Projection of Temperature and Precipitation using Multiple Linear Regression and Artificial Neural Network as a Downscaling Methodology for Upper Bhima Basin

Dattatray Kisan Rajmane, Milind Laxman Waikar

Abstract: Study of Climate change effect on water resources is very important for its effective management. Projection of temperature and precipitation can be performed by using General Circulation Model (GCM) outputs. GCM can make the projections of climate parameters with different emission scenarios at coarser scale. However hydrological models require climate parameters at smaller scale Downscaling technique is used for obtaining small scale climate variables from large scale variables of GCM outputs. In this study downscaling has been carried out by using Multiple Linear Regression (MLR) and Artificial Neural Network (ANN) techniques. Performance of MLR and ANN models has been evaluated considering Coefficient of determination value (R^2). It has been observed that ANN performs better against MLR Model, showed the results that rainfall distribution pattern is varied, in monsoon season rainfall decreases while it increases in post monsoon period. Due to its good evaluation performance such techniques can be applicable for downscaling purpose.

Keywords: Artificial neural network, General Circulation model, Multiple linear regression, Upper Bhima Basin

I. INTRODUCTION

In general climate change is defined as "The difference between long term mean values of a climate parameter, where the mean is taken over a specified interval of time usually a number of decades" [1]. The main cause of climate change is global warming. Global warming is caused because of increase in concentrations of gases like carbon dioxide, methane and nitrous oxide (Greenhouse gases) due to human activities such as industrialization, deforestation and burning fossil fuels. Impact of climate change is over the many fields like agricultural, medical, economical, water resources and others. As per Intergovernmental panel on Climate Change (IPCC), water scarcity has been expected in future in various seasons, changes in the seasonal distribution and amount of precipitation, rise in sea level are the main effects of climate change on water resources.[2] Therefore, it is very essential to project the climate parameters like temperature and precipitation to study its impact on water resources in particular area. This projection of climate parameters can be performed by using General Circulation Model, also called as Global Climate Model (GCM) GCM is a mathematical representation of general circulation between air, sea, cryosphere and land surface which simulates the time series of climate variables globally, considering the effects of greenhouse gases. It is a numerical model based on Navier-Stoke equation. GCM's can make the projection of possible future climate change parameters under different emission scenarios. These scenarios are formulated according to future greenhouse gases concentrations, land use and other driving forces. Different emission scenarios are RCP 2.0, RCP 4.5, RCP 6.0 and RCP 8.5. In this study RCP 6.0 scenario is considered. RCP 6.0 stand for Representative Concentration Pathway and 6.0 W/m^2 will be radiative forcing in the year 2100 relative to 1750 [3]. For RCP 6.0, prescribed CO_2 concentration is about 670 ppm A limitation of GCM is that it simulates the climate variables at coarser scale (low resolution), may be at 2.5^0 x 2.5^0 grid (1^0 is about 110 km). For the study of impact of climate change on water resources, most of the hydrological models require local scale (high resolution) input data. So, to convert this coarser scale output data of GCM to finer scale, downscaling techniques are used.

II. DOWNSCALING TECHNIQUE

Downscaling technique is used to bridge the gap between large scale variables (GCM outputs) and local scale hydro metrological variables like rainfall, runoff and temperature etc. downscaling can be carried out by two ways [4].

A. Dynamic Downscaling

It represents the use of high-resolution Regional Climate Models (RCM) which are nested with GCM [5]. It is a physical based model and suffers a drawback of its complicated design and high computation cost.

B. Statistical Downscaling

Statistical downscaling is the process of obtaining statistical relation between large scale atmospheric variables and local scale hydro metrological variables. Once the equation is formulated, it is applied to project future hydro metrological variables using the large-scale future climate variables (GCM outputs).

∗ Correspondence Author
   Dattatray Kisan Rajmane ∗, Department of Civil Engineering, Shri Guru Gobind Singhji Institute of Engg. & Tech., Nanded, India. E-mail: dkrajmane@gmail.com
   Dr. Milind Laxman Waikar, Department of Civil Engineering, Shri Guru Gobind Singhji Institute of Engg. & Tech., Nanded, India. E-mail: mlwaikar@sggs.ac.in

1

This technique is broadly divided into three categories i) Weather Typing ii) weather Generator iii) Regression based downscaling [6]. In this study, regression-based downscaling by multiple linear regression and Artificial Neural Network techniques are used for downscaling and projection of climate variables.

- *Multiple Linear Regression (MLR):*

Regression analysis is very useful for forecasting. In MLR, dependent variable (Predictand) is one and independent variables (Predictors) are more than one. The basic model is designed between dependent and independent variables by using least square method. A relation between dependent and independent variables is linear. Once the model is developed, it will be used for forecasting of predictand value corresponding the predictors values from GCM outputs.

