Journal of Agrometeorology 18 (1) : 113-119 (June 2016)

Constructing the downscale precipitation using ANN model over the Kshipra river basin, Madhya Pradesh

P. K. MEENA[*], DEEPAK KHARE and M. K. NEMA[1]

Department of Water Resources and Management,Indian Institute of Technology, Roorkee

[1]Water Resources Systems Division, National Institute of Hydrology, Roorkee, India

[*]Email: pramodcae@gmail.com

ABSTRACT

The present study is focused on simulating the impact of climate change on the behavior of precipitationof Kshipra river basin in Madhya Pradesh, India.Artificial neural network (ANN) model was used to construct of the downscale precipitation scenario. A General Circulation Model (GCM) viz. Hadley Centre Coupled Model, version 3 (HadCM3), from Hadley Center, UK has been used for the study. In Model, monthly weather data on the basis of rapid economic growth under A1B scenario (A balanced emphasis on all energy sources) were considered. The four predictor variables which are used in ANN model formulation are screened from a set of 26 predictors based on correlation analysis of observed precipitation. The basic ANN architecture was optimized for training of the model byfirst selecting the training algorithm and then varying the number of neurons in the hidden layer. Twelve different training algorithms have been used. Further, the model was evaluated by varying the number of neurons from 1 to 30 in the hidden layer.The performance of modelwas evaluated in terms of the correlation coefficient (R), mean square error (MSE), root mean square error (RMSE) and mean absolute error (MAE). The results of model revealed that the predicted precipitation and observed precipitation are better correlated (R=0.911 and 0.853 during training and validation runs) with back propagation variable learning rate “traingdx” algorithm.

Key Words: Climate change; downscaling; precipitation; ANN

The changing climate has been a great threat to not onlymankind but also for every life on the planet. According to intergovernmental panel on climate change (IPCC) fifth assessment report on managing the risks of extreme events and disaster to advance climate change adaptation (SREX) addressed that the changes of drought patterns during second half of the 20th century are based on its attributed impact on precipitation and temperature changes. Representative concentration pathways (RCP8.5) scenario confirms the global average surface temperature increase from 1951 to 2010. The greenhouse gases (GHGs) concentrations due to anthropogenic activities (IPCC 2013) are responsible for such augmentation (Stocker *et al.,* 2013). Hydrology, agriculture and water resources are the major areas which are affected on the platform of climate change (Barnett *et al.,* 2005; Piao *et al.,* 2010; Shi *et al.,* 2013; Li *et al.,* 2014). Mujumdar (2008) has presented an overview of the current scenario and recent work in India to assess the climate change impact on water resources. Riyu *et al.,* (2006) suggested that the warm Atlantic Multi decadal Oscillation enhances the Indian monsoon rainfall by setting up a positive tropospheric temperature anomaly in late summer/ autumn and resultant delayed withdrawal of monsoon. Ghosh *et al.,*(2010) assessed climate change impact in the Mahanadi river basin using the probabilistic approach. In this study, uncertainty model has been developed by statistical downscaling with bias correction. Downscaling involves conversion of large scale GCM outputs of climate variables to local scale hydrologic variables. GCMs are generally coarse in size, these areused to simulate climatic variables such as wind speed, sea level pressure etc. (Ghosh and Mujumdar, 2008), but these GCMs are poor in performance while predicting precipitation because it is inherently nonlinear and extremely sensitive to physical processes (Stockdale *et al.,*1998). Therefore, to bridge the gap between climatic variables to local hydrological variables and to account for the inaccuracies in describing precipitation extremes, downscaling methods are commonly used in practice (Willems *et al.,* 2012).Different downscaling techniques might be more accurate for different seasons, regions, and time periods and depending on the input feature dataset (Dibike *et al.,* 2008). A relatively new branch of nonlinear techniques, artificial neural networks (ANN or NN), has been applied not only as artificial intelligence but also as general, non–parametric "regression" tool. A neural network consists of layers of highly interconnected processing units, each containing a small amount of local "memory." The network is trained using an iterative method to adjust the weights of connections between these units. The present study has been carried out for the Kshipra basin with the objective to develop an Artificial Neural Network model for downscaling precipitation from large scale GCM output.

