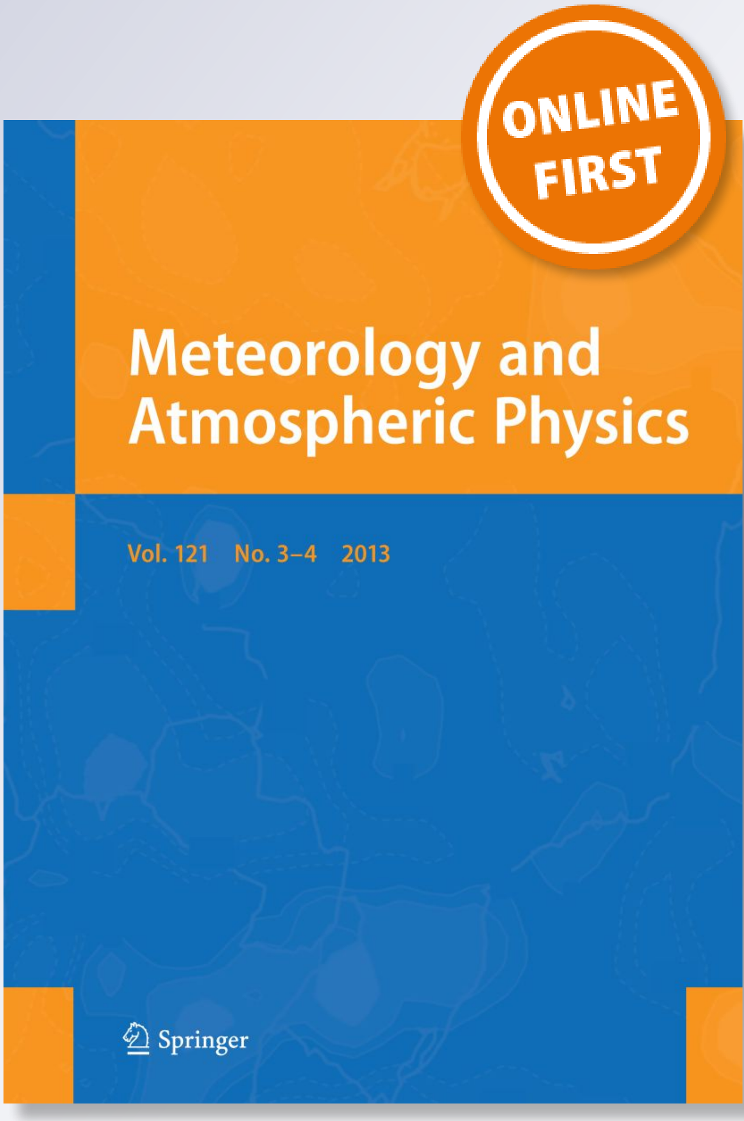
Empirical forecasting and Indian Ocean dipole teleconnections of south–west monsoon rainfall in Kerala

Kavya Johny, Maya L. Pai & S. Adarsh

Meteorology and Atmospheric Physics

ISSN 0177-7971

Meteorol Atmos Phys
DOI 10.1007/s00703-018-0620-7



Meteorology and Atmospheric Physics
https://doi.org/10.1007/s00703-018-0620-7

ORIGINAL PAPER

Empirical forecasting and Indian Ocean dipole teleconnections of south–west monsoon rainfall in Kerala

Kavya Johny[1] · Maya L. Pai[1] · S. Adarsh[2]

Received: 26 November 2017 / Accepted: 25 June 2018

Abstract

Rainfall is a vital hydrologic variable that has a direct and significant impact on the economic development of monsoon-dominated state of Kerala in southern India. An effective approach providing accurate prediction of rainfall makes it possible to take preventive and mitigation measures against natural disasters. In this study, the ensemble empirical mode decomposition (EEMD)–artificial neural networks (ANN)–multiple linear regression (MLR) hybrid approach is used to forecast the south–west monsoon (SWM) rainfall of Kerala. The EEMD of SWM rainfall of Kerala resulted in a set of orthogonal components of specific periodic scale. The non-linear components are identified and separately modeled using ANN and rest of the components are modeled using linear regression to get their values at a specific time *t*. Finally, the predicted modes are recombined to get the forecasts of a generic time *t*. The SWM rainfalls of 1871–1972 are used for model calibration and forecasts are made sequentially for 1973–2014 period, which clearly demonstrated its efficacy in handling non-linear part of SWM rainfall data with a predictive skill of 0.65 for validation data. Further, by considering a dataset of 1961–2014 period, this study has investigated the possible teleconnection of SWM rainfall of Kerala with the Indian Ocean dipole (IOD) using the cross-correlation and EEMD-based time-dependent intrinsic correlation (TDIC) analyses. Apart from the strong correlation in the trend component, the analysis has proved the dominancy of negative association of IOD with SWM of Kerala in different process scales with strong positive association at localized time spells. The forecasting strategy demonstrated in the study and the evidence of IOD–SWM rainfall link are an amendment to the efforts for improving the predictability of SWM rainfall in Kerala.

1 Introduction

Rainfall is an important hydrologic variable that plays a significant role in the water resources planning and management of India. In the monsoon-dominated states like Kerala located in southern India, rainfall plays a vital role in agriculture, power generation, travel and tourism and hence, significantly contributes to the economic development of the state. The major contributor of rainfall in Kerala is from Indian Summer Monsoon (falling in June–September months) and its accurate forecasting is important for the preparedness against natural disasters like floods and droughts, but it is a challenging task to the modelers. Therefore, any attempt for improvement in predictions needs to be encouraged. Numerous techniques have been used by the researchers in the past for the prediction of monsoon rainfall of India, which include simple statistical methods, clustering techniques, data-driven methods and hybrid approaches (Sahai et al. 2000; Abraham et al. 2001; Afshar and Fahimi 2012; Kashid and Maity 2012; Singh and Borah 2013; Beltran-Castro et al. 2013; Pai et al. 2014). The time series or data-driven techniques have received wide popularity in prediction of Indian summer monsoon rainfall (ISMR) due to the unknown and complex relationships between different variables relating to rainfall, the presence of non-linearities and instabilities in the rainfall events of ISMR, flexibility

Responsible Editor: A.-P. Dimri.

Kavya Johny
kavyajovita@gmail.com

Maya L. Pai
mayalpai@gmail.com

S. Adarsh
adarsh_lce@yahoo.co.in

[1] Department of Computer Science and IT, Amrita School of Arts and Sciences, Amrita Vishwa Vidyapeetham, Kochi, India

[2] Department of Civil Engineering, TKM College of Engineering, Kollam, India

Published online: 03 July 2018

of advanced forecasting techniques, etc (Sahai et al. 2000). For the effective implementation of time series models, identification of the most appropriate predictor variables is crucial. Numerous climatic oscillations such as ElNiño Southern Oscillation (ENSO), Quasi Biennial Oscillation (QBO), Atlantic Multidecadal Oscillation (AMO), and Equatorial Indian Ocean Oscillation (EQUINOO) are some among them (Walker 1923; Kripalani and Kulkarni 1997a, b; Krishna kumar et al. 1999; Kumar et al. 2006; Goswami et al. 2006; Gadgil et al. 2004).

