7th International Conference on Water Resources and Arid Environments (ICWRAE 7): 587-595
4-6 December 2016, Riyadh, Saudi Arabia

Prediction of Indian Summer Monsoon Rainfall Based on Scale Specific Control using MEMD- SLR Model

[1,2]S. Adarsh and [1]M. Janga Reddy

[1]Department of Civil Engineering, Indian Institute of Technology, Bombay, India
[2] Department of Civil Engineering, TKM College of Engineering, Kollam, India

Abstract: This paper proposes a new methodology for prediction of Indian Summer Monsoon Rainfall (ISMR) considering the time scale based decomposition of input variables. The multivariate dataset comprising the lagged rainfall values along with the current rainfall value are decomposed into different orthogonal modes using the Multivariate Empirical Mode Decomposition (MEMD) method. Then separate models are prepared to predict different modes (after identifying the modes of potential predictors as inputs based on the *p*-value statistics) using the Stepwise Linear Regression (SLR) fitting. The predicted modes are finally recombined to obtain the rainfall of the desired time step. The utility of proposed methodology is demonstrated by predicting rainfall in India for June, July, August and September months as well as for monsoon season keeping the data for 1871-1960 period for calibration and data of 1961-2010 period for validation. In all cases, it was found that the proposed methodology is resulting in superior performance as compared to the results reported in earlier studies using Artificial Neural Network (ANN). Further the efficacy of the proposed methodology is evaluated for the prediction of low and high seasonal rainfall by using various statistical performance measures.

Key words: Rainfall · Monsoon · Decomposition · MEMD · Regression

INTRODUCTION

Prediction of Indian Summer Monsoon Rainfall (ISMR) is a challenging problem for the meteorologists and hydrologists. The history of ISMR prediction perhaps started with Sir Henry Blanford [1] who tried to associate the ISMR with Himalayan snow cover. Later on Sir Gilbert Walker made significant contributions on finding the association of ISMR with different oceanic, meteorological and atmospheric parameters [2-4]. For the prediction of ISMR, mainly two types of studies were followed by the researchers over the years. In the first category of studies, the cause-effect relationships of ISMR are established through regression models in which they attempted to identify and incorporate the appropriate factors influencing monsoon rainfall and its prediction. Some of the recent studies that have considered the large scale circulation patterns as inputs for monsoon rainfall prediction for different parts of India falls in this category [5-9]. In the second category, many studies applied time series models for prediction of monsoon rainfall in different parts of the country [10-16]. In such studies lagged values of rainfall are considered as inputs and many of them used Artificial Neural Networks (ANN) as modeling tool [10-16]. Sahai *et al.*, [12] elaborately presented the logical reasoning behind the use of a time series approach for ISMR prediction. One of the important justifications made in the study is that the rainfall is the end product of different atmospheric processes which are related to different predictors.

Thus if there are any connections between ISMR and different predictors, all such information are embedded in the time series itself [12].

It is well understood fact that most of the hydrologic variables pose multi-scaling behavior and capturing information in multiple time scale may improve the prediction. In this context, the use of appropriate decomposition method such as wavelet transform or Empirical Mode Decomposition (EMD) were used as data preprocessing tools to model hydrological processes [17-19]. In the past, many EMD based hybrid models were proposed for simulation and forecasting of hydrologic variables [20-25]. Iyengar and Raghukanth [26, 27] applied a hybrid EMD–ANN method for summer monsoon rainfall

Corresponding Author: S. Adarsh, Department of Civil Engineering, Indian Institute of Technology Bombay, India.
E-mail: adarsh.sankaran@yahoo.co.in.

predictions in different parts of India. Most of such studies first performed the decomposition of available time series into multiple time scales and then identified the appropriate number of lagged values of modes to model the prediction of the concerned variable. But in multiscale hydrological modeling, often multiple factors need to be considered as predictors which cannot be performed by the traditional EMD or its variants. Identification of potential predictors from each of such variables at different time-scales is again a challenging step in modeling and none of the past studies have considered this vital information capturing step. Therefore, this study proposes a new modeling strategy in which identification of significant predictors at each time scale (omission of rest of them) is made to predict the modes of rainfall at current time step by considering multiple predictor variables.

The rest of the paper is organized as follows: first the description of proposed methodology and the algorithms of MEMD and SLR are briefly presented, followed by the description of datasets. The modeling and results are presented subsequently along with conclusions from the study.

