2080s and 2100s. The trend shows that the overall amount of rainfall will increase significantly in this region of all areas from 2081 to 2100 as compared to the base period due to climate change. While there is a mixed trend in rainfall, under CanESM2 scenarios of RCP2.6 and RCP4.5, the fluctuation in rainfall is slightly different for two scenarios i.e. RCP2.6

and RCP4.5 as compared to the emission scenario of RCP8.5. In RCP2.6, up to 2080s sometimes the rainfall is increasing and in some decades it goes down. In RCP4.5, the trend is slightly higher than the base period for the 2020s and 2050s and in 2080s and 2100 it increases abruptly.

The approximate predicted change up to the 21st century will be 19.51%, 18.50%, and 15.99% for Guntur, Krishna and Prakasam basin respectively under the scenarios of RCP8.5 (Table 9) and PBIAS under different scenarios is as shown in Fig. 9.

The ANN model calibration and validation results are shown in Fig. 10(a-i).

CONCLUSIONS

The present study attempted to forecast the rainfall pattern in three different basins of Andhra Pradesh. The present study tried to investigate the use of the ANN model as a tool for downscaling monthly precipitation for application in Climate Change studies. Firstly, three basins are chosen to perform the downscaling analysis using ANN model. The selection of predictors is based on the PCA was performed for the ANN
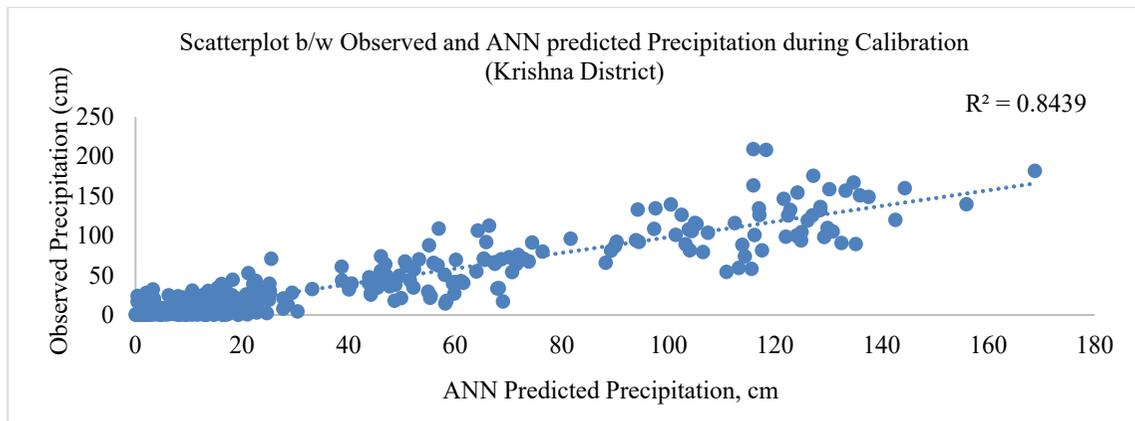