- *Artificial Neural Network (ANN):*

ANN has been developed as a generalization of mathematical models of human cognition or neural biology [7]. It is a network consists of nodes which are linked to each other by layer to layer. In typical neural network as shown in fig.1 [8], first layer is input layer and final layer is output layer. There can be one or more hidden layers consisting of one or more nodes. Each node is connected to the nodes in previous layer and next layer through links. Rumlhart (1986), rediscovered a backpropagation algorithm by which an error is propagated backward in the network through links and correspondingly the connection weights and bias values are adjusted [9].
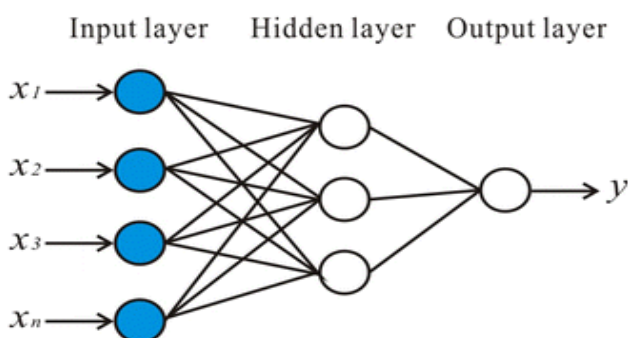

Fig 1. Feed forward back propagation neural network

Training process is stopped when there is no appreciable change in values of weights. After training of the model, it can be validated by using another data set which was not used for training purpose.

III. STUDY AREA

The Bhima river is a major tributary of Krishna river in India. It originates from a Jyotirling named Bhimashankar in Ambegaon Taluka of Pune district of Maharashtra. It flows southeast for 300km long and covering watershed area of 14,856 km^2 in Maharashtra state [10]. Ujjani dam is constructed on Bhima river near to village Ujjani. The basin area on upstream side of Ujjani dam is called as Upper Bhima basin. Upper Bhima basin shown in fig.2 is located within the coordinates of 17.18N to 19.24N latitude and 73.20E to 76.15E longitude. Rainfall variation in basin ranges from 415 mm to 4240 mm. The Ujjani dam provides wide range of economic, environmental and social benefits. Dam provides the facilities of limited flood control, drinking as well as

industrial water, hydroelectricity etc. Therefore, it is very essential to study the effect of climate change on the reservoir inflow to this dam for proper management of the water resources.
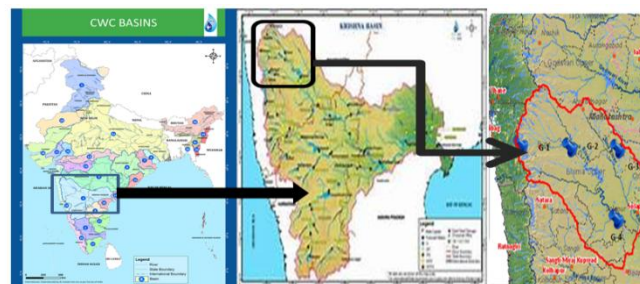

Fig 2. Basin map of India, Krishna Basin, and Upper Bhima basin (Source: India WRIS, http:// www.indiawris.gov.in)

IV. DATA COLLECTION AND PROCESSING

A. Data Collection

Three types of data have been used for this study.

- *Observed Temperature and Precipitation Data*

Observed temperature and rainfall (Predictands) data has been collected from Indian Metrology Department, Pune. It is available at 1^0 x 1^0 grid. Data of four grid stations as mentioned in Table1 lying within the Upper Bhima basin has been acquired. Data has been collected from Jan 1969 to Dec. 2015 period.

Table1 Station Points for Upper Bhima basin

Station	Latitude-Longitude	Place	State
1	17.25N -75.25E	Tembhurni	Maharashtra
2	18.25N -73.25E	Mulshi	Maharashtra
3	18.25N -74.25E	Supa	Maharashtra
4	18.25N -75.25E	Kashti Ghod	Maharashtra

- *National Centre for Environmental Protection (NCEP) Data*

NCEP data (Predictors for Training) is basically the data of observed atmospheric variables. Spatial resolution of this data is about 2.5^0 x 2.5^0. This data is downloaded from site https://sdsm.org.uk on daily basis then it is converted in average monthly scale.

- *Global Climate Model Data*

GCM (Predictors for Forecasting) data for GFDL-CM3 model was downloaded from https://esgf-node.llnl.gov web site on monthly basis considering RCP 6.0 scenario. It is available in. netcdf format, can be made in readable form by using ArcGIS, MATLAB or Panoply software.