Proposed Methodology: Identification of time scale of variability is an important step in multiscale hydrological simulation and forecasting, for which an appropriate decomposition technique need to be adopted. Huang *et al.,* [28] propounded a purely 'data adaptive' decomposition procedure namely Empirical Mode Decomposition (EMD) to analyze the non-linear and non-stationary time series. Some important merit of this technique are (i) it doesn't demand '*apriori*' fixing of decomposition levels and basis function type; (ii) and being data adaptive, it is unique and intuitive. Multivariate EMD proposed by Rehman and Mandic [29] 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 rainfall data and uses the stepwise linear regression (SLR)for building the regression models for each of the resulting components. The proposed methodology involves the following steps:

- Decomposition of rainfall and predictor variables (lagged rainfall values) using MEMD to get different orthogonal oscillatory modes called intrinsic mode functions (IMFs), each with specific time scale of variability.
- Build SLR models to predict each IMFs as a function of the corresponding IMF of different predictor variables.
- Refine the model by discarding the IMFs of the factors having the *p*-values greater than 0.1.
- Predict the IMFs (of current time step rainfall) at different time scales by the refined model.
- Add the predicted IMFs to get the current time step rainfall.

The above built model is designated as MEMD-SLR model in this paper. It takes general form

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

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 (lagged rainfalls); PV is the predictor variable; r_i is the regression coefficient; R is the rainfall.

The theoretical background of MEMD and SLR are presented in the following sections.

Multivariate Empirical Mode Decomposition (MEMD): A brief description of the MEMD algorithm, is presented below [29-31].

In this method, multiple envelops are produced by taking projections of multiple inputs along different directions in an m-dimensional space. Assuming $V(t)=\{v_1(t),v_2(t)....v_m(t)\}$ being the m vectors as a function of time t and $X^{\theta_k}=\{x_1^k,x_2^k,....,x_m^k\}$ denoting the direction vector along different directions given by angles $\alpha_k=\{\alpha_1^k,\alpha_2^k,...\alpha_{m-1}^k\}$ in a direction set X (k=1,2,3,….K, K is the total number of directions). 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:

- Generate a suitable set of direction vectors by sampling on a (m-1) unit hypersphere
- Calculate the projection $p^{\alpha_k}(t)$ of the datasets $V(t)$ along the direction vector X^{α_k} for all k
- Find temporal instants $t_i^{\alpha_k}$ corresponding to the maxima of projection for all k

- Interpolate $[t_i^{\alpha_k}, V(t_i^{\alpha_k})]$ to obtain multivariate envelop curves $e^{\alpha_k}(t)$ for all k
- The mean of envelope curves ($M(t)$) is calculated by $M(t) = \frac{1}{K}\sum_{k=1}^{K} e^{\alpha_k}(t)$
- Extract the 'detail' $D(t)$ using $D(t) = V(t)-M(t)$. If $D(t)$ fulfills the stopping criterion for a multivariate IMF, repeat the steps (2) to (5) to get the subsequent IMFs

Hammersley sampling sequence can be used for the generation of direction vectors and the stopping criteria reported in literature [30] can be used in the implementation of MEMD.

Stepwise Linear Regression (SLR): Stepwise linear regression is a regression method which involves successive addition and removal of terms from a multi-linear model based on their statistical significance. 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 a sufficient evidence to reject the null hypothesis, the term is added to the model. The method proceeds as follows [32].

- Fit the initial model.
- 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.
- 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=0.05 (entrance tolerance) and was taken out from the regression equation when the value of p=0.10 (exit tolerance). 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.

Table 1: The statistical properties of the dataset

Statistical Property	June	July	August	September	Seasonal
Mean (mm)	163.05	271.58	242.73	169.46	846.82
Minimum (mm)	78.20	117.60	144.10	77.20	603.90
Maximum (mm)	241.60	346.00	339.30	267.80	1020.10
Std. Dev.	36.23	38.40	38.08	37.45	83.69
Coeff. of Var.	22.22	14.14	15.69	22.10	9.88
Skewness	-0.06	-1.16	0.01	0.16	-0.53
Kurtosis	2.45	5.44	2.55	2.42	2.97

Dataset: The Indian Institute of Tropical Meteorology (IITM) Pune classified Indian subcontinent to 37 meteorological subdivisions. Long term spatially averaged data prepared by IITM Pune for all subdivisions at monthly scale (for 1871-2013) are available at http://www.tropmet.res.in/. Also the spatially averaged data of India as a whole (All India, AI) and five homogeneous rainfall regions are available. All India summer monsoon rainfall data for the four months June, July August and September (JJAS) for the period 1871-2010 is used in the present study. The data of initial 90 years during 1871-1960 are used for calibration and then the remaining data during 1961-2010 are used for validation. This dataset is proportioned in the above stated manner in order to make a comparison of the results by proposed model with that reported in other studies [16]. The descriptive analysis of the dataset used and its properties are provided in Table 1.

