

Finer Scale Rainfall Projections for Kerala Meteorological Subdivision, India Based on Multivariate Empirical Mode Decomposition

S. Adarsh

Abstract—This study proposes an innovative approach for statistical downscaling of rainfall based on scaling property of meteorological variables. The reanalysis data of five dominant meteorological variables mean sea level pressure, relative humidity, surface temperature, wind velocity (zonal and meridional components) extracted from National Centre for Environmental Prediction (NCEP) are used as predictors to project monthly rainfall of Kerala meteorological subdivision in India. The multiscale decomposition of predictor dataset of the region and the monthly rainfall of a specific grid point is performed simultaneously by employing the Multivariate Empirical Mode Decomposition (MEMD) technique. The individual modes are predicted by fitting stepwise linear regression (SLR) by considering the potential predictors based on p-value statistics. Subsequent addition of the predicted modes gives the monthly rainfall. The method is demonstrated by a specific grid point of Chalakkudi river basin in Kerala, India. The method is found to be superior over the linear regression and M5 model tree based transfer function approaches. Further, the MEMD-SLR hybrid model is used for rainfall projections of the state of Kerala under three representative concentration pathway scenarios (RCP2.6, RCP4.5 and RCP8.5) provided by Canadian Centre for Climate Modeling and Analysis (CCCMA).

Index Terms—Downscaling, Kerala, rainfall, MEMD.

I. INTRODUCTION

Modeling climate change impacts on hydrology has received considerable attention among the water resources community recently [1]–[8]. Transfer function based statistical downscaling is one of the most popular approaches for modeling hydrological impacts of climate change, owing to its simplicity. It is well known that most of the hydro-climatic time series data are possessing multiscaling behavior. This motivates the researchers to explore novel modeling practices which also include time scale information, to get accurate results in simulation and prediction of hydrological variables. Wavelet transform is one of the mathematical transformation methods which are used effectively as a data-pre processing tool in the modeling of hydrologic time series data. Hence the development of conjunction models based on discrete wavelet transform (DWT) is becoming popular for simulation and forecasting problems in hydrology [9], [10]. Adarsh and Reddy [11] applied a similar technique in the context of downscaling of rainfall of Orissa meteorological subdivision in India.

Most of the transfer function based downscaling approaches involves the step of identification of potential principal components (PCs) of hydrologic variables which often make the final model ‘black box’ in characteristics. i.e., in the final model the predictor set will be the PCs instead of the actual variables or its components. Moreover the accuracy of projections of hydrological variables by statistical downscaling may be constrained because of the non-stationary behavior of climatic data. It is believed that such limitations can be overcome to certain extent by carrying out the downscaling considering the real parameters instead of PCs and by following a data pre-processing based transfer function. Addressing the above issues, this paper presents a novel modeling strategy employing the multiscale decomposition of predictor (meteorological) variables employing the Multivariate Empirical Mode Decomposition (MEMD) technique. The specific objectives of the paper include : (i) to propose MEMD based transfer function model for location specific rainfall predictions in Kerala meteorological subdivision, India; (ii) compare the performance of MEMD based method with that of M5 Model Tree (MT) and multiple linear regression (MLR); (iii) apply the MEMD based model for location specific rainfall projections of the subdivision, and to draw broad inferences on future rainfall, under different scenarios of a specific General Circulation Model (GCM).

II. METHODOLOGY

Identification of time scale of variability is a challenging task in non-linear time series analysis. Huang *et al.*, [12] proposed a purely ‘data adaptive’ decomposition procedure namely Empirical Mode Decomposition (EMD) to analyse the non-linear and non-stationary time series. Unlike the case of DWT, it doesn’t demand ‘*a priori*’ fixing of decomposition levels and basis function type and being data adaptive, it is unique and intuitive. Multivariate EMD proposed by Rehman and Mandic [13] is an extension of the traditional EMD, which decomposes multiple time series simultaneously after identifying the common scales inherent in different time series of concern.