Fig.10(a): Scatterplot b/w observed and ANN predicted precipitation during Validation (Krishna District).
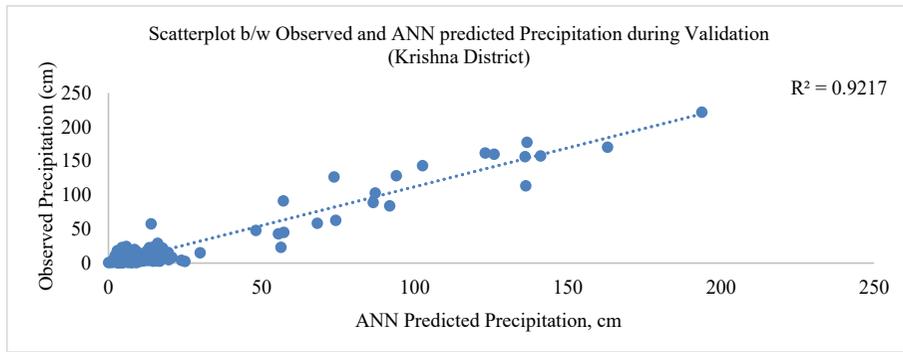
Fig.10(b): Scatterplot b/w observed and ANN predicted precipitation during Validation (Krishna District).
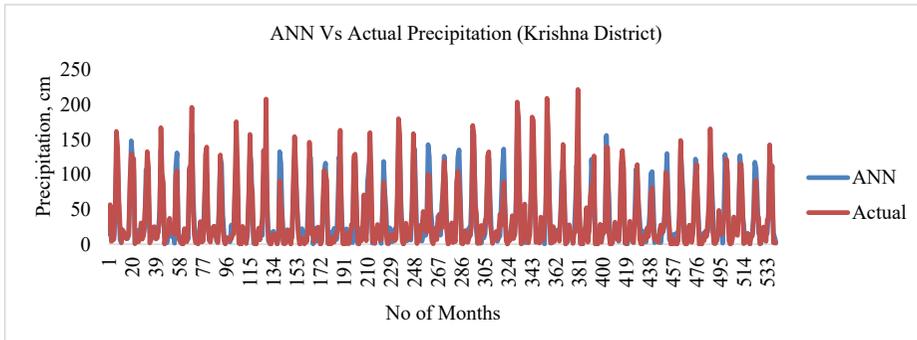

Fig. 10(c): Plot between ANN and actual precipitation (Krishna District).
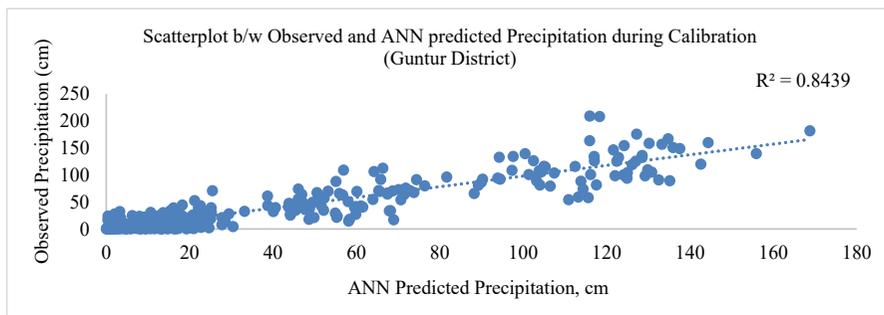

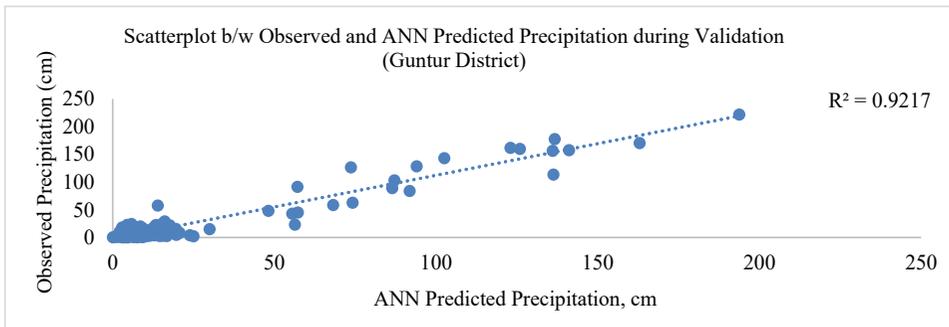Fig. 10(d): Scatterplot b/w observed and ANN Predicted Precipitation during Calibration (Guntur District).