RESULTS AND DISCUSSION

In the time series based ISMR prediction, the lagged rainfall values are used as inputs which are often considered as integral effect of different atmospheric processes/variables influencing monsoon, which inturn possess multiscaling property. Therefore performing the ISMR predictions incorporating the information pertaining to multiple time scale is of having considerable relevance. In some of the past studies it is proved that 5 lagged values are sufficient for ISMR predictions [12,14,16]. Hence in the present study, 5 lagged rainfall values are considered as inputs and therefore the predictions are available for the period 1876-2010. Models are prepared in such a way that the present year rainfall values for seasonal and June, July, August, September etc. are predicted based on the respective values for the previous

Table 2: Regression coefficient matrix for SLR models of different IMFs and residue. The figures in italics represents that the coefficients are not significant

	Mode							
Variable	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	Residue
$R_{seasonal(t-1)}$	-2.523	0.265	1.574	1.332	1.184	2.164	1.082	1.074
$R_{seasonal(t-2)}$	-3.709	-1.310	-2.021	1.311	0.321	-1.889	-0.205	0.272
$R_{seasonal(t-3)}$	-3.523	*-0.161*	1.343	-2.080	-1.080	-3.000	0.283	*0.000*
$R_{seasonal(t-4)}$	-2.168	-0.605	-0.624	*-0.525*	0.512	4.851	*0.040*	-0.182
$R_{seasonal(t-5)}$	-0.745	-0.242	*-0.033*	0.925	*-0.053*	-1.352	-0.047	-0.114

Table 3: Performance comparison between proposed MEMD-SLR model and ANN model (Singh and Borah [16]) for monsoon rainfall prediction

Performance	MEMD-SLR		ANN model	
Measure	Calibration	Validation	Calibration	Validation
R	0.972	0.950	0.92	0.85
RMSE	19.00	21.77	28.12	26.02
PP	0.766	0.687	0.65	0.70
Mean observed	854.931	833.034	855.3	833.03
Mean Predicted	854.931	837.450	852.4	825.95
SD observed	81.290	86.695	80.8	86.7
SD predicted	80.973	81.380	72.4	78.56

5 years. It is to be noted that the summation of the monthly values of JJAS gives the rainfall for monsoon season, which is referred as seasonal rainfall in this study. The input and output for the typical model for seasonal rainfall can be illustrated as:

$$R_{seasonal(t)} = f(R_{seasonal(t-1)}, R_{seasonal(t-2)}, R_{seasonal(t-3)}, R_{seasonal(t-4)}, R_{seasonal(t-5)})$$

All these six variables constitute the multivariate dataset for preparing the MEMD-SLR model. The tolerance parameters of MEMD are fixed as 0.075, 0.5 and 0.075 following the past studies [30, 33]. The decomposition resulted in 7 IMFs and residue, for all models except that for June month (in which the decomposition resulted in 6 modes and residue). In all cases maximum of 7-8 modes is considered to be stable as the expected maximum is log 2(N), N is the data length (135 in the study). Hence the decomposition is acceptable with the selected parameter setting. Here the stepwise regression models are prepared by considering the orthogonal mode of rainfall (at a particular time scale) at current time step as output and orthogonal modes of lagged rainfall at the same scale as inputs. The initial regression coefficients obtained during the preparation of SLR models (corresponding to different time scale), are summarized in Table 2. The numbers in italics indicate that those coefficients (and hence the IMF of the corresponding variable) is not significant at that time scale. For e.g., in IMF2, the rainfall with lag 3, in IMF3 rainfall of lag 5 etc. Such coefficients are assigned as zero and the regression coefficient matrix is revised to make the predictions at different time scales. Finally summation of all predicted modes helps to rebuild the rainfall at current time step. The proposed MEMD-SLR method is implemented and the seasonal rainfall are predicted (for calibration and validation) are shown in the form of time series plot as well as scatter plot in Fig. 1. The performance of the proposed method is compared based on the statistical performance evaluation of predictions made by study by Singh and Borah [16]. Different performance evaluation measures like correlation coefficient (R), Root Mean Square Error (RMSE) and Performance parameter (PP) [34]) and statistical properties of predictions like mean and standard deviation (SD). The performance evaluation statistics of seasonal predictions by both models are summarized in Table 3. MEMD-SLR models were also prepared for predictions of June, July, August and September rainfall and the performance evaluation of the different cases are summarized in Table 3.