MATERIALS AND METHODS

The present study has been carried out for the Kshipra basin which is located on the Malwa plateau in Western Madhya Pradesh (India) at an average altitude of 553 m above sea level. It originates at an elevation of about 560 meters from the hill near village Kampell (22° 31' N. and 76° E.) in the south east of Indore district and travels 195 km to meet river Chambal near village Kalahari (23° 53' N. And 75° 31' E.).Kshipra river basin is a southern tributary of Yamuna river basin which is second largest river basin of India. The studied basin has a catchment area of 5608 km^2. It is a seasonal river with plenty of water during the monsoon months June, July, August and September (JJAS), but the discharge goes on decreasing after monsoon and reduces to a trickle during the summers. Over the years the river has lost its perennial nature and now runs dry for a period of 5 to 6 months per year (NWM, 2011). The water of the Kshipra River is used for drinking, industrial and irrigation purposes.

The monthly average precipitation data collected from India Meteorological Department (IMD), Pune, India for the periods 1981 to 2010 has been used as the response variable. GCM Model used in this study is a coupled atmosphere-ocean general circulation model developed at the Hadley Centre, UK (i.e. HadCM3) with a horizontal resolution of 2.5° latitude and 3.75° longitude. The large-scale monthly predictors of GCM for A1B future scenarios (Rapid Economic Growth, a balanced emphasis on all energy sources) for 60 years (1981–2040) was obtained from the Canadian Climate Impacts Scenarios (CCIS) website. Among thevarious scenarios ofSpecial Report on Emissions Scenarios (SRES),the A1Bscenarios with monthly predictor variables, have been used for the ANN modeling. As the spatial scale of National Center of Environmental Prediction (NCEP) grid points and GCM grid points are not the same; therefore interpolation has been performed for the processing of the data. Here, two dimensional linear interpolationinverse distance method (IDW) by MATLAB programming have been used (Salvi *et al.,* 2011).

Selection of predictors

There have been many techniques used for selection of useful predictors in downscaling in terms of their performance (predictive power) with real data by many researchers worldwide.Bergant and Kajfez–Bogataj, 2005; Shongwe et al., 2006; Benestad et al., 2007; suggested that predictors should be selected using the following criteria: (a) the large-scale predictors should be physically relevant to the local-scale features and realistically simulated by GCMs, (b) the predictors are readily available from the archives of GCM output and reanalysis datasets, and (c) strongly correlated with the predicted.

In the present study,the predictor selection is carried out based on (Principle Component Analysis) the correlation between predictand (precipitation) and predictor (NCEP) (Meena *et al.,* 2014). The study has been carried out using four predictor variables which are screened from a set of 26 predictors mentioned in Table 1 (a). The variables shown in Table 1(b), are later used as input to the ANN model.

Development of ANN model has been followed in two stages. First training mode (1981 to 1998) and second validation phase (1999 to 2010). In training mode, the output links to as many of the input nodes as desired and pattern is defined. The network is adjusted according to mean square error. The validation dataset is used at this stage to ensure that the model did not over trained. The most useful neural network in function approximation is multilayer perception (MLP). It consists of an input layer, hidden layer(s) and output layer. During the training phase, the weights and biases of the network is optimized using an optimization algorithm (Meena *et al.,* 2014). Back Propagation (BP) networks are the most widely used ANN models. In fact the name back-propagation comes from the error term, which is propagated back through the network during learning and used to change the weights on the equation. The weights are changed using the following equation.

$$\Delta w_{ij} = -\eta \frac{\delta E}{\delta w_{ij}} + m\Delta w_{ij}(n-1) \qquad (1)$$