B. Data Processing

It is required to process the large-scale climate variables downloaded data before using it for model calibration.

- *Interpolation:*

NCEP data is available at 2.5^0 x 2.5^0 and GCM data is available at 2^0 x 2.5^0 resolution. IMD data resolution is 1^0 x 1^0.

Hence NCEP and GCM data are required to interpolate at the grid size of 1^0 x 1^0 to match the coordinates of four station points. Interpolation can be carried out by using MATLAB.

- *Standardization*

Standardization means rescaling the data to have mean value zero and standard deviation equal to one. It is the process of converting disparate dataset into common dataset format which gives consistent data form. Standardization is conventional method to remove the bias in GCM output data set [11].

$$\text{Standardized Value} = \frac{(\text{Original Value} - \text{Mean})}{\text{Standard Deviation}} \quad (1)$$

- *Selection of Predictors*

NCEP parameters and GCM outputs are called as predictors. These predictors are selected with the criteria of Pearson's correlation Coefficient [12]. This coefficient shows that whether the two variables (Predictor and Predictand) are strongly correlated, weakly correlated or independent. Here the predictand is temperature or precipitation. Following predictors shown in Table 2 were considered and used those which are strongly correlated with predictand at different stations.

Table 2. Predictor variables for downscaling Temperature and Precipitation

Predictor Variables	Notation	Units
Eastward wind@500hpa	p5_u	metres/second
Eastward wind@850hpa	p8_u	metres/second
Northward wind@500hpa	p5_v	metres/second
Northward wind@850hpa	p8_v	metres/second
Geopotential height @500hpa	p500	Metres
Geopotential height @850hpa	p850	Metres
Air pressure at sea level	Mslp	Pascal
Near surface relative humidity	Hurr	%
Near surface specific humidity	Hurs	Kilogram of vapour /kilogram of air
Surface air temperature @ 2m	Temp	Kelvin
Precipitation	prec	mm

V. METHODOLOGY

A. Using Multiple Linear Regression (MLR)

- Model Calibration, Validation and Evaluation

For calibration of the MLR model, NCEP selected predictor data from 1969 to 2005 has been used as input to the model and predictand (temperature or precipitation) from 1969-2005 as output of the model. Validation was done with the data from 2006 to 2015.The approach used for the calibration of the model is as below

$$y = A + B_1 x_1 + B_2 x_{2+} \text{---------------------} + B_n x_n \quad (2)$$

Where
y= Predictand (Temperature or Precipitation)
$x_1, x_2, \text{---------} x_n$ = Predictors (NCEP variables while Calibration of the model). Evaluation of the model was performed with the help of statistical parameter Coefficient of Determination (R^2 Value).

- Forecasting of Temperature and Precipitation

The calibrated and validated MLR model was used for forecasting the temperature and precipitation from 2006 to 2100. For this, GFDL-CM3 Model outputs from 2006 to 2100 under RCP 6.0 scenario were used as inputs to the calibrated model and corresponding temperature and precipitation were forecasted. These forecasted temperature and precipitation were plotted for different time periods under the three sets of categories namely 2020's (2020-29), 2050's (2050-59), 2080's (2080-89).

B. Using Artificial Neural Network (ANN)

With the help of same data used in MLR model, calibration, validation and testing of the model has been carried out. For calibration and validation of the model feed forward back propagation method was used. Levenberg-Marquardt ANN training algorithm was considered Different models have been prepared by varying the number of nodes in the hidden layers and finally the optimum was selected according to maximum R^2 and mean square error value as an evaluation criterion. An optimum model has been considered for the forecasting of the climate parameters.

