Symposium on Hydrology, Indian Association of Hydrologists (IAH)
22-23 December 2015, New Delhi, India

DEVELOPMENT OF NEURAL MODEL FOR PRECIPITATION CONSIDERING CLIMATE CHANGE UNDER RAPID ECONOMIC GROWTH

PRAMOD KUMAR MEENA[#1], DEEPAK KHARE[2] AND MANISH K NEMA[3]

[#1]Research Scholar, WRDM, IIT Roorkee, Email - pramodcae@gmail.com
[2]Professor & Head, WRDM, IIT Roorkee
[3] Scientist, WRSD, NIH Roorkee

ABSTRACT

Reliable forecasting of future precipitation influenced by climate change scenario is an important field of research. The present study is focused on simulating the impact of climate change on precipitation behavior of Kshipra River, Madhya Pradesh, India which is a southern tributary of Yamuna River basin. CGMs are the most reliable sources available for the future climate, downscaling involves conversion of large scale GCM outputs of climate variables to local scale hydrologic variables. An attempt has been made to downscale the GCMs using Artificial Neural Network (ANN).The developed model used HadCM3 monthly weather data under A1B scenario (Rapid Economic Growth, A balanced emphasis on all energy sources) to determine the monthly precipitation at a specific site. Model performance has been evaluated in terms of coefficient of correlation (R), Mean square Error (MSE), Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). It was observed that the ANN model predicted precipitation dependably correlated with observed precipitation. (R=0.911 and 0.853 during training and validation runs)

Key Words: *Climate Change; Downscaling; Precipitation; ANN*

1. INTRODUCTION

The changing climate has been a great threat to not only mankind but also for every life on the planet. This is going to play a major governing role in different aspect of the various cycles like hydrologic cycle, food chain cycle, geologic cycle, etc. Climate change occurs at both global and regional scales. The average global near surface air temperature has increased by 0.3^0C to 0.5^0C in the last century due to contribution of greenhouse gases and other human activities. There is an alarming increasing trend of temperature. The Intergovernmental Panel on Climate Change (IPCC) produced its fourth assessment report in 2007 and projected that there will be a 1.4 ^{0}C to 5.8 ^{0}C increases in globally averaged surface temperature during 1990 to 2100. 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, it used 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 to artificial intelligence but also as general, non–parametric “regression” tools. 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 is located on the Malwa plateau in Western Madhya Pradesh. To establish the true merits of ANNs forecasting skill with the specific objective; to development of an Artificial Neural Network model for predicting precipitation from large scale GCM output.

2. STUDY AREA AND DATA

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 originated 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 the 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 basin which has plenty of water during the monsoon months (June- October), 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 (HadCM3) with a horizontal resolution of 2.5° of latitude by 3.75° of longitude The large-scale monthly predictors of (HadCM3) for HadCM3 A1B (Rapid Economic Growth, A balanced emphasis on all energy sources) future scenarios for 60 years (1981–2040) obtain from the Canadian Climate Impacts Scenarios (CCIS) website. Among the SRES scenarios A1B, provides monthly predictor variables, which can be exclusively used for the ANN model. NCEP grid points and GCM grid points vary; therefore interpolation is needed for processing of the data. Here two dimension linear interpolations by MATLAB programming have been used.

3. MATERIALS AND METHODS

3.1 Selection of predictors

There have many techniques used for selection of useful predictors in downscaling in terms of their performance (predictive power) with real data by many researchers worldwide [Benestad et al., 2007; Shongwe et al., 2006; Tripathi et al., 2006; Bergant and Kajfez–Bogataj, 2005] 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 predictors are selected by the stepwise regression method. Stepwise regression consists of two main approaches namely forward selection and backward elimination. Here the combination of the two approaches is used. The predictor selection is carried out by the automatic procedure which includes simple correlation and partial correlation. The process will continue until it reaches the best f-test, t-test, adjusted R^2 (co-efficient of determination), Akaike information criterion (AIC), Bayesian information criterion and Mallows'Cp .The study has been carried out using four predictor variables that are screened from a set of 26 predictors. The variables are shown in Table 1, which later used as an input to ANN model.