In this equation η and m are known as learning rate and momentum coefficient respectively.These input signals get added up, and are fed into the activation function. The reaction signals of the neuron would then pass through a transfer function, which decides the strength of the out signal (Parida, 2012). Finally, the output signal is sent

Table 1(a): Name and description of all NCEP and GCM predictors

Sl No.	Atmospheric pressure level	Variables	Name	Unit
A.	1013.25 hPa (1)	Sea level Pressure	SAT	Pa
B.	1000 hPa (6)	Wind speed	p_f	ms^{-1}
		Zonal velocity	p_u	ms^{-1}
		Meridional velocity	p_v	ms^{-1}
		Vorticity	p_z	s^{-1}
		Wind direction	p_th	Degree
		Divergence	p_zh	s^{-1}
C	850 hPa (8)	Wind speed	p8_f	ms^{-1}
		Zonal velocity	p8_u	ms^{-1}
		Meridional velocity	p8_v	ms^{-1}
		Vorticity	p8_z	s^{-1}
		Wind direction	p8_th	Degree
		Divergence	p8_zh	s^{-1}
		Geopotential height	p850	M
		Relative humidity	r850	%
D	500 hPa (8)	Wind speed	MSW	ms^{-1}
		Zonal velocity	p5_u	ms^{-1}
		Meridional velocity	p5_v	ms^{-1}
		Vorticity	p5_z	s^{-1}
		Wind direction	p5_th	
		Divergence	p5_zh	s^{-1}
		Geopotential height	p500	M
		Relative humidity	RH	%
E	Near surface (3)	Specific humidity	shum	g/kg
		Surface air temperature	SAT	^0C
		Relative humidity	rhum	%

Table 1(b). Input variable for ANN modeling and Correlation coefficient of predictor for using the HadCM3 model

Model	Predictors
R-ANN-4	SAT, RH, SLP, MSW
Parameter	**Correlation coefficient with rainfall**
Surface air temperature (SAT)	0.271
Relative humidity @ 500 hpa (RH)	0.79
Sea level pressure (SLP)	-0.618
Meridional surface wind speed (MSW)	0.596

through all the output connections to other neurons

$$y_j = \int \{W_j \times X_i\} - \theta_j \quad (2)$$

The function f (x) is called as an activation function, the activation function enables a networkto map any non-linear process. The most commonly used function is the sigmoidal function expressed as:

$$f(x) = \frac{1}{1+e^{(-x)}} \quad (3)$$

The variables were selected according to the model R-ANN-4 for developing and evaluating the ANN models. The ANN model architecture is a single layer feed forward network, which is one of the simplest neural network and has

Table 2: Performance of ANN models based on various algorithms (training with 60% of the data).

Alog.	Model	R		MSE		RMSE		MAE	
		Trg	Val	Trg	Val	Trg	Val	Trg	Val
Trainlm	R-ANN-4	0.915	0.832	0.146	0.346	48.763	69.114	30.832	36.088
Traingd	R-ANN-4	0.865	0.826	0.269	0.301	60.28	73.987	38.406	38.219
Traingdm	R-ANN-4	0.862	0.822	0.274	0.311	60.721	73.241	37.408	38.722
Traingda	R-ANN-4	0.869	0.833	0.247	0.294	59.242	72.182	36.994	39.419
Traingdx	R-ANN-4	0.905	0.853	0.194	0.257	51.822	66.394	30.077	32.727
Traincgf	R-ANN-4	0.851	0.811	0.268	0.324	63.022	76.498	39.571	40.552
Traincgp	R-ANN-4	0.907	0.842	0.179	0.271	51.048	67.448	31.846	34.439
Traincgb	R-ANN-4	0.899	0.857	0.189	0.26	52.682	64.923	32.615	34.147
Trainscg	R-ANN-4	0.902	0.865	0.192	0.244	51.889	63.687	31.271	33.326
Trainbfg	R-ANN-4	0.91	0.852	0.175	0.257	49.754	66.895	30.356	33.895
Trainoss	R-ANN-4	0.881	0.847	0.227	0.275	57.212	68.476	36.113	36.366
Trainrp	R-ANN-4	0.904	0.854	0.191	0.26	51.317	66.25	31.03	33.446