Fig.10(e): Scatterplot b/w Observed and ANN Predicted Precipitation during Validation (Guntur District).
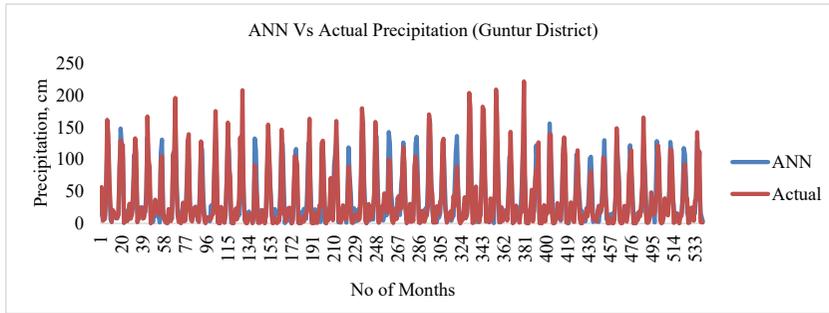
Fig.10(f): Plot between ANN and actual precipitation (Guntur District).
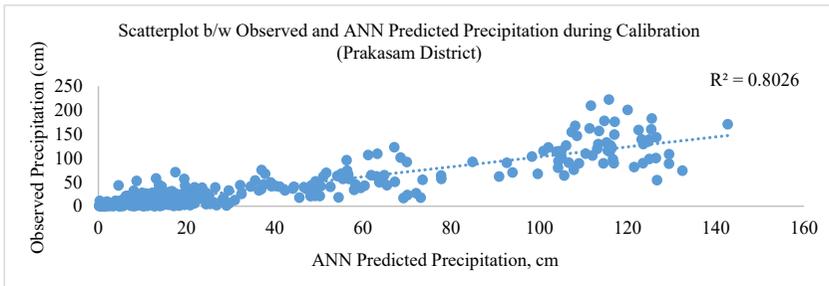

Fig.10(g): Scatterplot b/w observed and ANN predicted precipitation during calibration (Prakasam district).
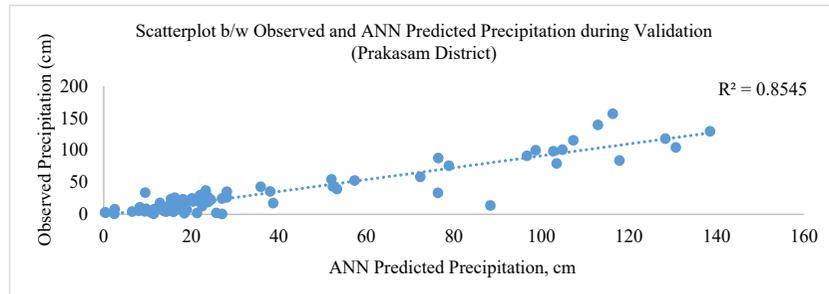

Fig.10(h): Scatterplot b/w Observed and ANN Predicted Precipitation during Validation (Prakasam district).
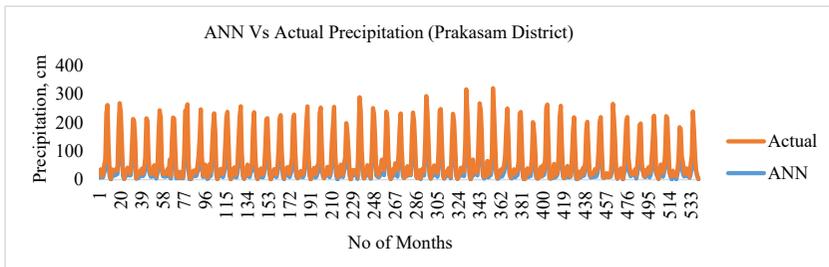

Fig.10(i): Plot b/w ANN and Actual Precipitation (Prakasam district).

model. After the calibration and validation of both the model, the sensitivity analysis was performed using R2, RMSE, and E value and also by plotting the observed versus simulated scatter plots for both. The analysis shows that the ANN is performed well in these three different basins of Andhra Pradesh. The Scenario Analysis was performed for all the regions using the ANN model under three different RCPs 2.6, 4.5 and 8.5 to visualize the impact of Climate Change

in the regions. The purpose behind finding out the projected change in precipitation in different regions of the state is to see the impact of Climate Change on Vulnerability. In this study, the historical data of three stations was used from the period of 1961-2000. The result of ANN model predicted that rainfall will be increased under the entire RCP scenario especially in the case of RCP8.5. In RCP2.6, there is little fall of rainfall in the 2020s and 2050s. The seasonal rainfall varies from one station to another, especially during the monsoon season where there will be very high rain, especially in RCP8.5 scenarios.