The identification of periodic scales is important in ISMR predictions and in the past, different researchers reported the signatures of both intra- and inter-annual variability and their links with global climate system in the ISMR series (Krishnamurthy and Kinter 2003; Iyengar and Raghu Kanth 2005). Mini et al. (2016) analyzed the rainfall data of Kerala over 1871–2011 period at annual, monthly and seasonal scale and reported strong inter-annual variability, reduction in monsoon rainfall an increase in pre- and post-monsoon rainfall of the state. For the identification of periodic components, multi-scale decomposition methods are suitable alternatives. Iyengar and Raghu Kanth (2005) used empirical mode decomposition (EMD) as a potential tool and described the link of Quasi Biennial Oscillation, ENSO, sunspot cycles, tidal forcing, etc., with ISMR by the comparison of their periodicities. Recently Adarsh and Janga Reddy (2017) used a noise-assisted variant of EMD and established such links through comparison of the periodicities and multi-scale correlation analysis. Even though the teleconnections of ISMR with the different climatic variables are expressed best for large spatial scales, the links are evident at subdivisional or river basin scale, studies of which have been performed by some investigators (Kashid and Maity 2012; Adarsh and Janga Reddy 2016). The potential of multi-scale decomposition is eventually helpful for improved forecasting of ISMR and Iyengar and Raghu Kanth (2005) presented a hybrid strategy involving artificial neural networks (ANN) and EMD for the same and applied for Karnataka subdivision in India (Iyengar and Raghu Kanth 2006). The state of Kerala is popularly known as the 'gateway of Indian monsoon' and its rainfall predictions are of particular importance and always watched carefully by the scientific community. Long-range or short-term seasonal predictions of rainfall in Kerala using different techniques have been attempted by the researchers in the past using different methods (Abraham et al. 2001; Guhathakurta 2006).

This study uses the noise-assisted variant of EMD, namely, ensemble EMD (EEMD) coupled with ANN as strategy for forecasting south–west monsoon rainfall in Kerala meteorological subdivision. The rest of the paper is organized as follows: Sect. 2 presents the theoretical background of the methodology; Sect. 3 presents the details of dataset used for the study; Sect. 4 presents results and discussion. It is followed by the final section which presents important conclusions of the study.

2 Materials and methods

2.1 Empirical mode decomposition

The empirical mode decomposition (Huang et al. 1998) is a data-adaptive signal decomposition technique gaining popularity for analyzing non-linear and non-stationary data. The decomposition process flow primarily involves the stages such as empirical fitting of splines through extrema points (maxima and minima points), computation of mean profile of spline fittings through extrema, computation of difference between original series and mean profile to produce a zero mean series called intrinsic mode functions (IMFs), etc. A typical IMF should satisfy two conditions: (1) in the whole data set, the number of extrema and the number of zero crossings must be either equal or differ at the most by one; and (2) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero. The process of generation of an IMF is executed over several iterations (so-called 'sifting' operation). Subsequently, on subtracting the evolved IMF from the original series, the candidate series for the generation of IMF2 is obtained. The process of generation of IMFs is continued till a typical mode evolved contains only one extrema or it is monotonic in nature. EMD replaces traditional methods for analyzing time series signals like wavelet analysis and Fourier analysis in the sense that this method does not use any mathematical functional form in the decomposition operation. More details on the algorithm can be found elsewhere (Huang et al. 1998; Rao and Hsu 2008).

2.2 Ensemble empirical mode decomposition (EEMD)

The decomposition operation of a time series is said to be effective only when IMFs are well-separable in their frequency contents; but if an IMF obtained by EMD possesses the characteristics of different frequency content (or on the other hand, the process scales get separated in different modes), the physical interpretation of the signals become very difficult and eventually it may lead to completely wrong conclusions. Over the years different researchers proposed noise-assisted and ensemble-averaged variants of EMD to overcome this serious drawback so-called 'mode mixing' and one such variant proposed by Wu and Huang (2005) called as ensemble empirical mode decomposition (EEMD) became popular as it is found to be effective in decomposition of non-linear and non-stationary time series signals. The typical steps of iterative procedure of EEMD are: (1)

addition of a white noise series to the given time series signal to form an ensemble of artificial signals; (2) decompose each artificial signal by EMD to evolve an IMF; (3) perform ensemble averaging to get the desired IMF.

2.3 Time-dependent intrinsic correlation

Time-dependent intrinsic correlation (TDIC), proposed by Chen et al. (2010) is a running correlation analysis method for determination of association of two correlated time series in different time scales. The method is a data-adaptive technique developed based on the HHT. In this method the following steps are involved:

1. Decompose the two correlated time series using EMD or its variants
2. Determine the mean periodicities of orthogonal components of the two series and identify the modes of similar periodicity
3. Perform HT of identified modes to obtain the instantaneous frequency (and hence instantaneous periods)
4. Fix the minimum sliding window size (t_d) as maximum instantaneous period between the two signals at the current position t_k, i.e., $t_d = \max(T_{1,i}(t_k), T_{2,i}(t_k))$, where $T_{1,i}$ and $T_{2,i}$ are instantaneous periods.
5. Estimate the size of sliding window as $t_w^n = \left[t_k - \frac{nt_d}{2} : t_k + \frac{nt_d}{2}\right]$, where n is any positive number (a multiplication factor for minimum sliding window size) and normally n is selected as 1 (Huang and Schmitt 2014)
6. Determine the correlation between the two modes for different sliding windows and estimate their statistical significance using t test. Proceed iteratively till the boundary of the sliding window exceeds the end points of the time series

The results of TDIC analysis are obtained in the form of a TDIC matrix, and its plot will be triangular in shape, with time (the centre points of sliding window) in the horizontal axis and the size of the sliding window as vertical axis. The color bar depicts the instantaneous correlation between the two modes. The correlation at the apex point will be the correlation coefficient between the modes considering the complete data length, if the data length is chosen as the maximum window size (Chen et al. 2010).

2.4 Methodology of rainfall prediction

A variant of rainfall prediction proposed by Iyengar and Raghu Kanth (2005) is used in this study for prediction of SWM rainfall of Kerala. The different steps in the rainfall prediction of Kerala subdivision using EEMD–ANN–MLR is shown as a flowchart in Fig. 1. First, the data are decomposed into empirical modes and a final residue using EEMD. The histograms are used to check the normality of each mode that follows Gaussian distribution. Based on the normality tests, the linear and non-linear parts are separated. The non-linear part of the signal (R) is predicted separately using the ANN. A non-linear autoregressive neural network is used for forecasting of k $(t+1)$ from the time series where k represents the non-linear IMF and t represents the time period. The partial auto-correlation function (PACF) is used to find the significant inputs required for prediction of k $(t+1)$. The linear part l_{n+1} is predicted using the multiple linear regression using the equation:

$$l_{n+1} = C_1R_n + C_2l_{n-1} + C_3l_{n-2} + C_4l_{n-3} + C_5l_{n-4} + C_6,$$

where $l_j = R_j - k_j$, k_j represents the sum of non-linear IMFs at time j and Ci represents the regression coefficients. The final forecast is the sum of the individual forecasts of linear and non-linear parts.