(Trg-Training, Val- Validation)

Table 3: Performance of ANN models based on different number of neuron (training with 60% of the data)

Alog.	Model	R		MSE		RMSE		MAE	
		Trg	Val	Trg	Val	Trg	Val	Trg	Val
i	**N1**	0.835	0.797	0.321	0.347	67.313	79.822	44.087	42.779
ii	**N2**	0.862	0.821	0.266	0.31	60.702	74.213	37.043	37.809
iii	**N3**	0.897	0.86	0.214	0.258	53.136	64.478	32.911	34.042
iv	**N5**	0.852	0.799	0.282	0.339	62.941	79.247	40.064	41.192
v	**N7**	0.905	0.853	0.194	0.257	51.822	66.394	30.077	32.727
vi	**N9**	0.806	0.765	0.386	0.4	71.025	84.791	44.595	42.739
vii	**N11**	0.817	0.764	0.43	0.488	70.958	88.328	51.138	51.913
viii	**N13**	0.895	0.821	0.212	0.319	53.429	73.821	32.879	37.104
ix	**N14**	0.904	0.826	0.199	0.318	51.407	70.805	32.495	37.395
x	**N17**	0.909	0.84	0.18	0.287	50.129	68.136	31.106	35.713
xi	**N19**	0.917	0.831	0.164	0.299	47.757	69.584	29.156	36.11
xii	**N21**	0.818	0.75	0.345	0.438	68.888	85.405	41.898	43.676
xiii	**N23**	**0.911**	**0.853**	**0.176**	**0.262**	**49.439**	**66.734**	**30.551**	**35.438**
xiv	**N25**	0.906	0.833	0.186	0.297	50.882	71.078	32.32	38.579
xv	**N27**	0.91	0.85	0.189	0.269	49.913	66.881	29.652	33.026
xvi	**N30**	0.767	0.726	0.657	0.735	84.656	98.781	53.409	57.203

(Trg-Training, Val- Validation)

been successfully used as criteria for both model development and evaluation (Govindaraju, 2000; Maier and Dandy 2000).The number of hidden layers is one. The transfer function from input to hidden layer is Tan-Sigmoid Transfer Function (Tansig) and from hidden layer to output layer is Linear Transfer function (Purelin). The performance function used for training and testing of networks is MSE (Mean Squared Error.

Performance parameters

Correlation coefficient, mean square error (MSE), root mean square error (RMSE) and mean absolute error (MAE) were used to evaluate the performance of the model.

RESULTS AND DISCUSSION

ANN models were developed for prediction of future precipitation of the Kshipra River basin using future climatic variables obtained from the GCM simulation. The "mapstd" function available in MATLAB was used for scaling all input and target data for zero mean and standard deviation of one. In this study, HadCM3 – GCM model under A1B Scenario was applied for providing the input parameters to ANN