The present study uses MEMD for the decomposition of the datasets and uses the stepwise linear regression (SLR) for building the regression models for each of the components. The proposed methodology involves the following steps:

- 1) Decomposition of rainfall and predictor variables using MEMD to get different orthogonal oscillatory modes called intrinsic mode functions (IMFs), each of specific

Manuscript received March 1, 2016; revised May 10, 2016.

S. Adarsh is with TKM College of Engineering Kollam 691005, India (e-mail: adarsh_ice@yahoo.co.in).

time scale of variability

- 2) Build SLR models to predict each IMFs as a function of the corresponding IMF of different predictor variables.
- 3) Refine the model by discarding the IMFs of the factors having the p -value is greater than 0.1
- 4) Predict the IMFs at different time scales by the refined model
- 5) Add the predicted IMFs to get the predicted rainfall

The above built model is designated as MEMD-SLR model in this paper, while SLR model refers the stepwise linear regression model between rainfall and the predictor variables (5 meteorological parameters in this study). The MEMD –SLR model takes general form

$$OM_{Ri} = \sum_{i=1}^{NP} r_i OM_{PV_i} \text{ and } R = \sum_{i=1}^M OM_{Ri} \quad (1)$$

where OM denote an orthogonal mode (an IMF or the residue), M is the total number of decomposed modes, NP is the number of predictor variables; PV is the predictor variable; r_i is the regression coefficient; R is the rainfall

The theoretical background of MEMD, SLR and M5 Model Tree are presented in the following sections.

A. Multivariate Empirical Mode Decomposition (MEMD)

A brief description of the MEMD algorithm, is presented below [13]-[15].

In this method, multiple envelopes are produced by taking projections of multiple inputs along different directions in an m -dimensional space. Assuming

$V(t) = \{v_1(t), v_2(t), \dots, v_m(t)\}$ being the m vectors as a function of time t and $X^{\theta_k} = \{x_1^k, x_2^k, \dots, x_m^k\}$ denoting the direction vector along different directions given by angles $\theta_k = \{\theta_1^k, \theta_2^k, \dots, \theta_{m-1}^k\}$ in a direction set X ($k=1, 2, 3, \dots, K$, K is the total number of direction). It can be noted that the rotational modes appears as the counterparts of the oscillatory modes in EMD or its variants. The IMFs of m temporal datasets can be obtained by the following algorithm :

- 1) Generate a suitable set of direction vectors by sampling on a $(m-1)$ unit hypersphere
- 2) Calculate the projection $p^{\theta_k}(t)$ of the datasets $V(t)$ along the direction vector X^{θ_k} for all k
- 3) Find temporal instants $t_i^{\theta_k}$ corresponding to the maxima of projection for all k
- 4) Interpolate $[t_i^{\theta_k}, V(t_i^{\theta_k})]$ to obtain multivariate envelop curves $e^{\theta_k(t)}$ for all k
- 5) The mean of envelope curves ($M(t)$) is calculated by

$$M(t) = \frac{1}{K} \sum_{k=1}^K e^{\theta_k(t)}$$

- 6) Extract the ‘detail’ $D(t)$ using $D(t) = V(t) - M(t)$. If $D(t)$ fulfills the stoppage criterion for a multivariate IMF, apply the above procedure to $V(t)$ the above procedure to $V(t) - D(t)$, otherwise apply it to $D(t)$.

Hammersley sampling sequence can be used for the generation of direction vectors [14] and the stoppage criteria reported in literature [16], [17] can be used in the implementation of MEMD.

B. Stepwise Linear Regression (SLR)

Stepwise linear regression is a method for adding and removing terms from a multi-linear model based on their statistical significance in a regression process. The method begins with an initial model and then compares the explanatory power of incrementally larger and smaller models. At each step, the p -value of an F -statistic is computed to test models with and without a potential term. If a term is not currently in the model, the null hypothesis is that the term would have a zero coefficient if added to the model. If there is sufficient evidence to reject the null hypothesis, the term is added to the model. Conversely, if a term is currently in the model, the null hypothesis is that the term has a zero coefficient. If there is insufficient evidence to reject the null hypothesis, the term is removed from the model. The method proceeds as follows [15], [18].