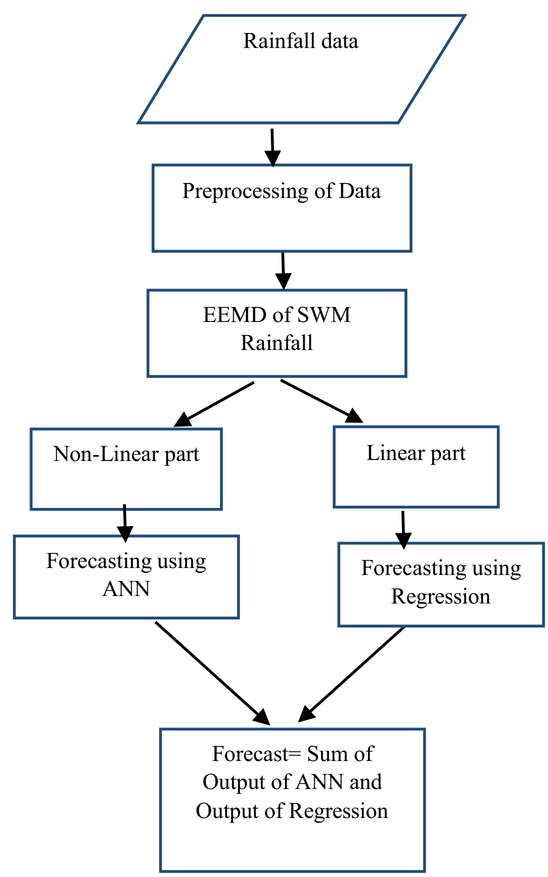


Fig. 1 Flowchart of EEMD–ANN–MLR method for SWM rainfall forecasting

3 Data

Indian Institute of Tropical Meteorology (IITM) Pune defined 36 meteorological subdivisions in India based on rainfall homogeneity. The monthly area weighted precipitation considering representative rainfall stations (at least one per district) have been prepared first in 1994 (Parthasarathy et al. 1994). Updating of this database has been carried out several times by the scientists of IITM Pune and the monthly data of Kerala meteorological subdivision (Lat. 8.29°N–11.52°N and Lon. 75.22°E–76.57°E) for 143 years (1871–2014 period) are extracted from http://www.tropmet.res.in. Then the south–west monsoon (SWM) seasonal rainfall of the subdivisions is computed by aggregating the monthly rainfalls of June, July, August and September months. To investigate the possible teleconnection of SWM rainfall of Kerala with the Indian Ocean dipole (IOD), the IOD index of 53 years (1961–2014) is accessed from IRI/LDEO Climate Data Library (http://iridl.ldeo.columbia.edu).

4 Results and discussion

The SWM rainfall data of Kerala subdivision is first decomposed using EEMD by selecting control parameters such as noise standard deviation, ensemble number and number of sifting iterations as 0.02, 100 and 10, respectively. The EEMD package in MATLAB platform available at (http://perso.ens-lyon.fr/patrick.flandrin/emd.html) is used for the above procedure after making necessary modifications. The EEMD of SWM rainfall series resulted in five IMFs and a final residue. The modes obtained by EEMD along with their probability density functions (PDFs) are presented in Fig. 2. The basic properties of the modes obtained are analyzed in detail. The mean period of each of the signals is calculated by counting the number of zeroes and extrema. The mean period of modes, % variability explained by different modes, correlation of different modes with original signal among different modes, etc., are presented in Table 1.

From Table 1, it is noted that IMF1 is having a process scale of 2.95 years which is typical of Quasi Biennial Oscillation (with period of two and half years approximately). Similarly, IMF2, IMF3 and IMF4 show the periodicities close to phenomena or oscillations like ENSO, sunspot cycle and tidal forcing, respectively. The residue shows a

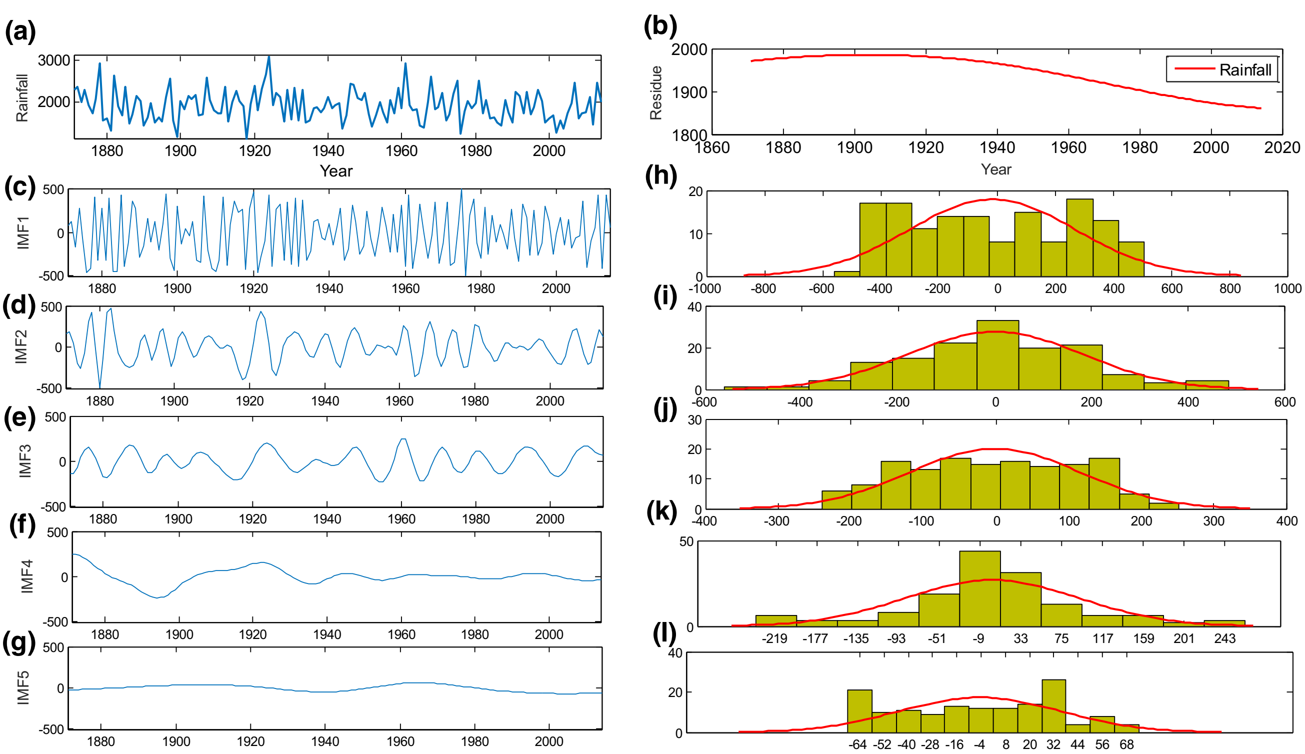


Fig. 2 Orthogonal modes of SWM rainfall of Kerala (1871–2014) and the histograms of IMFs. **a** The rainfall time series and **b** the corresponding residues. **c–g** The different IMFs and **h–l** the respective histograms