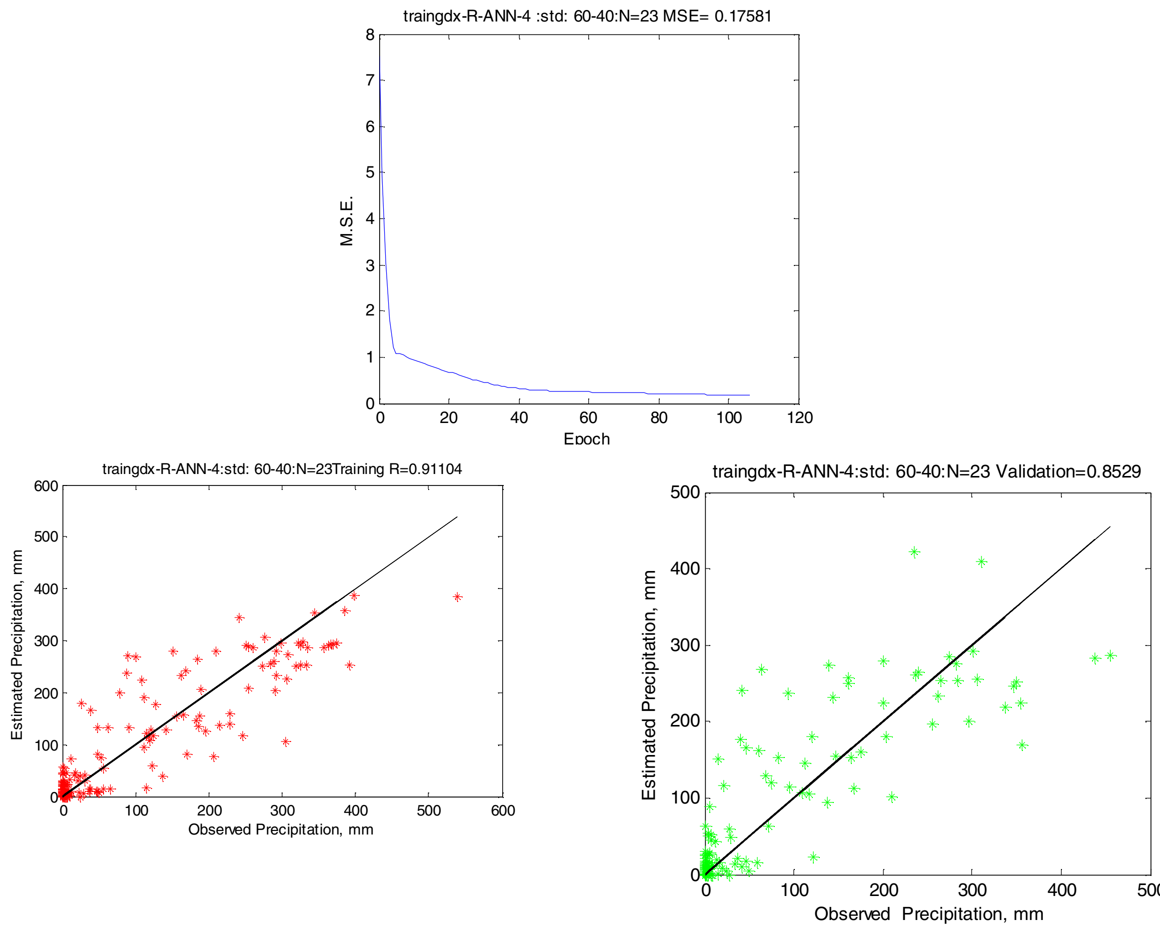

Fig. 1: Performance of "traingdx" training algorithm and 23 neurons with 60% length of data set used for training and the rest of that used for validation, ANN modeling (precipitation).

model based downscaling method.

Development of ANN Model

In this section, comparative study is carried out to find out the optimal number of hidden neurons required for best performance of the model as well as to find out the best algorithm for training the ANN model for large scale GCM. Twelve different algorithms are used and an attempt is made to choose the best algorithm for training the ANN model (Meena *et al.,* 2014).

Training algorithms for ANN based mean monthly precipitation model of Kshipra basin

For developing the ANN based monthly precipitation prediction model, performance of 12 training algorithms were evaluated. The model "R-ANN-4" was developed using Levenverg Marquardt Algorithm (trainlm). The best training algorithm in the hidden layer of ANN model can be determined by trial and error, at which a model perform better. In "traingdx" training algorithms the MSE of scale output and target is 0.194 and 0.257 during training and validation, respectively. Coefficient of correlation between estimated precipitation and observed precipitation is 0.905 and 0.853 during training and validation, respectively. RMSE has been worked out as 51.822 and 66.394 and MAE as 30.077 and 32.727 during training and validation respectively. In comparison of performance parameters presented in Table 2, it can be stated that model "R-ANN-4" trained with "traingdx" algorithm, "mapstd" normalization function performed best. This network architecture has been further improved varying neuron in the hidden layer.