- 1) Fit the initial model.
- 2) If any terms not in the model have p -values less than an entrance tolerance (that is, if it is unlikely that they would have zero coefficient if added to the model), add the one with the smallest p value and repeat this step; otherwise, go to step 3.
- 3) If any terms in the model have p -values greater than an exit tolerance (that is, if it is unlikely that the hypothesis of a zero coefficient can be rejected), remove the one with the largest p value and go to step 2; otherwise, stop.

Depending on the terms included in the initial model and the order in which terms are moved in and out, the method may build different models from the same set of potential terms. The method terminates when no single step improves the model. In this study, an IMF of a parameter was added to the regression equation when the value of $p \leq 0.05$ and was taken out from the regression equation when the value of $p \geq 0.10$. Then the values of rainfall at the measurement scale are predicted by summing up the predicted IMFs and residue of rainfall. Different performance evaluation measures are used to judge the quality of overall prediction of rainfall considering the observed rainfall and predicted rainfall.

C. M5 Model Tree

Model Tree (MT) is a popular machine learning technique used for solving regression problems through classification and decision making. MT was originally proposed by Quinlan [19]. Model Trees follows a modular approach so that the entire domain is divided into sub-domains and multi-linear regression models are developed for each sub-domain. Therefore it formulates many piecewise linear models to approximate the non-linear relationship between the predictors and the predict and.

The first stage in preparation of MT based models involves use of a splitting criterion to create a decision tree. Depending upon the method of splitting of the domain, there exist different learning algorithms for model trees. The one which uses standard deviation reduction (SDR) as the splitting criteria is called as M5 learning algorithm [20]. In this method the standard deviation of the class values that

reach a node is treated as a measure of the error at that node and the expected reduction in this error as a result of testing each attribute at that node is calculated. The computation of SDR can be represented as follows:

$$\sigma_R = \sigma(N) - \sum \frac{|T_i|}{|N|} \sigma(T_i) \quad (2)$$

where, σ_R = standard deviation reduction; N is the total number of training samples; T_i is the training samples of i^{th} sub-domain; $\sigma(N)$ and $\sigma(T_i)$ are the standard deviations of complete training samples, and T_i sub-domain samples respectively. A typical representation of this model could be in the form: $O = a_0 + a_1x_1 + a_2x_2 + \dots$ where, O is the output values of the sub-domain; a_0, a_1, a_2, \dots are the coefficients of linear regression; x_1, x_2, \dots are the input values of the sub-domain. The splitting continues till the class values of all the instances that reach a node varies negligibly or only a few instances remain. Then the model improvement is done by two ways - performing ‘pruning’ and ‘smoothing’ operations, which reduce the effect of ‘overfitting’ and sharp discontinuities between different sub-classes, which happen especially when the dataset is very small [20]. A more detailed description of the theory behind model trees and pictorial representations can be found elsewhere [20], [21]. The use of M5 model tree reduces the effort for setting of algorithm specific user defined parameters [22] and another major advantage of M5 Model Tree is that, it combines several simple linear relations and hence more transparent and acceptable by decision makers [23], [24].

III. DATA

Indian Institute of Tropical Meteorology, Pune defined 36 meteorological subdivisions in India. Kerala meteorological subdivision, located on western coast of India lies between 8°N to 17°N in latitude and from 70° to 80°E in longitude, is chosen as the study area. The monthly precipitation data of the subdivision for 45 years (from January 1961 to December 2005) is computed from the daily gridded rainfall data at 0.25° resolution prepared by the Indian Institute of Tropical Meteorology, Pune (<http://www.tropmet.res.in>) [25]. The region is called as ‘the gate way of Indian monsoon’ and climatologically one of the most important region in Indian subcontinent. It is customary that in the absence of reliable meteorological datasets for longer periods, the data provided by meteorological data centers such as National Centre for Environmental Prediction (NCEP) can be used for climatic downscaling studies [1]-[5]. Five meteorological variables such as mean sea level pressure, u wind at 850 mb, v wind components at surface, near surface temperature, specific humidity, as these variables are proven to be strongly correlated with the rainfall in Kerala meteorological subdivision in past studies [26]. It is to be noted that a smaller NCEP grid comprising 4 grid points is compared in the present study. To alleviate the possible undesirable effects of spatial heterogeneity of meteorological data in modeling, owing to the direct use of the selected predicted variables