Table 1 The mean period (in years), variance of modes and correlation between modes of SWM rainfall of Kerala

Mode number	VE (%)	Mean period (years)	Correlation coefficients					
			Data	IMF1	IMF2	IMF3	IMF4	IMF5
Data	–	–	1	0.7182	0.506	0.336	0.257	0.082
IMF1	60.19	2.95		1	−0.054	−0.059	−0.007	−0.045
IMF2	23.33	6.42			1	0.164	0.015	−0.018
IMF3	10.17	12.10				1	0.029	−0.106
IMF4	5.13	21.11					1	−0.029
IMF5	1.17	61						1

maximum rainfall during a period between 1910 and 1935 and shows a decrease in the amount of rainfall after 1940 which is evident in the panel (b) of Fig. 2. The residue displays a slowly varying pattern that limits the monsoon rainfall throughout the years from 1871 to 2014. Figure 2 shows that IMF1, IMF3 and IMF5 are non-Gaussian (with multimodality), which rejects the null hypothesis that the signals follow a normal distribution. The remaining IMFs are unimodal. The bimodal and multimodal behavior indicates very strong non-linear characteristics of the signals in dynamics of their related process (Fan and Yao 2003).

For modeling SWM rainfall of Kerala the procedure explained in Sect. 2.4 is invoked. To demonstrate the efficacy of the proposed hybrid approach (which accounts non-linearity separately), two different models are considered. In the first model (model 1) IMF1 and IMF5 are considered as non-linear part and rest of the modes are considered as linear part. In the second model (model 2) in addition to IMF1 and IMF5, IMF3 is also considered as non-linear part (which is evident from histogram plots presented in Fig. 2), and IMF2, IMF4 and residue are considered as linear part. In both models, first 70% of the data are used for model calibration and remaining 30% of the data are predicted sequentially (with one step ahead at a time). The selection of appropriate inputs is crucial in time series modeling. So to identify the proper inputs, Partial Auto-Correlation (PAC) analysis of different modes is performed. In this analysis, we assume that x_i as the output variable on the condition that the PAC at lag k is out of the $[-1.96/\sqrt{n}, 1.96/\sqrt{n}]$ range, where n is the data length at 95% confidence interval and x_{i-k} is one of the input variables. If none of the PAC coefficients is not in the 95% confidence interval, the previous one value x_{i-1} is selected as the input variable and the formulas for computation of PAC coefficients can be found elsewhere (Huang et al. 2014). The results of PAC analysis are presented in Fig. 3. Figure 3 show that two lagged values are sufficient to model IMF1,

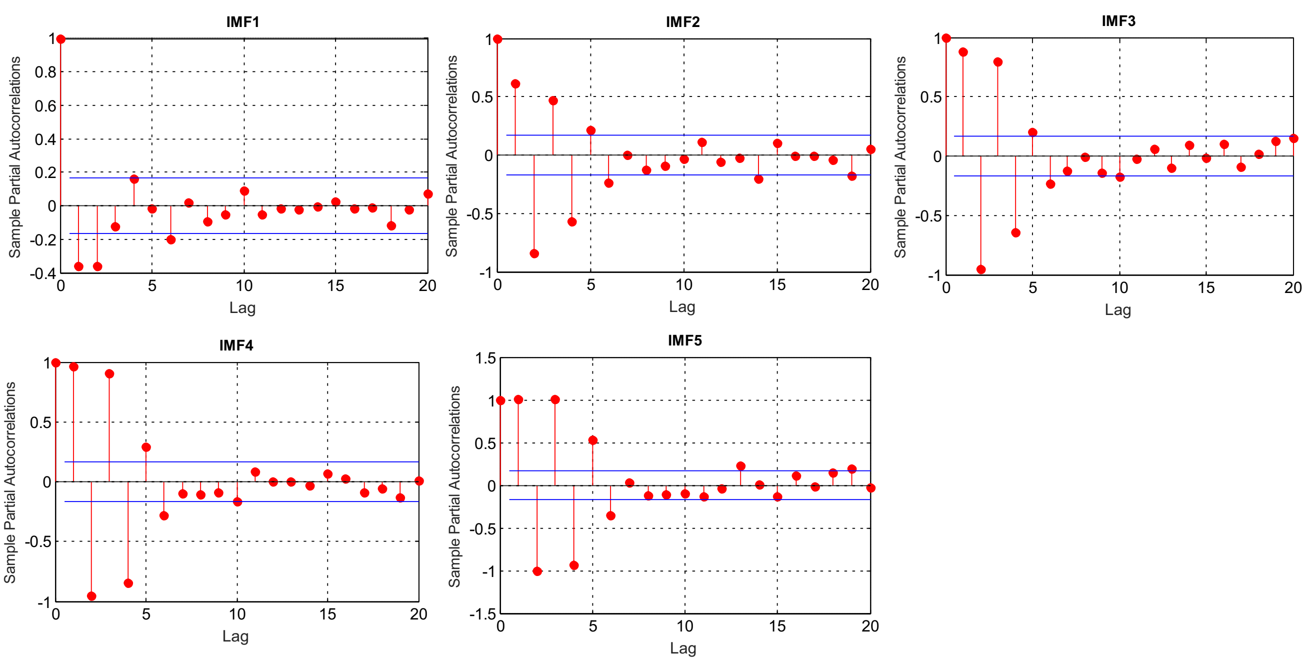


Fig. 3 Partial auto-correlation plots of different IMFs of SWM rainfall of Kerala

while five past values are required to get the value of mode IMF5 and in the case of the latter model, the IMF3 requires six past values to get the predictions of a generic time *t*. Hence the predictions are available from 1876 for the former model and from 1877 onwards for the latter model. The non-linear part (IMF1 and IMF5 in model 1; IMF1, IMF3, IMF5 in model 2) is modeled using non-linear autoregressive neural network. A single hidden layer of ten hidden neurons; sigmoid transfer function is used to train the network. The Levenberg–Marquardt (LM) algorithm is used for model calibration, considering the datasets of 1871–1972.

Finally, the predicted values of IMF1, IMF5 and linear part are combined to get one step ahead forecast of SWM rainfall of Kerala by model 1. In model 2, the forecast is made by combining the predicted values of IMF1, IMF3, IMF5 and linear part. The forecasting is sequentially continued to get the rainfall of 1973–2014 period. The performance evaluation of forecasts is done by different statistical measures. In addition to the common measures such as correlation coefficient (R) and mean absolute percentage error (MAPE), the mean square skill score (MSSS) and normalized root mean square error (NRMSE) are also computed.