VI. RESULTS AND DISCUSSION

The Coefficient of Determination values by both the techniques were found satisfactory for calibration and validation of the models. ANN showed better R square value against MLR. It is may be due to the fact that ANN can be able to capture the nonlinear relation between the predictors and predictand. It is given in Table 3. GFDL-CM3 model outputs were given as input to the calibrated optimum model and future temperature and precipitation were projected by both the models. These forecasted temperature and precipitation are represented in graphical form from fig.4 to fig. 9.
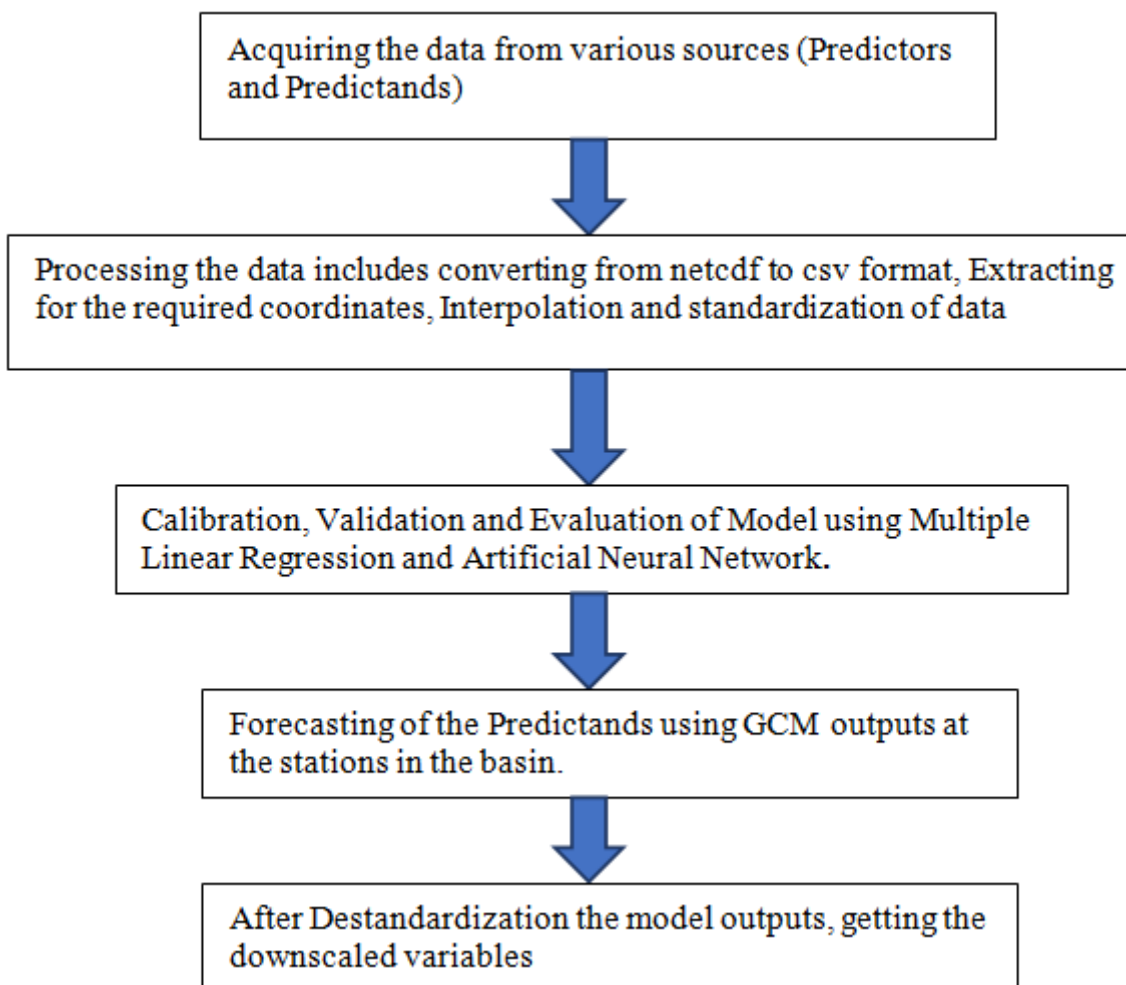
Fig 3. Flow Chart of the present methodology

Table 3. Comparison of R^2 Value

	Climate Parameter	CALIBRATION				VALIDATION			
		Station 1	Station 2	Station 3	Station 4	Station 1	Station 2	Station 3	Station 4
	Tmax	0.932	0.948	0.947	0.956	0.938	0.929	0.947	0.938
	Tmin	0.953	0.973	0.946	0.958	0.941	0.962	0.969	0.937
MLR	Rainfall	0.894	0.5	0.5	0.51	0.887	0.09	0.53	0.13
		CALIBRATION				VALIDATION			
		Station 1	Station 2	Station 3	Station 4	Station 1	Station 2	Station 3	Station 4
	Tmax	0.977	0.982	0.977	0.982	0.957	0.946	0.938	0.946
	Tmin	0.981	0.989	0.979	0.985	0.938	0.967	0.966	0.936
ANN	Rainfall	0.954	0.849	0.815	0.888	0.878	0.5	0.67	0.606