(without going for the computations of PCs), this study uses a finer NCEP grid when compared with that of whole southern India, as adopted in some past studies [26]. The NCEP/NCAR reanalysis data [27] (<http://www.cdc.noaa.gov/cdc/reanalysis/reanalysis.html>) for the period, for the spatial extent 7.5°N-12.5°N and 75°E-77.5°E for the period 1961-2005 are used as predictor variables. This model considers 30 years of data (1961-1990) for training and the data of the rest of the period (1991-2005) used for validation.

IV. STATISTICAL DOWNSCALING OF RAINFALL DATA

The MEMD-SLR hybrid approach is used as the transfer function for downscaling of rainfall different geographical grid points in the Kerala subdivision. To demonstrate the application of the proposed methodology, a typical grid point in the Chalakudi basin (Point 11) in the central Kerala is considered, the location of which is shown in Fig. 1. The data of the meteorological variables (from NCEP) and the rainfall at grid point 11 are decomposed into several number of oscillatory modes.

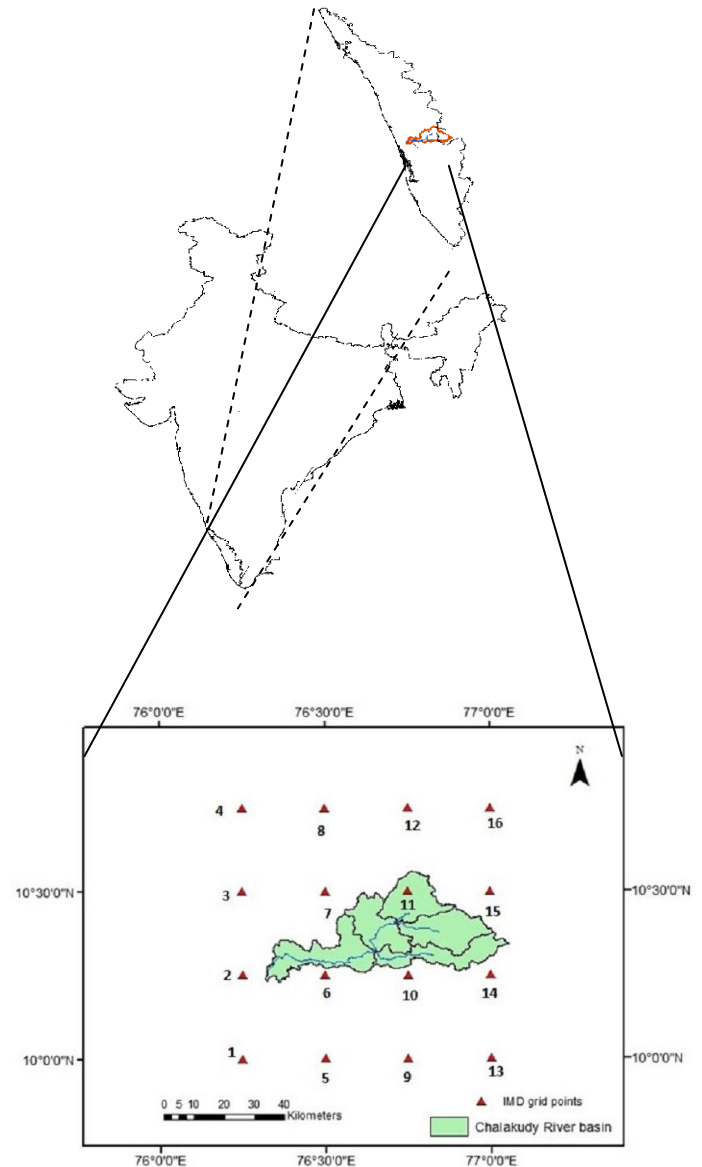


Fig. 1. Location map of the study area and demonstration of IMD grid points in Chalakudi basin.