$\mathrm{MSSS} = 1 - \frac{\mathrm{MSE}}{\mathrm{MSE_c}}$,in which $\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(x_{\mathrm{o}i} - x_{\mathrm{f}i})^2$ and $\mathrm{MSE_c} = \frac{1}{n}\sum_{i=1}^{n}(x_{\mathrm{o}i} - \bar{x})^2$, where MSE is the mean square error of forecasts and $\mathrm{MSE_c}$ is mean square error of climatology; $x_{\mathrm{o}i}$ is the observed rainfall, $\bar{x}$ is the mean rainfall; $x_{\mathrm{f}i}$ is the forecasted rainfall and n is the number of observations. It gives a measure how best are the predictions over the mean and a positive value indicates acceptable predictions with a value close to unity indicating the best predictive skill. The NRMSE is the normalized version of RMSE stated by Karthikeyan and Nagesh Kumar (2013) computed as:

$$\mathrm{NRMSE} = \frac{\sqrt{\frac{\sum_{i=1}^{n}(x_{\mathrm{o}i} - x_{\mathrm{f}i})^2}{n}}}{(x_{\max} - x_{\min})},$$

where $x_{\max}$ and $x_{\min}$ are maximum and minimum of observed values, respectively. Here RMSE is normalized with reference to the range of observed values to control the range of error outcomes. An NRMSE value of greater than one means that the average root of squared deviation of predicted values from the observed (numerator) is greater than the range of observed values (denominator) suggesting poor model performance (Karthikeyan and Nagesh Kumar 2013). Ideally the error value should be 0. The results of performance statistics are summarized in Table 2.

From Table 2, it is evident that model 2 performs better than model 1 (with better R and MSSS values and low error statistics like NRMSE and MAPE) which clearly infer that the inclusion of IMF3 as a non-linear part in modeling improves the predictability of SWM rainfall of Kerala. Thus, it is evident that the proposed hybrid modeling approach is capable of accounting the effect of non-linear part in SWM rainfall predictions. The plots of observed and predicted values of rainfall for training and validation are presented in Fig. 4.

Table 2 Performance evaluation of predictions

Evaluation criteria	Model 1		Model 2	
	Calibration	Validation	Calibration	Validation
R	0.666	0.543	0.716	0.653
MSSS	0.442	0.287	0.51	0.389
NRMSE	0.145	0.22	0.135	0.204
MAPE	11.82	13.52	11.06	12.20

The method used in the study received a prediction skill of 0.65 for validation data (for model 2). Recalling that typical ISMR predictions based on physically based empirical models and IMD operational forecasts display a prediction skill of 0.51 for ISMR predictions of 1989–2012 (Wang et al. 2015; Adarsh and Janga Reddy 2017) period, this skill score is encouraging for SWM rainfall forecasting. Despite the availability of many physical models for operational forecasts, prediction of ISMR is challenging owing to the complexity of physical processes of monsoon, influence of large number of parameters, presence of non-linearities and instabilities of the observed time series, etc. At the same time, the use of physical models is complex, computationally expensive, time consuming and involves huge data requirement, which offers difficulties to the modelers. The time series and statistical approaches can be chosen as suitable alternatives for forecasting owing to the simplicity n implementation, as fairly stable relationship exists between ISMR and some of the dominant predictors. Here it is believed that if there are connections between ISMR and different predictors, all such information may be embedded in the time series itself and it is quite logical to follow a time series approach. Also, the inclusion of appropriate predictor variables (which are often climate indicators or oscillations) may improve the predictability efforts of rainfall. In this context, in the next section the influence of IOD on SWM rainfall is investigated, using the HHT-based running correlation analysis.

4.1 IOD teleconnections of SWM rainfall of Kerala

In the equatorial zone, Quasi Biennial Oscillation (QBO) of the wind flows from easterlies to westerlies in the tropical stratosphere with an average period of 28–29 months. ENSO consists of ElNiño and LaNiña which is a phenomenon that recurrently occurs within a period that ranges from 2 to 10 years (Varikoden and Preethi 2013). The sunspot cycle (time period of which is 11 years on an average), is the period within which a variation has been seen in the

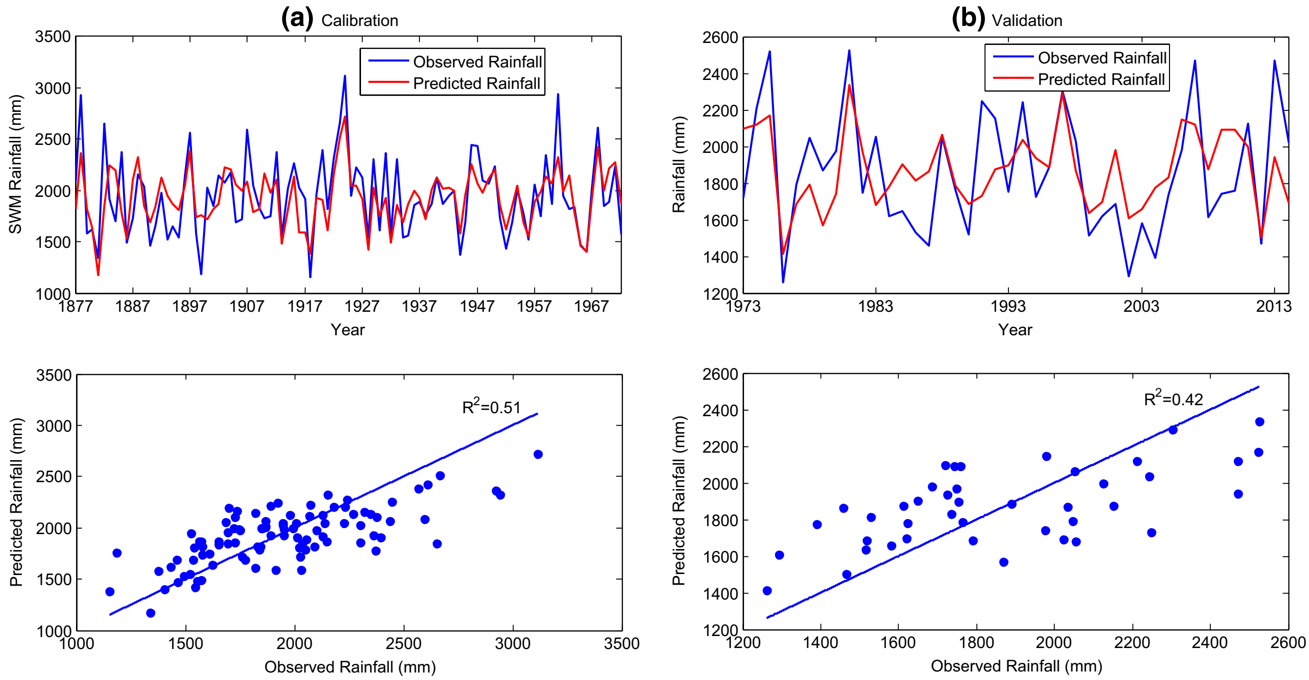


Fig. 4 Predictions of SWM rainfall of Kerala for calibration period (1877–1972) and validation period (1973–2014). Upper panels show the time series plots and lower panels show scatter plots