In all the scenarios, it can be visualized that the change in the rainfall is considerable after the 2050s. The changes in rainfall up to 2050s are not notable. However, in 2080, and 2100, the predictions are alarming in nature. The Seasonal Analysis of rainfall data also shows the shifting of the monsoon trend up to September especially in the case of RCP8.5, besides, the supposed rise of rainfall due to global climate change. Finally, the conclusions (findings) are summarized as:

1. The approximate predicted change up to the 21st century will be 19.51%, 18.50% and 15.99% for Guntur, Krishna and Prakasam basin respectively under the scenarios of RCP8.5.

2. The approximate predicted change up to the 21st century will be 8.59%, 8.09% and 10.70% for Guntur, Krishna and Prakasam basin respectively under the scenarios of RCP4.5.

3. In RCP2.6 scenario, there is little fall of rainfall in the years 2020s and 2050s and in RCP4.5 Scenario, there will be little rainfall in the year 2020.

4. In RCP8.5 scenario, there will be high rainfall in the years 2080 and 2100, and rainfall is highest in Guntur district, second highest Krishna district.

5. As per RCP8.5, the rainfall in Guntur districts in the year 2100 is nearly double that in the year 2080. The rainfall in Krishna and Prakasam districts in the year 2100 is nearly 1.5 times that in the year 2080.

Hence, based on the conclusions of this research it is suggested that we should develop better mitigation measures to counter such heavy rainfall trends in the future. The policymakers and the local governments should focus on the better planning of water management in the state especially for water storage capacity and harmless drainage system.

REFERENCES

Beaumont, R. 2012. An introduction to principal component analysis & factor analysis using SPSS 19 and R (psych package). Journal of Geophysical Research, (April), P.No.24-32.

Cardona, O.D., Van Aalst, M.K., Birkmann, J., Fordham, M., Mc Gregor, G., Rosa, P., Pulwarty, R.S., Schipper, E.L.F., Sinh, B.T., Décamps, H. and Keim, M., 2012. Determinants of risk: Exposure and vulnerability. In: Managing The Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change, pp. 65-108, Cambridge University Press.

Chen, H., Xu, C.Y. and Guo, S. 2012. Comparison and evaluation of multiple GCMs, statistical downscaling and hydrological models in the study of climate change impacts on runoff. Journal of Hydrology, 434: 36-45.

Chu, J. T., Xia, J., Xu, C. Y. and Singh, V. P. 2010. Statistical downscaling of daily mean temperature, pan evaporation and precipitation for climate change scenarios in Haihe River, China. Theoretical and Applied Climatology, 99(1-2): 149-161.

Crawford, T., Betts, N. L. and Favis-Mortlock, D. 2007. GCM grid-box choice and predictor selection associated with statistical downscaling of daily precipitation over Northern Ireland. Climate Research, 34(2): 145-160.

Feyissa, G., Zeleke, G., Bewket, W. and Gebremariam, E. 2018. Downscaling of future temperature and precipitation extremes in Addis Ababa under climate change. Climate, 6(3): 58.

Ghosh, S. and Misra, C. 2010. Assessing hydrological impacts of climate change: modeling techniques and challenges. The Open Hydrology Journal, 4: 115-121.

Hannan, S. A., Manza, R. R. and Ramteke, R. J. 2010. Generalized regression neural network and radial basis function for heart disease diagnosis. International Journal of Computer Applications, 7(13): 7-13.

Huang, J., Zhang, J., Zhang, Z., Xu, C., Wang, B. and Yao, J. 2011. Estimation of future precipitation change in the Yangtze River basin by using statistical downscaling method. Stochastic Environmental Research and Risk Assessment, 25(6): 781-792.

Joshi, R. 2016. Artificial neural network (ANN) based empirical interpolation of precipitation. International Journal of Mathematical, Engineering and Management Sciences, 1(3): 93-106.

Kazmi, D. H., Rasul, G., Li, J. and Cheema, S. B. 2014. Comparative study for ECHAM5 and SDSM in downscaling temperature for a geo-climatically diversified region, Pakistan. Applied Mathematics, 2014 (January), 137-143.

Kim, S., Kwak, J., Kim, H. S., Jung, Y. and Kim, G. 2016. Full access nearest neighbor-genetic algorithm for downscaling of climate change data from GCMs. Journal of Applied Meteorology and Climatology, 55(3): 773-789.