The tolerance parameters of MEMD are fixed as 0.075, 0.5 and 0.075 following the past studies [17]. The decomposition resulted in 6 modes and residue, which is less the expected maximum of $\log_2(N)$, N is the data length (540 in the study). The modes obtained are shown in Fig. 2.

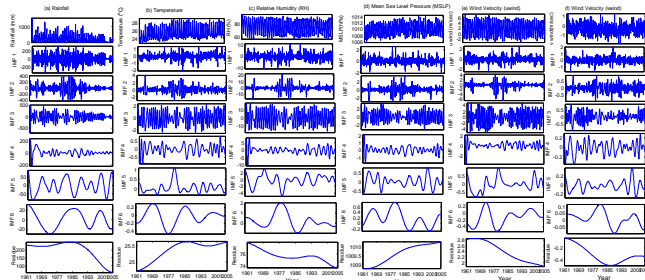


Fig. 2. Decomposition of (a) rainfall and (b)-(f) the four meteorological variables pertaining to the grid point 11 in Chalakkudi basin.

The SLR models are fitted for each rotatory mode for the training period (1961-1990) by considering the corresponding modes of meteorological variables as inputs. The input selection is done in such a way that the impact of inclusion of a parameter (successively) is assessed based on the p value of fitting (such that p value is less than 0.05). For brevity, the linear models of different modes are not presented here. The final addition of predicted components gives the rainfall for the training period. On supplying the modes for validation period as input to the model at a specific time scale, the modes of rainfall for validation period for the corresponding scale are predicted. Eventually the predicted modes are added to obtain the rainfall for the validation period. By considering the actual meteorological parameters for the training period (1961-1990) as input MLR and M5 model tree (MT) based models are developed. By giving the meteorological parameters for the validation period as input to the respective models, the rainfall for the validation period is obtained. Then a rigorous performance evaluation exercise is carried out to investigate the efficacy of the MEMD-SLR transfer function model. The evaluation criteria such as coefficient of determination (R^2), Nash Sutcliffe Efficiency (NSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), mean Bias Error (MBE) and Scatter Index (SI) are considered as evaluation criteria. The evaluation statistics are summarized in Table I. Table I show that the validation data the MEMD-SLR model is superior, as the model results are associated with less error (RMSE, MAE, MBE, SI) and better fitting (R^2 and NSE).

TABLE I: PERFORMANCE EVALUATION OF DIFFERENT TRANSFER FUNCTION MODELS

Criteria	Training			Validation		
	MLR	MT	MEMD-SLR	MLR	MT	MEMD-SLR
R^2	0.643	0.772	0.707	0.454	0.469	0.583
NSE	0.643	0.771	0.705	0.055	0.15	0.496
RMSE	157.1	125.4	142.8	147.2	162.3	107.6
MAE	106.9	73.82	96.335	101.9	104.9	74.8
MBE	-0.02	-0.03	-4.036	-68.3	-76.2	-12.0
SI	0.697	0.558	0.634	1.183	1.304	0.865

Further for the validation data, the performance of the three models was depicted by time series plot and the Taylor diagram (which considers the standard deviation of

predictions and correlation coefficient between observed and predicted rainfall). The plots are given in Fig. 3. The position coordinates of the performance statistics different models in the Taylor diagram (Fig. 3) clearly decipher the superiority of MEMD-SLR transfer function model.

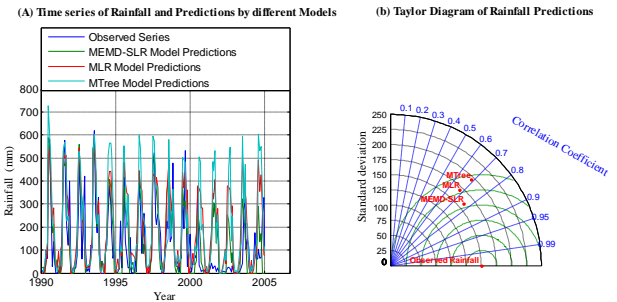


Fig. 3. (a) Time series plot of observed rainfall and predictions based on the three transfer function models; (b) Taylor diagram of predictions based on three transfer function models.