amount of magnetic flux that rises to the sun's surface. The tidal force is an important factor for the monsoon rainfall in India extends from 20 to 24 years. Many of the past studies identified the influence of these factors in prediction of ISMR at All India or regional scales (Adarsh and Janga Reddy 2016, 2017). It is well-proven that the effect of Indian Ocean Dipole (IOD) is crucial in the variability of ISMR (Saji et al. 1999; Bala and Singh 2008; Krishnaswamy et al. 2015). IOD is the difference in sea surface temperatures (SST) of Arabian Sea (AS) and Bay of Bengal (BoB). IOD has a positive value when the SST of AS is more than BoB and vice versa. During the months from June to August the south–west winds flowing from high pressure to low pressure area create a trough that induces a rough weather on Arabian Sea. It is quite logical to hypothesize its powerful impact on the rainfall in the coastal regions like Kerala. There are three types of IODs: early IOD, normal IOD and prolonged IOD. The early IOD peaks in the month of July to August and thereby plays a vital role in the monsoon rainfall. The normal IOD and prolonged IOD peak in the months of September–November.

To investigate the link between IOD and SWM of Kerala, first the monthly monsoon rainfall and IOD Indices of monsoon period are decomposed using EEMD method. The decomposition resulted in seven modes for SWM rainfall of Kerala and six modes for IOD and the modes are presented in Fig. 5. The mean periods of orthogonal modes of SWM rainfall and IOD along with the % variability explained by the models are presented in Table 3. The modes are presented in Fig. 6 to get better insight on their association. From Fig. 6 it is noted that there is considerable difference in the pattern of the modes of IMF4 and IMF5 (with less number of cycles in the mode of IOD, which is also reflected in the mean period presented in Table 3). Thus, to identify the true association, IMF4 of SWM rainfall is compared with the IMF3 of IOD; similarly, IMF5 of SWM rainfall is compared with IMF4 of IOD. Thus, panels (f) and (g) show the resemblance between the two in many of the short time spells/windows (associated with some phase shift in certain cases). This observation is important in explaining the association between the two series. Further by considering the complete data length, the correlation between the modes of IOD and that of SWM rainfall is computed. The correlation values are presented in Table 4. From Table 4 it is clear that the two are strongly associated in the low frequency mode only, which is a typical character of climatic teleconnections, but the plots show that one cannot ignore the strong association between the two series in shorter time spells. To capture such masked associations, the TDIC analysis is performed. From Table 3 it is clear that the first four modes are having similar periodicity and first these modes are selected for the in-depth TDIC analysis, to get more insight into their association at different process scales. Then the links IMF4 (SWM)–IMF3 (IOD) and IMF5 (SWM)–IMF4 (IOD) (which are having similar periodicity) are also investigated. The different TDIC plots are presented in Fig. 7. Figure 7 shows that the correlation between IOD and SWM of Kerala is dominantly

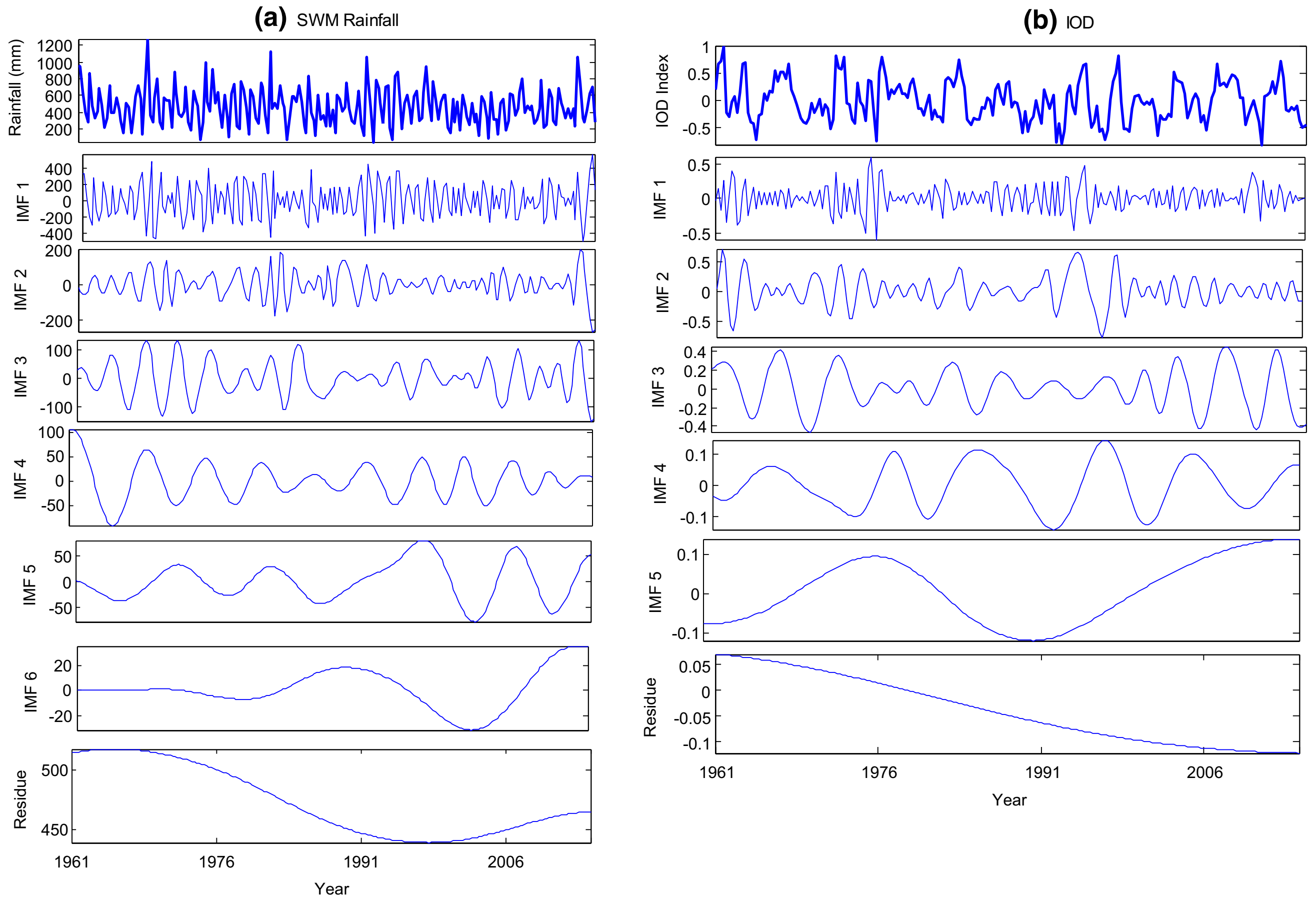


Fig. 5 Orthogonal modes of monthly SWM rainfall of Kerala and monsoon IOD index for 1961–2014 period **a** SWM rainfall; **b** IOD index

Table 3 Percentage variability explained and mean period (in years) of modes of SWM rainfall of Kerala and the IOD index