The results of performance evaluation for different rainfall time series (for JJAS data and MEMD-SLR model predictions are provided in months) are provided in Table 4 and time series plots in Fig. 2.

Table 3 and Table 4 clearly show that different error statistics and agreement measures are better for MEMD-SLR model when compared to the ANN model both for calibration and validation stages. The time series plots (Fig. 1 and Fig. 2) also shows good agreement of the proposed model with the observed data for different rainfall time series. From Table 3 and Table 4, it can also be noticed that the standard deviations of predictions show better matching with that of observed series for different cases which clearly show the superiority of the proposed approach. To get a better insight into this aspect, the numerical difference between SD of predicted series from that of observed data is computed for both MEMD-SLR model and Singh and Borah [16] model.

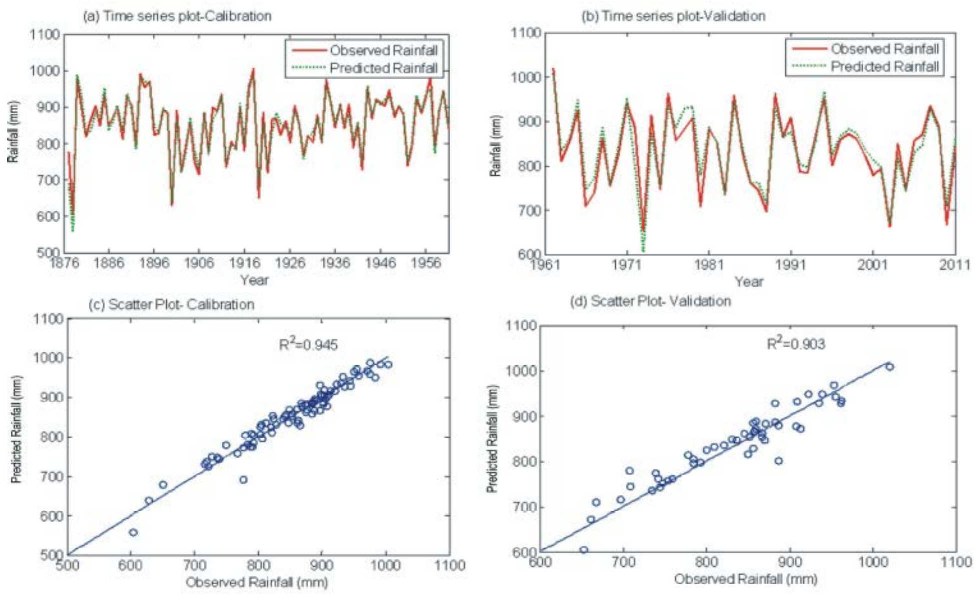

Fig 1: Results of proposed MEMD-SLR model for monsoon rainfall prediction during calibration period (1876-1960) and validation period (1961-2010).

Table 4: Performance comparison between proposed MEMD-SLR model and ANN for different months of monsoon season

	June				July			
	MEMD-SLR		Singh and Borah (2013)		MEMD-SLR		Singh and Borah (2013)	
Performance Measure	Calibration	Validation	Calibration	Validation	Calibration	Validation	Calibration	Validation
R	0.965	0.948	0.89	0.87	0.969	0.918	0.88	0.89
RMSE	9.694	11.810	17.3	18.5	9.470	15.458	16.01	20.09
PP	0.741	0.656	0.53	0.46	0.754	0.585	0.58	0.46
Mean observed	164.400	160.766	164.5	160.76	276.289	263.566	276.4	263.57
Mean Predicted	164.247	164.117	161.01	157.06	276.315	264.896	272.4	260.01
SD observed	37.413	34.372	37.2	34.37	38.517	37.206	38.3	37.21
SD predicted	36.389	29.256	30.3	30.11	37.574	39.158	31.4	32.01
		August					September	
R	0.969	0.958	0.92	0.91	0.952	0.945	0.9	0.88
RMSE	10.192	9.440	14.1	15.12	12.077	11.192	18.08	20.00
PP	0.752	0.712	0.66	0.54	0.693	0.671	0.54	0.41
Mean observed	242.618	242.910	242.7	242.91	171.624	165.792	171.7	165.79
Mean Predicted	242.618	243.718	241.5	239.56	171.597	165.243	168.6	163.02
SD observed	41.057	32.805	40.8	32.8	39.349	34.062	39.1	34.06
SD predicted	38.979	29.896	35.2	27.9	36.305	30.124	32.5	28.01