Hence on comparison of Table I and Fig. 3 it is clear that the MEMD-SLR model is performing reasonably well and better than that of MT and MLR models. The MEMD-SLR models are developed for 132 grid points falling in the Kerala subdivision and those models are applied for rainfall projections by considering the meteorological data provided by Canadian Centre for Climate Modeling and Analysis (CCCMA) for the three Coupled Model Inter-comparison Project Phase 5 (CMIP5) based Representative Concentration Pathway (RCP) scenarios (RCP2.6, RCP 4.5 and RCP8.5). Finally, spatially averaged representative monthly rainfall of the subdivision is computed for each scenario. The cumulative distribution function (CDF) of rainfall for the period 2006-2096, for the three scenarios is prepared using Weibull plotting position formula and the result is presented in Fig. 4. Moreover, the CDFs of rainfall for the three scenarios for the three slices of period [2006-36 (2020s the near future), 2037-2056 (the 2050s) and 2057-2096 (2080s, the far future)]. The CDFs obtained for different slices for the three scenarios are presented in Fig. 5.

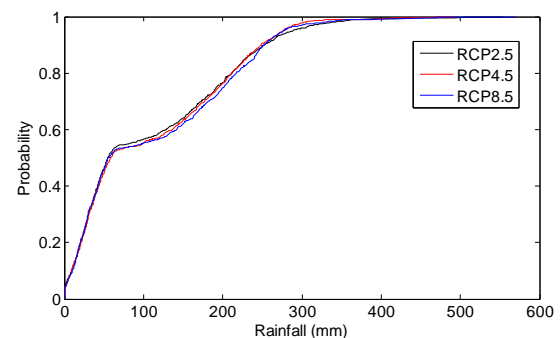


Fig. 4. CDFs of monthly rainfall of Kerala for 2006-2096.

From Fig. 5 it is noticed that the rainfall value at which the CDF approaches 1 is the highest for the RCP 8.5 scenario, hence among the three scenarios, the highest rainfall is projected by this scenario.

After careful perusal of Fig. 5, (by noticing the value of rainfall at which the CDF reaches 1), it is inferred that, highest rainfall is projected by RCP8.5 scenario in the near future period (2020s). In the far future, highest rainfall is projected by RCP2.5, but with a marginal difference with

RCP8.5 scenario. In all the slices the maximum rainfall projected by RCP4.5 scenario is the lowest among the three projections.

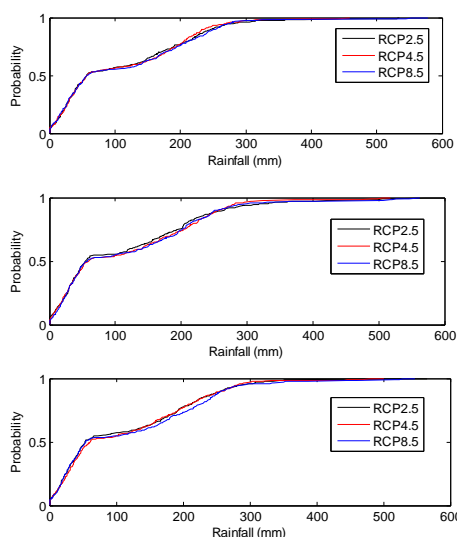


Fig. 5. CDFs of monthly rainfall of Kerala for the three time slices of the present century (2006-2096) for different RCP scenarios.

This study proposed an alternative strategy for transfer function based downscaling of rainfall. The method is based on the MEMD technique. The technique simultaneously decomposes all the potential predictor variables of interest into multiple time scales. The ‘mode alignment’ property of MEMD ensure the same number of modes for each variable when compared with EMD, while it doesn’t demand the ‘*a priori*’ fixation of decomposition levels like the discrete wavelet transform. The proposed approach considers the meteorological parameters directly rather than the PCs. Hence the proposed method offers certain advantages in station specific downscaling operations and more appealing to the public domain. The present study provided broad inferences on rainfall for the different time slices of the century (2001-2100) under different CMIP5 scenarios but to draw concrete conclusions on the expectancy of actual future rainfall, the modeling exercise using multiple GCM data and subsequent uncertainty analysis of the results is quite essential.