Mode number	% Variability explained		Mean period (years)	
	SWM rainfall	IOD	SWM rainfall	IOD
IMF1	78.49	21.79	1.11	1.01
IMF2	8.78	37.65	2.06	2.48
IMF3	6.30	30.48	3.79	5.54
IMF4	2.34	3.33	6.55	12.00
IMF5	2.29	4.23	18.00	72.00
IMF6/residue	0.43	2.53	24.00	72.00
Residue	1.37		72.00	

negative in different process scales and the association is best reflected in IMF3 and IMF4 (3–12 years of periodicity from Table 3). Recalling that typical periodicity of ENSO is 3–8 years, this observation is quite convincing. However, there are localized positive association in different process scales (i.e., for different IMFs) and this dynamics is quite rich in IMF1. In IMF3 (after 2006), IMF4 (SWM) and in IMF3 (IOD) also such transitions are noticeable. This transition in the nature of correlations (from negative to positive or vice versa in time domain) might be due to the influence of local perturbations or other large-scale climatic oscillation processes which modulate the SWM–IOD links and apart from statistical analysis; a proper meteorological analysis accounting the physical process is essential to identify the exact reason behind such switchovers in correlations. The patterns noticed in Fig. 7 reveal that there is a connection between SWM Kerala rainfall and IOD in different process scales. It can be concluded that a change in monsoon IOD is producing proportionately more impacts on the SWM Kerala rainfall. Hence capturing information on the SWM–IOD link from different process scales may eventually help in improving the predictability efforts of SWM rainfall of Kerala.

5 Conclusion

The prediction of SWM rainfall is very important for planning and devising different strategies for the economic growth of the monsoon-dominated state like Kerala. In this study, a hybrid strategy involving EEMD, ANN and MLR

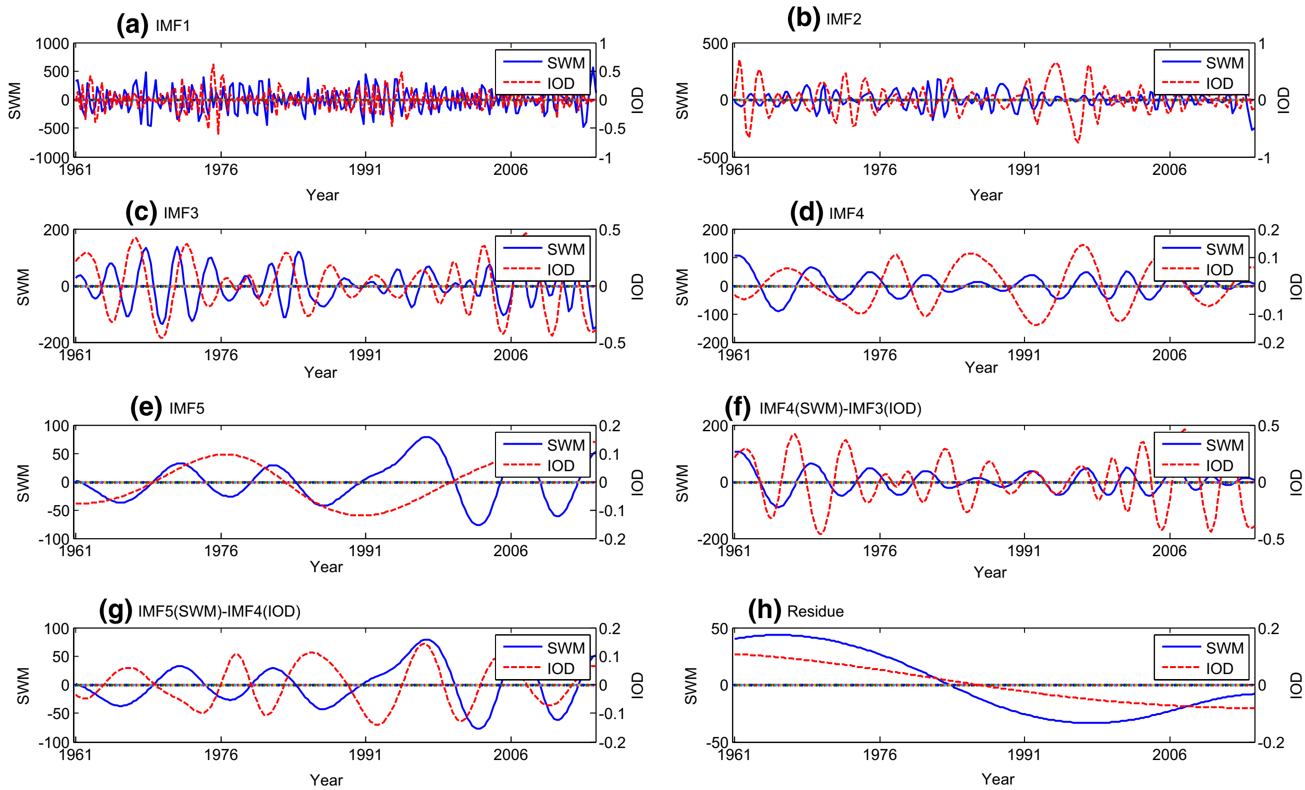


Fig. 6 Comparison between orthogonal modes of SWM rainfall and monsoon IOD index. **a–d** The comparison of modes of same index number. **f** The comparison between IMF4 of SWM rainfall with IMF3 of IOD. **g** The comparison between IMF5 of SWM rainfall with IMF4 of IOD. **h** The comparison between the residues of SWM rainfall and IOD

Table 4 Correlation between modes of IOD and modes of SWM rainfall of Kerala

Modes of IOD	Modes of SWM rainfall						
	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	Residue
IMF1	0.042	0.015	−0.015	0.018	−0.056	0.016	−0.013
IMF2	0.027	−0.103	−0.036	−0.033	−0.014	0.027	−0.009
IMF3	−0.002	0.049	0.029	−0.070	0.040	−0.069	0.045
IMF4	−0.006	0.005	−0.067	−0.217	0.060	−0.033	−0.049
IMF5	−0.008	−0.045	−0.019	−0.033	−0.031	−0.056	0.076
Residue	−0.032	0.016	0.031	0.052	−0.139	−0.066	0.916

is presented for forecasting of SWM rainfall of Kerala, in which EEMD is used for decomposition of rainfall time series. The non-linear part is modeled using ANN, linear part by multiple linear regression and the final recombination of the predicted modes enables the prediction of SWM rainfall. The proposed hybrid approach involving ANN is found to be efficient in handling the non-linear part of the SWM rainfall data and displayed a predictive skill of 0.65, which is encouraging on considering complex dynamics of Indian summer monsoon system. Moreover, this study used EEMD-based TDIC analysis for investigating the teleconnection of SWM of Kerala with IOD. The TDIC analysis showed that both the strength and nature of association between the SWM of Kerala and IOD differ with the process scale and over the time domain, with a dominancy of negative association in their link and localized reversals in the nature of association.