The results are presented in Fig. 3 and Fig. 4. Fig. 3 and Fig. 4 clearly show that the deviations in predictions are much less than that of Singh and Borah [16] model for all cases except that for the validation data of June rainfall. Here it can be concluded that even though overall statistics is better in this case for MEMD-SLR model, there exists large deviation in predictions. The statistics of the dataset used show that the highest variance is associated with rainfall during this month, i.e., coefficient of variation (mean /SD expressed as percentage) of the dataset is the highest for June month (22.33), followed by that for September (21.8) while that for July and August are relatively less (13.9 and 15.6) respectively. This is obvious because the monsoon dynamics might not have fully developed in the month of June. In such a case the scale separation naturally will not be effective in capturing the variance of the data sets, as the time series approach involves the basic assumption that the rainfall is the end product of various atmospheric processes which are related to different predictors and the information of connections between ISMR and different predictors are embedded in the time series itself [12].

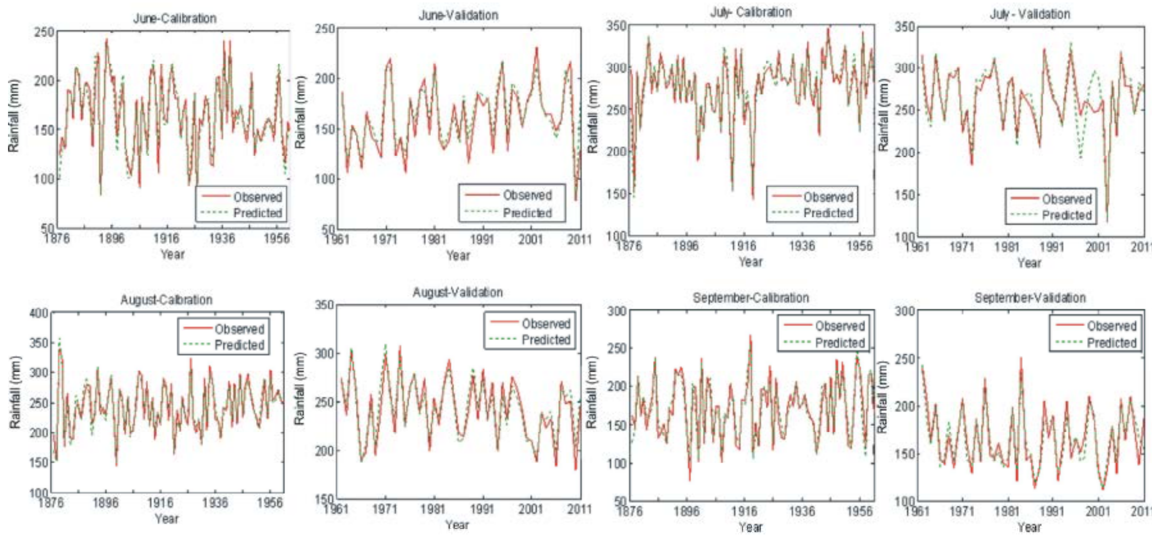

Fig. 2: Time series plots of observed and predicted rainfall (MEMD-SLR Model) of June, July, August and September months for calibration period (1876-1960) and validation period (1961-2010).

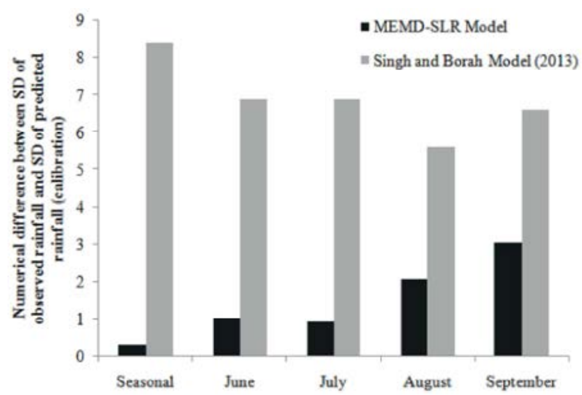

Fig 3: Performance comaprison of MEMD-SLR model and ANN model based on SD of predicions (calibration data)