Tmax
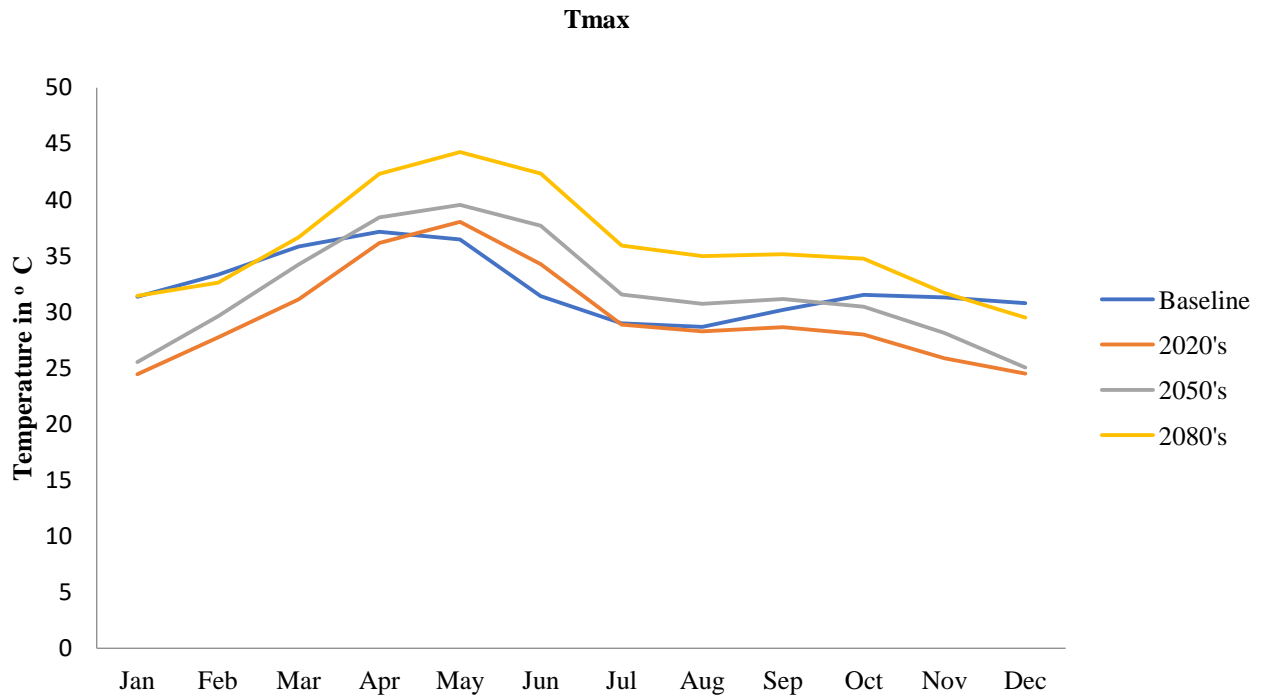
Fig 4. Max. Temperature at station 4 by MLR

Tmax

Fig 5. Max. Temperature at station 4 by ANN

Projection of Temperature and Precipitation using Multiple Linear Regression and Artificial Neural Network as a Downscaling Methodology for Upper Bhima Basin
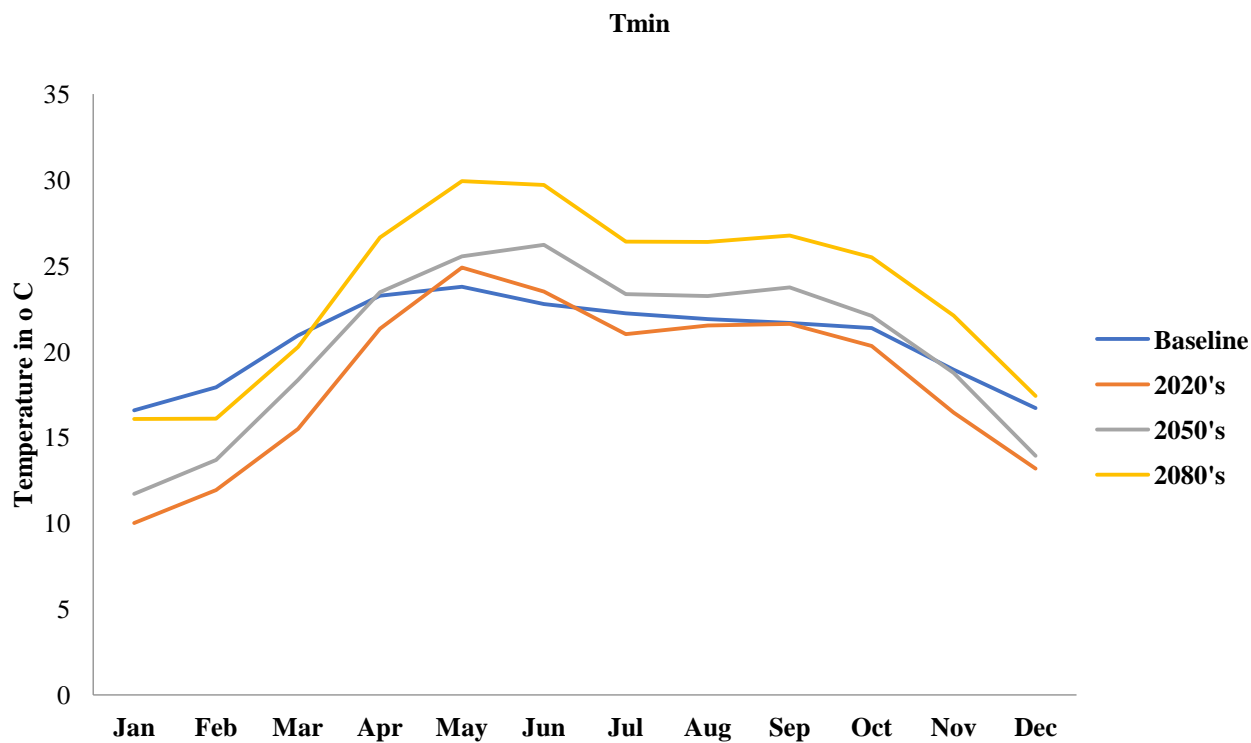
Tmin