V. CONCLUSIONS

The non-stationary behavior of hydrologic time series is more noticeable because of changing climate. Hence alternate modeling practices are to be explored for statistical downscaling of hydro-climatic variables, with the aim to improve the accuracy of forecasting non-stationary hydrologic time series data. This paper proposes an alternative method for transfer function based statistical downscaling, where in a multiscale decomposition step is used as data pre-processing tool and the rotatory components (modes) obtained are later on recombined to facilitate the rainfall predictions. The proposed method is demonstrated by application to downscale GCM outputs to project monthly rainfall in Kerala meteorological subdivision in India. The rigorous statistical performance evaluation showed that the MEMD based transfer function model performs better when

compared with M5 Model Tree and MLR based models. Further, the developed MEMD-SLR hybrid model is applied for projecting rainfall of Kerala subdivision under three RCPs provided by Canadian Centre for Climate Modeling and Analysis (CCCMA). It is found that highest rainfall is projected by RCP8.5 likely occurrence in the near future (2020s); while the extreme rainfall projected by the RCP4.5 scenario is the lowest among the three scenario based projections.

ACKNOWLEDGMENT

The present research work is supported by the research seed money scheme “Finer Scale Rainfall Projections for Kerala Meteorological Subdivision by Statistical Downscaling of GCM simulations” (RSM073/2013) of Centre for Engineering Research and Development (CERD) under the Government of Kerala. The author also thanks Mrs. Sreedevi S, Research scholar, Indian Institute of Technology Bombay India and Ms. Priya Philip, Assistant Professor, Amal Jyothi College of Engineering Kanjirappally Kerala, India for their constant support.

REFERENCES

- [1] S. Ghosh and P. P. Mujumdar, “Non-parametric methods for modeling GCM and scenario uncertainty in drought assessment,” *Water Res. Research*, vol. 43, no. 7, 2007.
- [2] S. Ghosh, “SVM-PGSL coupled approach for statistical downscaling to predict rainfall from GCM output,” *J. Geophys. Res. Atmos.*, vol. 115, no. D22, 2010.
- [3] S. M. Asokan and D. Dutta, “Analysis of water resources in the Mahanadi River Basin, India under projected climate conditions,” *Hydrol. Processes*, vol. 22, pp. 3589–3603, 2008.
- [4] M. K. Goyal, D. H. Burn, and C. S. P. Ojha, “Evaluation of machine learning tools as a statistical downscaling tool: Temperatures projections for multi-stations for Thames River Basin, Canada,” *Theor. Appl. Climatol.*, vol. 108, no. 3-4, pp. 519–534, 2012.
- [5] K. Salvi, S. Kannan, and S. Ghosh, “High-resolution multisite daily rainfall projections in India with statistical downscaling for climate change impacts assessment,” *J. Geophysical Research. Atmosphere*, vol. 118, pp. 3557–3578, 2013.
- [6] M. S. Parvez and G. M. Henebry, “Projections of the Ganges–Brahmaputra precipitation — Downscaled from GCM predictor,” *J. Hydrol.*, vol. 517, pp. 120–134, 2014.
- [7] K. Shashikanth, C. G. Madhusoodhanan, S. Ghosh, T. I. Eldho, K. Rajendran, and R. Murtugudde, “Comparing statistically downscaled simulations of Indian monsoon at different spatial resolutions,” *J. Hydrol.*, vol. 519, pp. 3163–3177, 2015.
- [8] K. Salvi, S. Ghosh, and A. R. Ganguli, “Credibility of statistical downscaling under non-stationary climate,” *Clim. Dyn.*, vol. 40, no. 5, pp. 1991–2003, 2016.
- [9] Y.-F. Sang, “A review on the applications of wavelet transform in hydrology time series analysis,” *Atmos. Res.*, vol. 122, pp. 8–15, 2013.
- [10] V. Nourani, A. H. Baghanam, J. Adamowski, and O. Kisi, “Applications of hybrid wavelet–Artificial Intelligence models in hydrology: A review,” *J. Hydrol.*, vol. 514, no. 6, pp. 358–377, 2014.
- [11] S. Adarsh and M. J. Reddy, “Wavelet model tree conjunction approach for rainfall projections under representative concentration pathways,” *Hydro International at IIT Madras*, pp. 29–37, 2013.
- [12] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis,” in *Proc. Royal Soc. London, Series A*, vol. 454, pp. 903–995, 1998.
- [13] N. Rehman and P. P. Mandic, “Multivariate empirical mode decomposition,” in *Proc. Royal Soc. Series A*, vol. 466, pp. 1291–1302, 2010.
- [14] G. Huang, Y. Su, A. Kareem, and H. Liao, “Time-frequency analysis of non-stationary process based on multivariate empirical mode decomposition,” *J. Eng. Mechanics*, vol. 142, no. 1, 2016.