Table 1. 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

3.2 Design of 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 this 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). The most popular algorithm is the error backpropagation neural network (BPNN) model which has been used in this study. Back Propagation (BP) networks are the most widely used ANN models. They are a gradient descent technique which minimizes the network error function. 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.

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

In this equation η and m are known as learning rate and momentum coefficient respectively. The functional diagram of an artificial neuron is shown in Fig.1 there are weighted input connections to the artificial neurons. 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 decided the strength of the out signal (Sablani, 2010; Parida, 2012). Finally, the output signal is sent through all the output connections to other neurons

$$y_j = \int \{W_j \times X_i\} - \theta_j \tag{2}$$

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

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

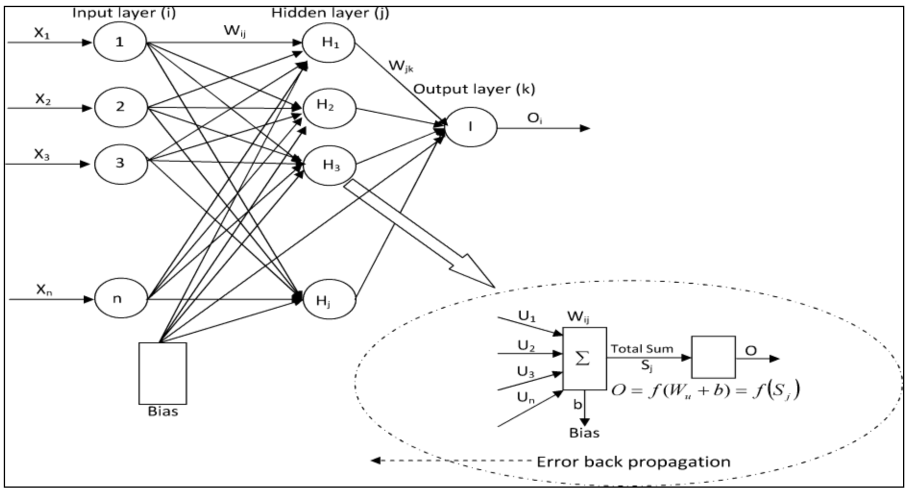

Fig.1: An artificial neuron showing its function

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 been successfully used the prediction of the nonlinear process (Maier and Dandy 2000; Govindaraju, 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 of training and testing of networks used are MSE (Mean Squared Error). The various combinations of hidden nodes and training function were done to arrive at optimum combinations to give less error (Hardaha et al., 2013, Meena et al., 2014). The network iterations (Epochs) were kept at 500. The neural network utility file is edited in MATLAB (7.6 Version). The input variable selection, input data source file, network option, training function, setting for the data for training, validation, plotting the predicting values and saving the network are created and run in MATLAB software.

3.3 Evaluation Indicators

Correlation coefficient, Mean square error (MSE), Root mean square error (RMSE) and Mean Absolute error (MAE) were used as the model develops parameters as well as the criteria for evaluation.

4. 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 applied for providing the input parameters to ANN model based downscaling method.

4.1 Development of ANN Model

The GCM models are HadCM3 a 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 GCM model. Twelve different algorithms are trying to choose the best algorithm for training the ANN model (Meena et al., 2014).

4.2 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.

Table 2. Performance of various algorithms ANN models (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)

4.3 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 essential part ANN model development. The model R-ANN-4 with learning function "traingdx" and normalization function "mapstd" trained with 60 percent of data has been 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. 2. 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. In comparison of performance parameters presented in Table 3, it can be stated that model "R-ANN-4" trained with "trainsgdx" algorithm, "mapstd" normalization function and 23 neurons performed best.

Table 3: Performance of neural network with different number of neutrons for ANN based monthly precipitation modeling

S.N	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)

4.4 Simulation of future monthly precipitation of the Kshipra river basin

The above analysis reveals that the ANN model for monthly precipitation estimation trained with "Traingdx" algorithm and hidden neurons of 23 with 60% length of data set used for the training and 40% of data set used for the validation was found better than other ANN model. This network architecture has been used for estimation the future precipitation of the Kshipra river basin. Fig. 3

shows the future monthly precipitation estimation (2011-2040) of the Kshipra river Basin. Precipitation in the Kshipra River basin has an increasing trend.