Selection of optimum number of neurons in the hidden layer for ANN based mean monthly precipitation model for Kshipra basin

Increasing the number of neurons in the hidden layer, the network gets an over fit, that is the network have problem to generalize. To determine the optimum number of neurons, at which network should have to perform its best, trial and error method is applied. Selection of optimum number of neurons is an essential part of ANN model development. The model R-ANN-4 with training algorithm "traingdx" has been

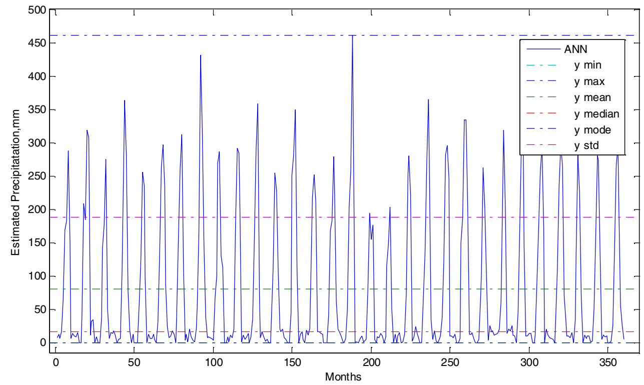

Fig. 2: Performance ANN modelfor precipitation in future (monthly) until 2040

evaluated for optimum number of neurons. Neurons in the hidden layer have been varied from 1 to 30. Performance of the model with 23 neurons has been depicted in Fig. 1. The model gets trained at 105 epochs with MSE 0.176.The correlation coefficient between estimated precipitation and observed precipitation is 0.911 and 0.853 during training and validation, respectively; RMSE and MAE during training is 49.439 and 30.551, respectively; whereas during validation, RMSE and MAE is 66.734 and 35.438 respectively. By the comparison of performance parameters presented in Table 3, it can be stated that model "R-ANN-4" trained with "trainsgdx" algorithm, and 23 neurons performs best.

Simulation of future monthly precipitation of the Kshipra river basin

The above analysis reveals that the ANN model for monthly precipitation, trained with "Traingdx" algorithm and hidden neurons of 23 performs better. This network architecture has been used for estimation the future precipitation of the Kshipra river basin. Fig. 2 shows the future monthly precipitation estimation (2011-2040) of the Kshipra river Basin. Precipitation in the Kshipra River basin has an increasing trend.

CONCLUSION

The present study was based on neural methods for site-specific precipitation for Kshipra river basin. The input variables from GCM data were used in study and the parameters considered monthly Surface Air Temperature, Relative Humidity @ 500 hpa, Sea Level Pressure, Meridional Wind Speed whereas, the response variable was monthly precipitation. Hence, the present attempts build a parsimonious (i.e. minimum number of parameters and more predictive power) ANN based model in neural network module of MATLAB and giving special concern to precipitation prediction.Beside this, the performance evaluation of the proposed model was also carried out by comparing observed and simulated precipitation.

Model performance has been evaluated in term of R, MSE, RMSE and MAE. ANN model R-ANN-4 with "traingdx" algorithm, 23 numbers of neurons, 60 percent and 40 percent length of record for training and validation was found with best predictive powers. The highest value of the coefficient of correlation between estimated and observed precipitation was found to be 0.911 and 0.853 during training and validation, respectively and MSE traced out 0.176 and 0.262 for the same.

The results are very encouraging and suggest the usefulness of neural network based modeling technique for downscaling of precipitation. ANN can be a useful tool for downscaling the various GCM products for hydrologic modelling and for planning and management of water resources thereafter.

REFERENCES

Barnett, T.P, Adam, J.C and Lettenmaier, D.P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, 438(7066):303-309.