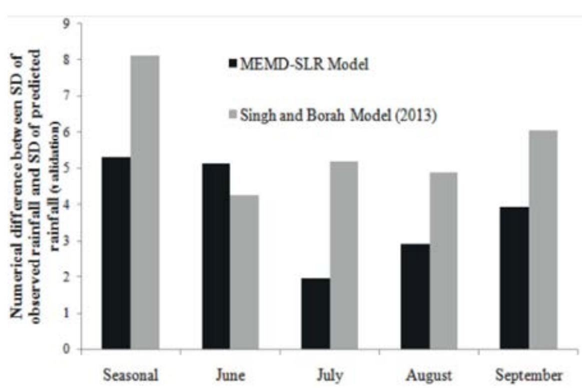

Fig. 4: Performance comaprison of MEMD-SLR model and ANN model based on SD of predicions (validation data)

In the modeling of extreme flow conditions like flood and drought, the low and high rainfall have considerable importance, unfortunately most of the models fail to capture such extreme values. The seasonal monsoon rainfall exceeding (mean+SD) are considered to be high rainfall, less than mean are considered to be low rainfall. The time bar graphs of predictions of high and low rainfall are presented in Fig. 5. After segregating the calibration and validation time series into high and low rainfall, the performance of predictions MEMD-SLR model is evaluated. On segregating in this manner, 11 data points from calibration dataset and 8 points from validation data set are found to be falling in high rainfall category. Also 37 points from calibration data and 21 points from validation data belong to the low rainfall category. Different performance measures like correlation coefficient (R), Nash Sutcliffe Efficiency (NSE), Index of Agreement (IA) [9], RMSE, Mean Absolute Error (MAE), Mean Bias Error (MBE), basic statistical properties of predictions such as mean, standard deviation, minimum and maximum rainfall are presented in Table 5.

A careful perusal of Table 5 shows that the predictions of extreme rainfall are also reasonably accurate by the proposed MEMD-SLR method. Different performance measures like NSE, IA are found to be good in the context of rainfall predictions and different error statistics like RMSE and MAE are relatively less and in similar lines as that reported in some earlier studies on the prediction of Indian monsoon rainfall [35]. Also the statistical properties of the predictions are closer to that of observed data. This clearly infers that the proposed

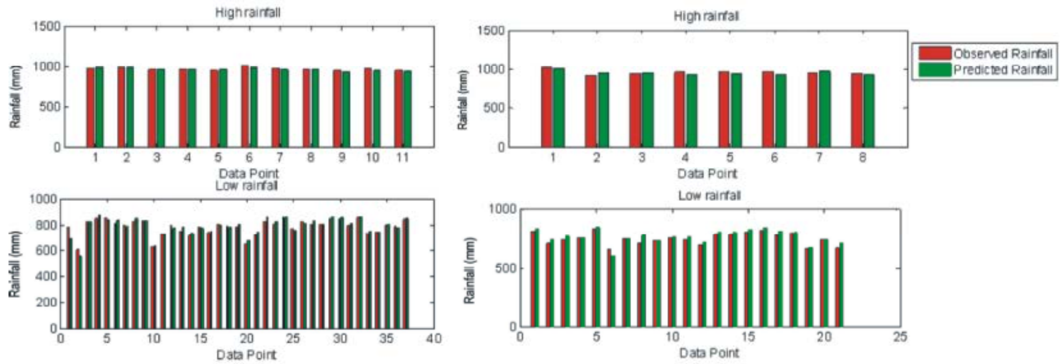

Fig 5: Plots showing observed and predicted rainfall extremes by the proposed model

Table 5: Performance assessment of High and low rainfall prediction by MEMD-SLR model