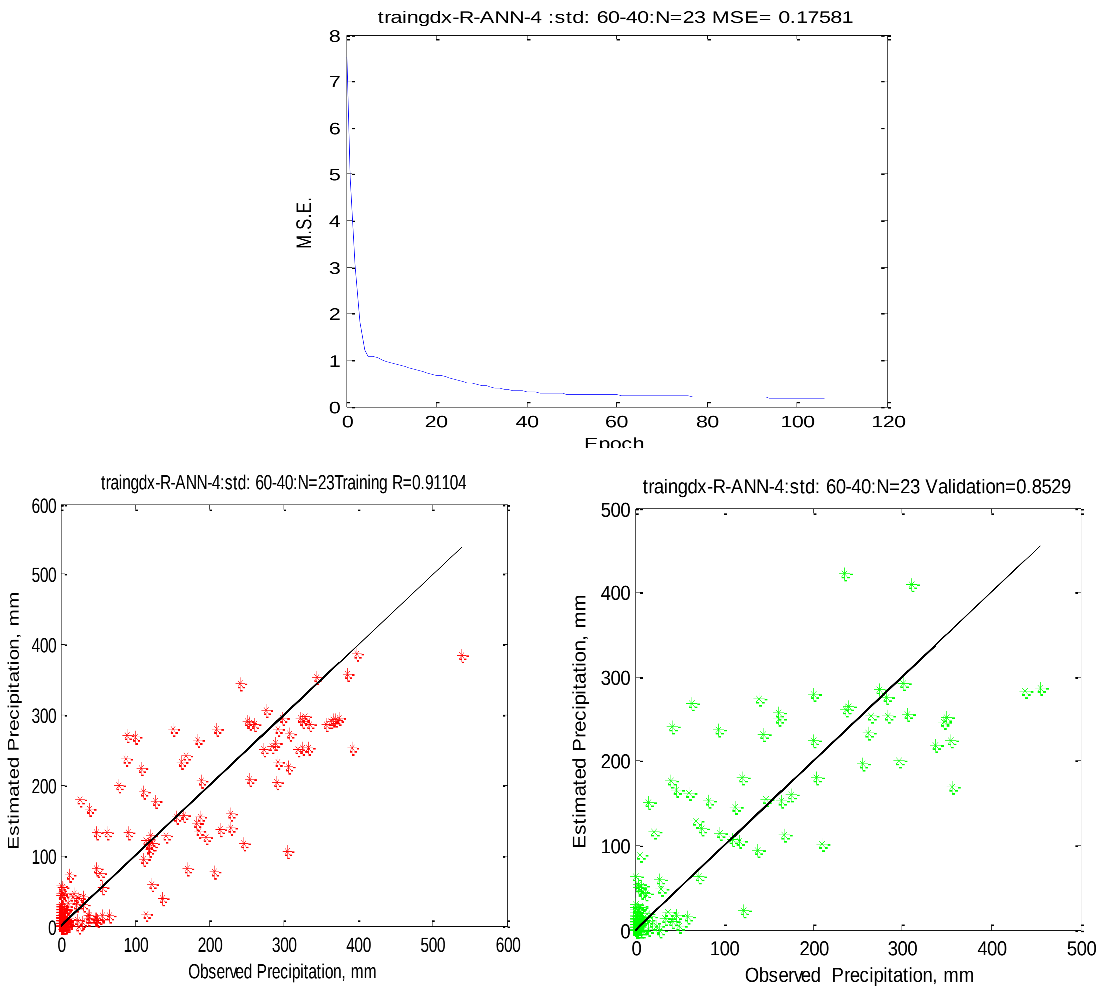

Fig. 2 Performance of "traingdx" training algorithm with 23 neurons with 60% length of data set used for training and the rest of that used for validation ANN modeling (precipitation).