Kumar, D., Bhishm, S. K. and Khati, S. 2012. Black box model for flood forecasting. Journal of Civil Engineering (IEB), 40(June 2011): 47-59.

Luo, Y., Ficklin, D. L., Liu, X., and Zhang, M. (2013). "Assessment of climate change impacts on hydrology and water quality with a watershed modeling approach." Science of the Total Environment, Elsevier B.V., 450–451, 72–82.

Kusunoki, S. and Arakawa, O. 2015. Are CMIP5 models better than CMIP3 models in simulating precipitation over East Asia? Journal of Climate, 28(14): 5601-5621.

Mahmood, R. and Babel, M. S. 2014. Future changes in extreme temperature events using the statistical downscaling model (SDSM) in the trans-boundary region of the Jhelum river basin. Weather and Climate Extremes, 5(1): 56-66.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D. and Veith, T. L. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE, 50(3): 885-900.

Nourani, V., Razzaghzadeh, Z., Baghanam, A. H. and Molajou, A. 2019). ANN-based statistical downscaling of climatic parameters using decision tree predictor screening method. Theoretical and Applied Climatology, 137(3-4): 1729-1746.

Onyutha, C., Tabari, H., Rutkowska, A., Nyeko-Ogiramoi, P. and Willems, P. 2016. Comparison of different statistical downscaling methods

for climate change rainfall projections over the Lake Victoria basin considering CMIP3 and CMIP5. Journal of Hydro-Environment Research, 12: 31-45.

Pervez, M. S. and Henebry, G. M. 2014. Projections of the Ganges-Brahmaputra precipitation - Downscaled from GCM predictors. Journal of Hydrology, 517: 120-134.

Rajan, S. 2014. Statistical downscaling of gcm output, hydrological simulation and generation of future scenario using variable infiltration capacity (VIC) model for the Ganga basin, India. Uttarakhand: Andhra University, 132.

Randall, D. A., Wood, R. A., Bony, S., Colman, R., Fichefet, T., Fyve, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., Stouffer, R. J., Sumi, A. and Taylor, K. E. 2007. Climate models and their evaluation. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, 591-662.

Shukla, R., Khare, D. and Deo, R. 2016. Statistical downscaling of climate change scenarios of rainfall and temperature over Indira Sagar canal command area in Madhya Pradesh, India. Proceedings - 2015 IEEE 14th International Conference on Machine Learning and Applications, ICMLA 2015, 313-317.

Tan, M. L., Ibrahim, A. L., Yusop, Z., Chua, V. P. and Chan, N. W. 2017. Climate change impacts under CMIP5 RCP scenarios on water resources of the Kelantan River Basin, Malaysia. Atmospheric Research, 189: 1-10.

Wilby, R., Dawson, C. and Barrow, E. 2002. SDSM - a decision support tool for the assessment of regional climate change impacts. Environmental Modelling & Software, 17(2): 145-157.

Wilby, R. L. and Dawson, C. W. 2015. Statistical DownScaling Model-Decision Centric (SDSM-DC) Version 5.2 Supplementary Note 6 March 2015, 5.

Xiao, R., He, X., Zhang, Y., Ferreira, V. G. and Chang, L. 2015. Monitoring groundwater variations from satellite gravimetry and hydrological models: A comparison with in-situ measurements in the mid-atlantic region of the United States. Remote Sensing, 7(1): 686-703.

Xu, R., Chen, N., Chen, Y. and Chen, Z. 2020. Downscaling and projection of multi-CMIP5 precipitation using machine learning methods in the upper Han river basin. Advances in Meteorology, Atmospheric Research, 2020.

Yang, P., Xia, J., Zhan, C., Qiao, Y. and Wang, Y. 2017. Monitoring the spatio-temporal changes of terrestrial water storage using GRACE data in the Tarim River basin between 2002 and 2015. Science of the Total Environment, 595: 218-228.

Zhang, Y., You, Q., Chen, C. and Ge, J. 2016. Impacts of climate change on streamflows under RCP scenarios: A case study in Xin River Basin, China. Atmospheric Research, 178-179: 521-534.