Benestad, R. E, Hanssen-Bauer, I, and Førland, E. J. (2007). An evaluation of statistical models for downscaling precipitation and their ability to capture long-term trends. *Intern. J. Climatol.,* 27(5): 649–655.

Bergant, K, and Kajfez-Bogataj, L. (2005). N-PLS regression as empirical downscaling tool in climate change studies. *Theor. Applied Climatol*., 81: 11–23.

Dibike, Y, Gachon, P, St-Hilaire, A, Ouarda, T. B. M. J, and Nguyen, V. T. V. (2008). Uncertainty analysis of statistically downscaled temperature and precipitation regimes in Northern Canada. *Theor. Applied Climatol.,* 91(1): 149–170

Ghosh, S, and Misra, C. (2010). Assessing hydrological impacts of climate change: Modeling techniques and challenges. *Open Hydrol. J.*, 4(1):115-121

Ghosh, S, and Mujumdar, P. P. (2008). Statistical downscaling of GCM simulations to streamflow using relevance vector machine. *Advances Water Resou.*, 31(1):132-146.

Govindaraju, R. S. (2000). Artificial neural networks in hydrology. I: Preliminary concepts. *J. Hydrol.Engg.*, 5(2): 115-123

Li, Hongwei, andXiaoping, Yang., (2014). Temperate dryland vegetation changes under a warming climate and strong human intervention—with a particular reference to the district XilinGol, Inner Mongolia, China. *CATENA*, 119: 9-20.

Maier, H. R., Dandy, G. C. (2000). Neural networks for prediction and forecasting of water resources variables: a review of modeling issue and application. *Environ. Modeling Softw.*, 15, 101–124

Meena, P. K, Hardaha, M. K, Khare, D., and Mondal, A. (2014). Neural Method for Site-Specific Yield Prediction. In Proceedings of the Third International Conference on Soft Computing for Problem Solving, Springer India, 237-246

Mujumdar, P. P. (2008). Implications of climate change for sustainable water resources management in India" Physics and Chemistry of the Earth, Parts A/B/C, 33(5): 354-358

NWM (2011). TA 7417- IND: Support for the National Action Plan on Climate Change Support to the National Water Mission. Government of India.

Parida, P.K, (2012). Artificial Neural Network Based Numerical Solution of Ordinary Differential Equations (Doctoral dissertation, NIT, ROURKELA).

Piao, S, Ciais, P, Huang, Y, Shen, Z, Peng, S, Li, J, and Zhou, L, (2010). The impacts of climate change on water resources and agriculture in China. *Nature*, 467(7311) : 43-51.

Riyu, Lu, Dong, B, and Ding, H, (2006). Impact of the Atlantic Multi-decadal Oscillation on the Asian summer monsoon", *Geophy. Res. Letters*, 33; (L24701), doi10.1029/2006GL027655.

Salvi, K, Kannan, S, Ghosh, S, (2011). Statistical downscaling and bias-correction for projections of Indian rainfall and temperature in climate change studies. In 4th International Conference on Environmental and Computer Science, Chicago. pp. 7-11.

Shongwe, E. M, Landman, W. A, and Mason. S. J. (2006). Performance of recalibration systems for GCM forecasts for Southern Africa" *Intern. J. Climatol.,* 26: 1567–1585

Stockdale, T, Anderson, D,Alves, J. (1998). Global seasonal rainfall forecasts using a coupled ocean–atmosphere model, *Nature*, 392(3): 1996-1999

Stocker, T.F, Dahe, Q, Plattner, G,K, (2013). Climate Change 2013: The Physical Science Basis. Working Group I Contribution to the Fifth Assessment Report of IPCC: Summary for Policymakers.

Willems, P, Nielsen, K. A, Olsson, J, and Nguyen V. T. V. (2012). Climate change impact assessment on urban rainfall extremes and urban drainage: Methods and shortcomings" *Atmosph. Res.*, (103): 106-118.

Received : January 2016 ; Accepted : April 2016