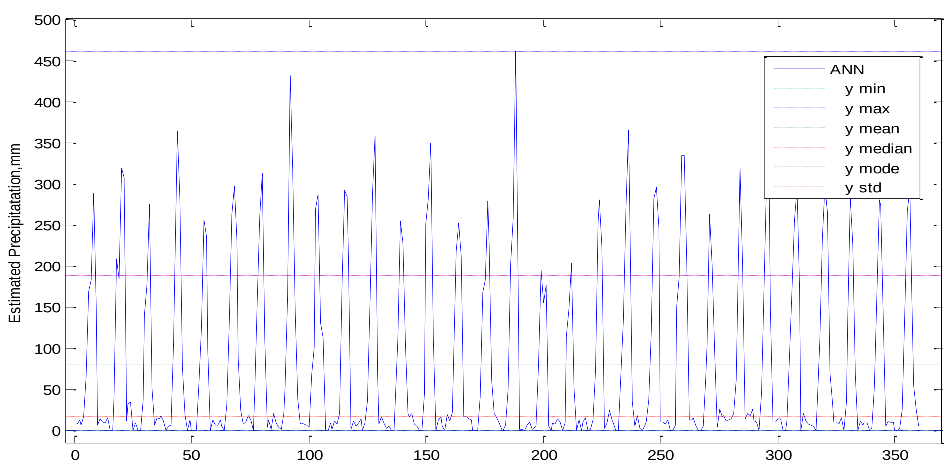

Fig. 3 Performance ANN model for precipitation in future (monthly) until 2040

5. CONCLUSION

Neural based adaptive climate change prediction for precipitation was undertaken for the Kshipra river basin. A Back Propagated Neural Network (BPNN) has been developed for precipitation Modelling. The predictor variable were monthly Surface Air Temperature, Relative Humidity @ 500 hpa, Sea Level Pressure, Meridional Wind Speed whereas, the response variable was monthly precipitation. An Attempt was made to developed parsimonious ANN model in neural network module of MATLAB. Model performance has been evaluated in term of R, MSE, RMSE and MAE. ANN model 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.

6. REFERENCES

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

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". International Journal of Climatology, *27(5): 649–655.*

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", Theoretical and Applied Climatology, *91(1): 149–170*

Govindaraju, R. S. (2000) "Artificial neural networks in hydrology. I: Preliminary concepts", Journal of Hydrologic Engineering, *5(2): 115-123*

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

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

Hardaha, M. K., Chouhan, S. S., and Ambast, S.K. (2013). "Application of Artificial Neural Network in Predicting Farmers' Response to Water Management Decisions on Wheat Yield" J. of Agricultural Engineering, *49 (3): 32 – 40*

IPCC: Summary for policymakers. In Parry Canziani, M., Palutik, O., linden, J. V., and Handson, C. (2007) "Climate change (2007). Impacts adaptation & vulnerability. Contribution of working group to fourth Assessment Report of the Intergovernmental panel on Climate Change II", Cambridge: Cambridge University Press, *7-22.*

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 Governments of Punjab, Madhya Pradesh Tamil Nadu and Asian Development Bank

Parida, P. K. (2012) “Artificial Neural Network Based Numerical Solution of Ordinary Differential Equations”

Riyu, Lu, Dong, B., and Ding, H. (2006) “Impact of the Atlantic Multi-decadal Oscillation on the Asian summer monsoon”, Geophysical Research Letters, Vol. 33, L24701, doi10.1029/2006GL027655.

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

Sablani, S. S. (2010) “Artificial Neural Network Modeling”, Handbook of Food and Bioprocess Modeling Techniques, 375

Shongwe, E. M., Landman, W. A., and Mason. S. J. (2006) “Performance of recalibration systems for GCM forecasts for Southern Africa” International Journal of Climatology *26: 1567–1585*

Tripathi, S., Srinivas, V. V., and Nanjundiah, R. S. (2006) “Downscaling of precipitation for climate change scenarios: a support vector machine approach”, Journal of Hydrology, DOI: 10.1016/j.jhydrol.2006.04.030.

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” Atmospheric Research, *(103): 106-118*