- [15] W. Hu and B-C. Si, "Soil water prediction based on its scale-specific control using multivariate empirical mode decomposition," *Geoderma* vol. 193-194, pp. 180-188, 2013.
- [16] N. E. Huang, M. C. L. Wu, S. R. Long, S. S. P. Shen, W. Qu, P. Gloersen, and K. L. Fan, "A confidence limit for the empirical mode decomposition and Hilbert spectral analysis," in *Proc. Royal Soc. Series A*, pp. 2317-2345, 2003.
- [17] G. Rilling, P. Flandrin, and P. Gonçalves, "On empirical mode decomposition and its algorithms," in *Proc. IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP-03*, Grado (Italy), pp. 8-11, June 2003.
- [18] N. R. Draper and H. Smith, *Applied Regression Analysis*, Hoboken, NJ: Wiley-Interscience, pp. 307-312, 1998.
- [19] J. R. Quinlan, "Learning with continuous classes," in *Proc. Australian Joint Conference on Artificial Intelligence*, World Scientific Press, Singapore, pp. 343-348, 1992.
- [20] V. Jothiprakash and A.S. Kote, "Effect of pruning and smoothing while using M5 Model Tree technique for reservoir inflow prediction," *J. Hydrol. Eng.*, vol. 16, no. 7, pp. 563-574, 2011.
- [21] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann, San Francisco, CA, USA, 2005.
- [22] K. K. Singh, P. Pal, and V. P. Singh, "Estimation of mean annual flood in Indian catchments using back-propagation neural network and M5 model tree," *Water Res. Manage.*, vol. 24, pp. 2007-2019, 2011.
- [23] D. P. Solomatine and Y. Xue, "M5 model trees and neural networks: Application to flood forecasting in the upper reach of Huai River in China," *J. Hydrol. Eng.*, vol. 9, no. 6, pp.1-10, 2011.
- [24] B. Bhattacharya and D. P. Solomatine, "Machine learning in sedimentation modeling," *Neural Networks*, vol. 19, no. 2, pp. 208-214, 2006.
- [25] D. S. Pai, S. Latha, M. Rajeevan, O. P. Sreejith, N. S. Satbhai, and B. Mukhopadhyay, "Development and analysis of a new high spatial resolution(0.25 X 0.25) long period(1901-2010) daily gridded rainfall dataset over India," National Climate Centre Research Report No. 1/2013, 2013.
- [26] S. Sreedevi, "River basin scale hydrological and water quality analysis due to climate change," Ph.D evaluation report submitted to Indian Institute of Technology Bombay, 2014.
- [27] E. Kalnay, M. Kanamitsu *et al.*, "The NCEP/NCAR 40-year reanalysis project," *Bull. Amer. Meteor. Soc.*, vol. 77, pp. 437-471, 1996.



S. Adarsh was born in the state of Kerala, India. He completed his B.Tech degree in civil engineering in the year 2003 from TKM College of Engineering Kollam. He joined as an assistant professor in the civil Engineering Department of the same institute in 2004. He completed the M.Tech in water resources engineering from IIT Bombay in 2009. He is currently pursuing his Ph.D at IIT Bombay on multiscale analysis of hydrologic time series. He published 16 papers in international journals of high reputation and many in conferences. His areas of research interest include hydrologic time series analysis, water resources systems, data driven techniques for hydrologic modeling and hydro climatology.