Fig 6. Min. Temperature at station 4 by MLR
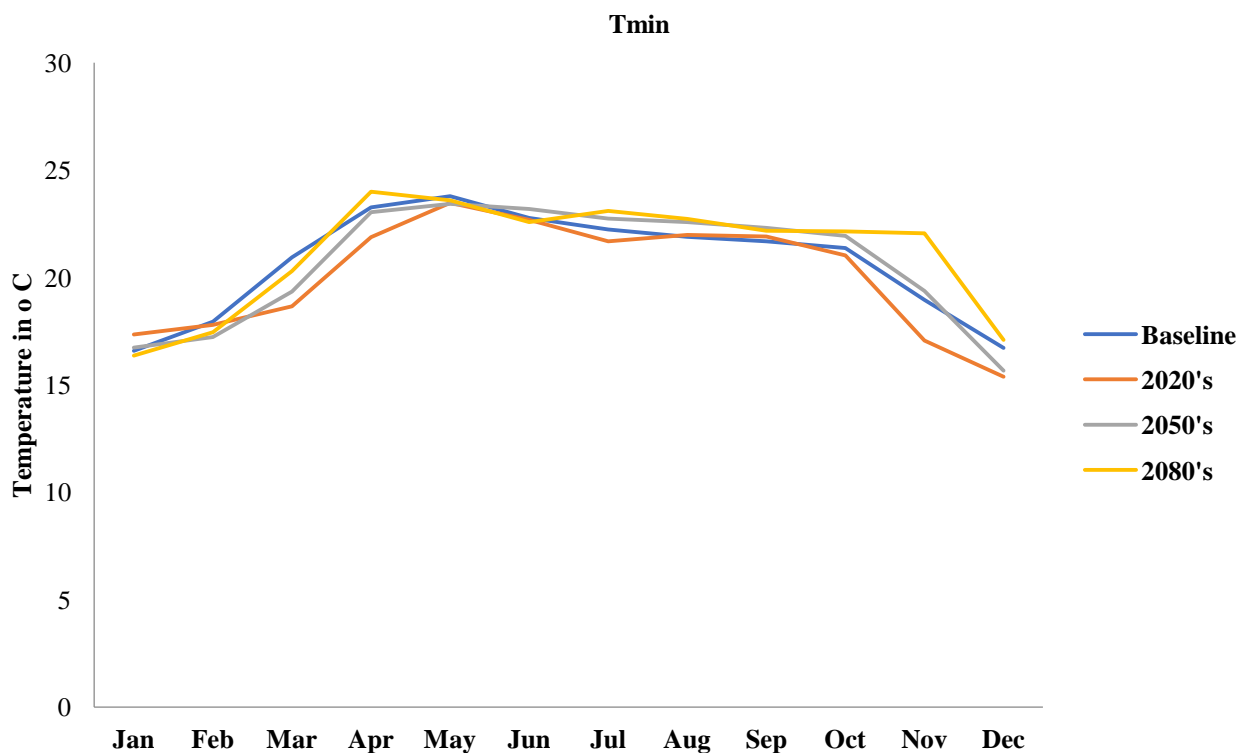
Tmin