Performance Measure	High rainfall Calibration	High rainfall Validation	Low rainfall Calibration	Low rainfall Validation
R	0.678	0.729	0.943	0.922
NSE	0.304	0.464	0.859	0.705
IA	0.581	0.558	0.829	0.760
RMSE	15.698	20.101	22.543	27.470
MAE	13.225	17.775	16.509	20.948
MBE	5.619	5.305	-4.509	-16.088
SI	0.016	0.021	0.029	0.037
Max. Observed	1004.200	1020.100	853.100	831.000
Max. Predicted	988.378	1008.082	869.724	849.088
Min. Observed	944.000	922.400	603.900	652.800
Min. Predicted	929.103	928.187	556.513	603.750
Mean Observed	968.518	956.175	781.278	748.819
Mean Predicted	962.899	950.870	785.787	764.907
SD Observed	19.737	29.342	60.949	51.810
SD Predicted	18.438	26.554	67.031	58.586

approach show better generalization capability on considering extreme rainfall predictions also. Overall, the proposed strategy of 'decomposition and exclusion' is found to be a promising modeling practice. The MEMD is quite successful in information capturing at multiple time scales and SLR is capable to identify the potential inputs at different time scales. This facilitates to retain the potential input and omit the less significant input at different time scales, which cannot be achieved through conventional modeling methods. Thus by using the hybrid MEMD-SLR method an improved performance can be obtained for prediction of Indian summer monsoon rainfall.

CONCLUSIONS

Prediction of Indian summer monsoon rainfall (ISMR) is a very challenging task yet important scientific problem for hydro-climatologists. This study proposes a new method for prediction of ISMR considering the time scale based decomposition of lagged rainfall inputs. The decomposition of multivariate dataset is performed by Multivariate Empirical Mode Decomposition (MEMD) method and separate models are built for prediction of different modes of rainfall using stepwise linear regression (SLR). In the model calibration stage, the less influential inputs at each time scale are excluded based on the *p*-value statistic. Results clearly exhibited the superiority of the proposed MEMD-SLR hybrid model over an existing ANN model when applied for prediction of seasonal (monsoon), June, July, August and September month's rainfall. Also the efficacy of the proposed approach is proved by evaluating the prediction of high and low magnitude monsoon rainfall through a multitude of performance evaluation criteria.

REFERENCES

1. Blanford, H.F., 1884. On the connection of the Himalayan snow with dry winds and seasons of droughts in India. Proceedings of Royal Society London 37: 3-22
2. Walker, G.T., 1910. Correlation in seasonal variations of weather. II, Memoirs of the Indian Meteorological Department, 21(Part 2): 22-45.
3. Walker, G.T., 1923. Correlation in seasonal variations of weather VIII, A preliminary study of world-weather. Memoirs of the Indian Meteorological Department, 24(Part 4): 75-131.
4. Walker, G.T., 1933. Seasonal weather and its prediction, British Association for Advances in Science, 103: 25-44.
5. Gadgil, S., P.N. Vinayachandran, P.A. Francis and S. Gadgil., 2004. Extremes of the Indian summer monsoon rainfall, ENSO and equatorial Indian Ocean oscillation, Geophysical Research Letters, 31: L12213. [doi:10.1029/2004GL019733].

6. Maity, R. and D. Nagesh Kumar, 2006. Bayesian dynamic modeling for monthly Indian summer monsoon rainfall using El Niño-Southern Oscillation (ENSO) and Equatorial Indian Ocean Oscillation (EQUINOO), Journal of Geophysical Research - Atmospheres, 111, D07104, [doi:10.1029/2005JD006539].
7. Maity, R. and D. Nagesh Kumar, 2006. Hydro-climatic association of monthly summer monsoon rainfall over India with large-scale atmospheric circulation from tropical Pacific Ocean and Indian Ocean, Atmospheric Science Letters, 7(4): 101-107.
8. Shukla, R.P., K.C. Tripathi, A.C. Pandey and I.M.L. Das, 2011. Prediction of Indian summer monsoon rainfall using Niño indices: A neural network approach, Atmospheric Research, 102(1-2): 99-109.
9. Kashid, S.S. and R. Maity, 2012. Prediction of monthly rainfall on homogeneous monsoon regions of India based on large scale circulation patterns using Genetic Programming, Journal of Hydrology, 454-455: 26-41.
10. Goswami, P. and Srividya, 1996. A novel neural network design for long range prediction of rainfall pattern, Current Science, 70: 447-457.
11. Guhathakurta, P., M. Rajeevan and V. Thapliyal, 1999. Long range forecasting Indian summer monsoon rainfall by a hybrid principal component neural network model, Meteorology atmospheric Physics, 71: 255-266.
12. Sahai, A.K., M.K. Soman and V. Satyan, 2000. All India summer monsoon rainfall prediction using an artificial neural network, Climate Dynamics, 16: 291-302.
13. Guhathakurta, P., 2006. Long-range monsoon rainfall prediction of 2005 for the districts and sub-division Kerala with artificial neural network, Current Science, 90: 773-779.
14. Chakraverty, S. and P. Gupta, 2008. Comparison of neural network configurations in the long-range forecast of southwest monsoon rainfall over India, Neural Computing with Applications, 17: 187-192.
15. Chattopadhyay, S. and G. Chattopadhyay, 2001. Identification of the best hidden layer size for three layered neural net in predicting monsoon rainfall in India, Journal of Hydroinformatics, 10(2): 181-188.
16. Singh, P. and B. Borah, 2013. Indian summer monsoon rainfall prediction using artificial neural network, Stochastic Environmental Research and Risk Assessment, 27(7): 1585-1599.
17. Karthikeyan, L. and D. Nagesh Kumar, 2013. Predictability of non-stationary time series using wavelet and EMD based ARMA models, Journal of Hydrology, 502: 103-119.
18. Nourani, V., A.H. Baghanam, J. Adamowski and O. Kisi, 2014. Applications of hybrid wavelet-Artificial Intelligence models in hydrology: A review, Journal of Hydrology, 514(6): 358-377.
19. Zhu S., J. Zhou, L. Ye and C. Meng, 2016. Streamflow estimation by support vector machine coupled with different methods of time series decomposition in the upper reaches of Yangtze River, China. Environmental Earth Sciences, 75:531, [doi: 10.1007/s12665-016-5337-7]
20. Napolitano, G., F. Serinaldi and L. See, 2011. Impact of EMD decomposition and random initialization of weights in ANN hindcasting of daily streamflow series: An empirical examination, Journal of Hydrology, 406(3-4): 199-214.
21. Hu J., J. Liu, Y. Liu and C. Gao, 2013. EMD-KNN model for annual average rainfall forecasting, Journal of Hydrologic Engineering, 18(11): 1450-1457.
22. Wang W.C., K.W. Chau, D.M. Xu and X.Y. Chen, 2015. Improving forecasting accuracy of annual runoff time series using ARIMA based on EEMD decomposition, Water Resources Management, 29(8): 2655-2675.
23. Zhao Xue-hu and X. Chen, 2015. Auto regressive and ensemble empirical mode decomposition hybrid model for annual runoff forecasting, Water Resources Management, 29(8): 2913-2926.
24. Huang, S., J. Cheng, Q. Huang and Y. Chen, 2014. Monthly streamflow prediction using modified EMD-based support vector machine, Journal of Hydrology, 511: 764-775.
25. Bai, Y., P. Wang, J. Xie, J. Li and C. Li, 2015. Additive model for monthly reservoir inflow forecast, Journal of Hydrologic Engineering, 10.1061/(ASCE)HE.1943-5584.
26. Iyengar, R.N. and T.S.G. Raghu Kanth, 2005. Intrinsic mode functions and a strategy for forecasting Indian monsoon rainfall. Meteorology and Atmospheric Physics, 90: 17-36.
27. Iyengar, R.N. and S.T.G. Raghu Kanth, 2006. Forecasting of seasonal monsoon rainfall at subdivision level, Current Science, 91(3): 350-356.
28. Huang, N.E., Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung and H.H. Liu, 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proceedings of Royal Society London, Series A, 454:903-995.

29. Rehman, N. and P.P. Mandic, 2010. Multivariate empirical mode decomposition, Proceedings of Royal Society London, Series A, 466: 1291-1302.
30. Hu, W. and B.C. Si, 2013. Soil water prediction based on its scale-specific control using multivariate empirical mode decomposition, Geoderma, 193-194: 180-188.
31. Huang, G., Y. Su, A. Kareem and H. Liao, 2016. Time-frequency analysis of non-stationary process based on multivariate empirical mode decomposition, Journal of Engineering Mechanics, ASCE, 142(1), [doi: 10.1061/(ASCE)EM.19437889.0000975]
32. Draper, N.R. and H. Smith, 1998. Applied Regression Analysis, Hoboken, NJ: Wiley-Interscience, pp: 307-312.
33. Rilling, G., P. Fladrin and P. Goncalves, 2003. On empirical mode decomposition and its algorithms, In Proceedings of IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP-03, Grado (Italy), pp: 8-11.
34. Wilks, D.S., 1995. Statistical methods in atmospheric sciences. Academic Press, New York.
35. Kashid, S.S., S. Ghosh and R. Maity, 2010. Streamflow prediction using multi-site rainfall obtained from hydroclimatic teleconnection, Journal of Hydrology, 395: 23-38.