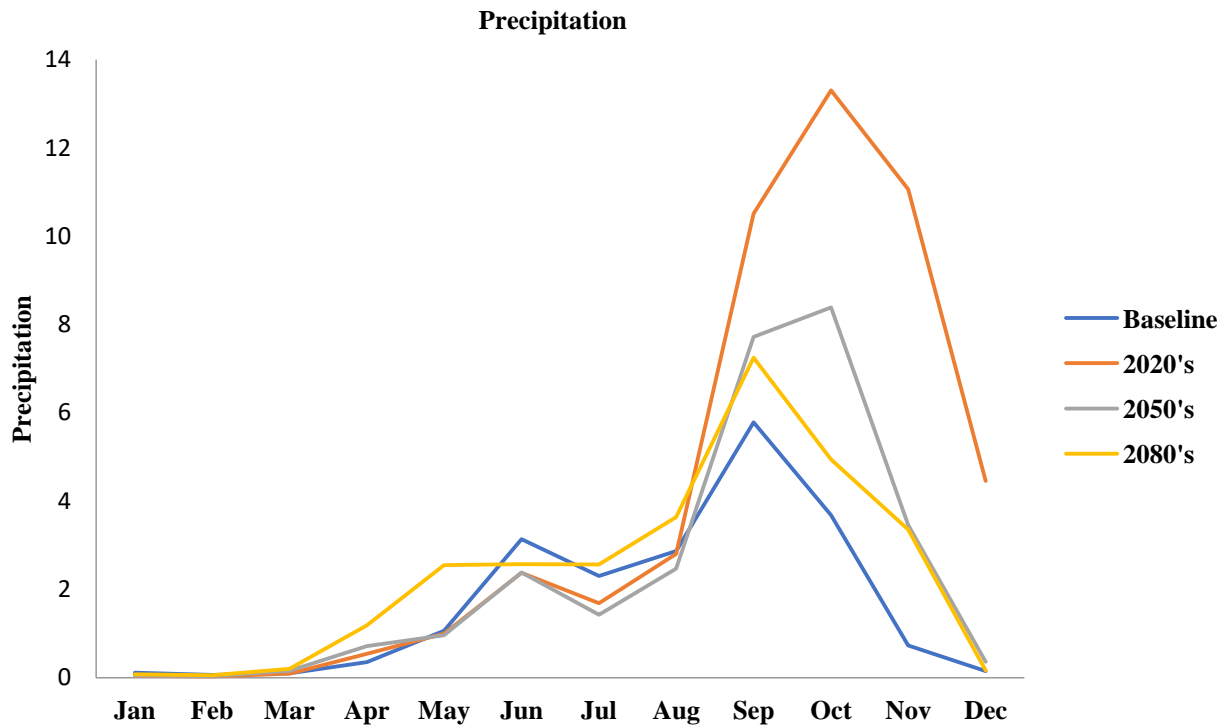
Fig 7. Min. Temperature at station 4 by ANN
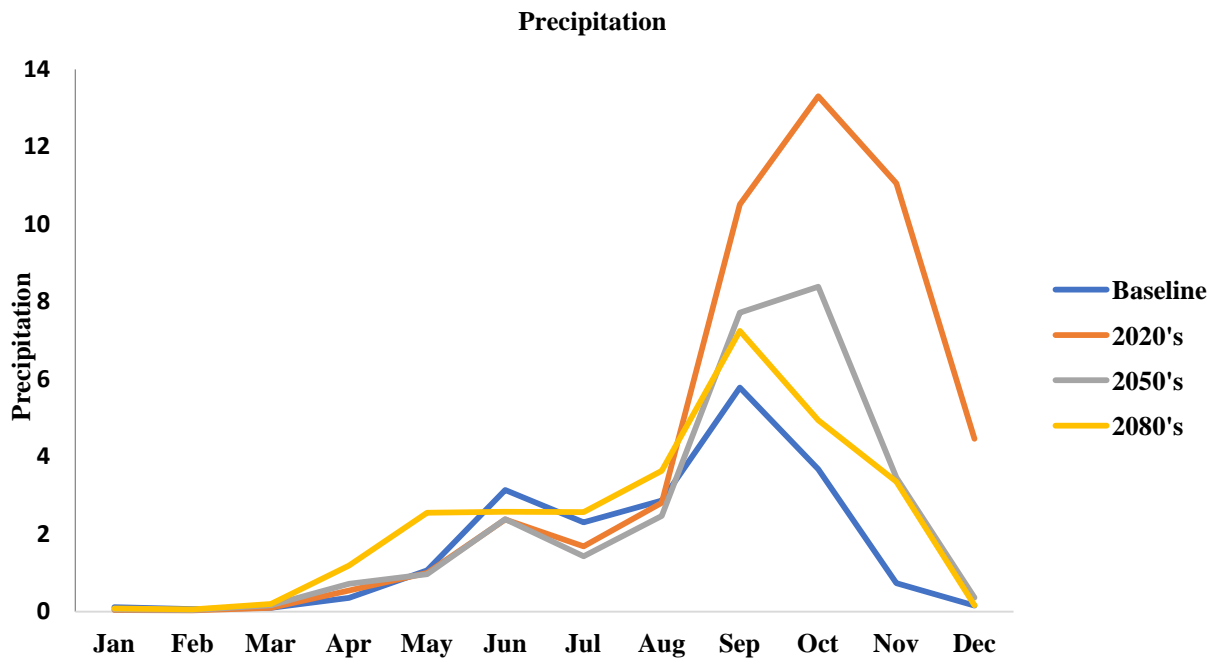
Fig 8. Precipitation at station 4 by MLR

Fig 9. Precipitation at station 4 by ANN

Results showed that maximum temperature increases in the monsoon season. As shown in fig 4 & 5, in the month of July an average maximum temperature increases from 28^0 C to 33^0 C. About 5^0 C rise in temperature in 2080's time period with respect to baseline period, results showed by the model. From the figure 6, increase in minimum temperature has been observed whereas fig.7 by ANN indicates that there is not significant change in the minimum temperature. It is also found form fig. 8 & 9 that there will be increase in precipitation during September to December while in monsoon it is almost similar to baseline period. The significant change in precipitation has been observed from 2050's to 2080's as compared to 2020's to 2050's. Peak point of precipitation is observed in September and October.

VII. CONCLUSION

The present paper investigated the applicability of Multiple Linear Regression and Artificial Neural Network. It is observed that there is a change in temperature, precipitation in the basin in future periods.

7

Even though the uncertainties associated with GCM and ANN models are not considered in the modeling, the model results showed changes in rainfall distribution pattern. Accordingly, it has to consider in reservoir operation system. Overall amount of temperature and hence precipitation has been found increased in future periods. This projected climate parameters can be used to study its influence on runoff and inflow to the reservoir.

Conflict of interest: On behalf of all the authors, the corresponding author states that there is no conflict of interest.

REFERENCES

1. Climate Change 2007: Synthesis Report (IPCC)
2. IPCC Report 2014.
3. B. C. Hewitson and R.G. Crane, "Climate downscaling: techniques and application", Clim. Res., vol.7, (1996), 85–95.
4. R.L. Wilby, M.L. Wigley, D. Conway, P.D. Jones, B.C. Hewitson, J. Main and D.S. Wilks, "Statistical downscaling of general circulation model output: a comparison of methods", Water Resour. Res., 34,(1998), 2995–3008.
5. A. J. Askew, "The influence of climate change and climate variability on the hydrologic regime and water resources", IAHS Publication,Vol.168, (1987), pp. 421-430.
6. E.P. Salathe, Jr., P.W. Mote and M.W. Wiley, "Review of scenario selection and downscaling methods for the assessment of climate change impacts on hydrology in the United States Pacific Northwest", Int. J. Climatol., Vol. 27, (2007), 1611–1621.
7. Rao S. Govindaraju, "Artificial neural network in hydrology I: Preliminary concepts", Journal of Hydrologic Engineering, Vol. 5, April 2000, 115-123
8. Ahmed Amara Konate Heping Pan, Nasir Khan and Jie Huai Yang, " Generalized regression and feed forward back propagation neuralnetwork in modelling porosity from Geophysical well logs", Journal of Petroleum exploration and production Technology, Vol.5, (2015), 157-166.
9. David E. Rumelhart, Geoffrey E. Hintont & Ronald J. WilliamsJ. Wang, "Learning representations by back-propagating errors", Nature Publishing Group, Vol.323 (9), 1986, pp 533-536.
10. Government of Maharashtra Hydrology Project II, Water Resources Department, "Real Time Streamflow Forecasting and Reservoir Operation System for Krishna and Bhima River Basins in Maharashtra" (RTSF & ROS), Inception Report, December 2011,
11. Subimal Ghosh, P.P. Mujumdar, "Statistical Downscaling of GCM simulations to streamflow using relevance vector machine", Advances in Water Resources, Vol 31, July 2007, pp. 132-146
12. K. Pearson, "Mathematical contributions to the theory of evolution III regression heredity and panmixia. Philos". Trans. R Soc. Lond. Ser., 187, (1896), 253–318.

AUTHORS PROFILE

Dattatray Rajmane has been working as Asst. Professor in Sinhgad College of Engineering, Pune since 1996. He has completed his Master's degree in Hydraulics. Seven number of papers are on his credit. Presently He is pursuing Ph.D. at S.G.G.S.I.E.&T., Nanded. He is a member of ISTE, ISH.

Dr. Milind Waikar has been working as Professor in Civil Engineering in S.G.G.S.I.E.&T., Nanded, since 2001. He has about 34 years of Teaching Experience. He has completed M. Tech from IIT, Kharagpur and PH.D. from University of Roorki. 47 number of papers are on his credit and written 5 number of books.