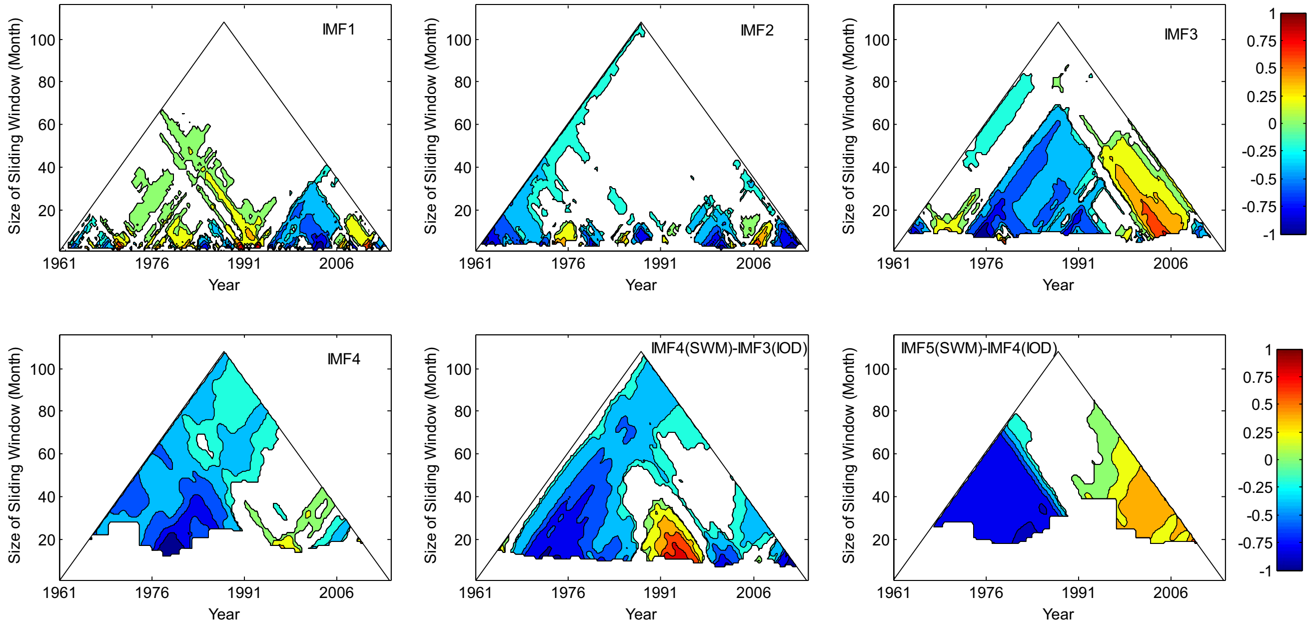


Fig. 7 TDIC plots of modes of SWM rainfall of Kerala and that of monsoon IOD index

Acknowledgements The authors thank the Indian Institute of Tropical Meteorology Pune and Climate Data Library for the providing the necessary data in the public domain with the intention of promotion of non-commercial scientific research. The authors also thank Amrita School of Arts and Sciences for the facilities provided.

References

Abraham A, Steinberg D, Philip NS (2001) Rainfall forecasting using soft computing models and multivariate adaptive regression splines. IEEE SMC Trans Spec Issue Fusion Soft Comput Hard Comput Ind Appl 1:1–6

Adarsh S, Janga Reddy M (2016) Analyzing the hydro-climatic teleconnections of summer monsoon rainfall in Kerala, India using multivariate empirical mode decomposition and time dependent intrinsic correlation. IEEE Geosci Remote Sens Lett 13(9):1221–1225

Adarsh S, Janga Reddy M (2017) Multiscale characterization and prediction of monsoon rainfall in India using Hilbert Huang transform and time dependent intrinsic correlation analysis. Meteorol Atmos Phys. https://doi.org/10.1007/s00703-017-0545-6

Afshar NR, Fahimi H (2012) Rainfall forecasting using Fourier series. J Civil Eng Archit 6(9):1258–1262

Bala I, Singh OP (2008) Relationship between Indian Ocean dipole and summer monsoon. Mausem 59(2):167–172

Beltran-Castro J, Valencia-Aguirre J, Orozco-Alzate M, et al (2013) Rainfall forecasting based on ensemble empirical mode decomposition and neural networks. In: Rojas I, Joya G, Gabestany J (eds) Advances in computational intelligence. IWANN 2013. Lecture Notes in Computer Science, vol 7902. Springer, Berlin, Heidelberg, pp 471–480. https://doi.org/10.1007/978-3-642-38679-4_47

Chen X, Wu Z, Huang NE (2010) The time-dependent intrinsic correlation based on the empirical mode decomposition. Adv Adapt Data Anal 2:233–265

Fan J, Yao Q (2003) Nonlinear time series: nonparametric and parametric methods. Springer, New York

Gadgil S, Vinayachandran PN, Francis PA, Gadgil S (2004) Extremes of the Indian summer monsoon rainfall, ENSO and equatorial Indian Ocean oscillation. Geophys Res Lett 31:L12213. https://doi.org/10.1029/2004GL019733

Goswami BN, Madhusoodanan MS, Neema CP, Sengupta D (2006) A physical mechanism for North Atlantic SST influence on the Indian summer monsoon. Geophys Res Lett 33:L02706. https://doi.org/10.1029/2005GL024803

Guhathakurta P (2006) Long-range monsoon rainfall prediction of 2005 for the districts and sub-division Kerala with artificial neural network. Curr Sci 90(6):773–779

Huang Y, Schmitt FG (2014) Time dependent intrinsic correlation analysis of temperature and dissolved oxygen time series using empirical mode decomposition. J Mar Syst 130:90–100

Huang S, Chang J, Huang Q, Chen Y (2014) Monthly streamflow prediction using modified EMD-based support vector machine. J Hydrol 511:764–775

Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Yen NC, Tung CC, Liu HH (1998) The empirical mode decomposition and the Hilbert spectrum for non-linear and non-stationary time series analysis. Proc R Soc Lond Ser A454:903–995

Iyengar RN, Raghu Kanth TSG (2005) Intrinsic mode functions and a strategy for forecasting Indian monsoon rainfall. Meteorol Atmos Phys 90:17–36

Iyengar RN, Raghu Kanth TSG (2006) Forecasting of seasonal monsoon rainfall at subdivision level. Curr Sci 91(3):350–356

Karthikeyan L, Nagesh Kumar D (2013) Predictability of nonstationary time series using wavelet and EMD based ARMA models. J Hydrol 502:103–119

Kashid SS, Maity R (2012) Prediction of monthly rainfall on homogeneous monsoon regions of India based on large scale circulation patterns using genetic programming. J Hydrol 454:26–41

Kripalani RH, Kulkarni A (1997a) Climatic impacts of El Niño/La Nina on the Indian monsoon: a new perspective. Weather 52:39–46

Kripalani RH, Kulkarni A (1997b) Rainfall variability over south East Asia—connections with Indian monsoon and ENSO extremes: new perspectives. Int J Climatol 17:1155–1168

Krishna Kumar KB, Rajagopalan B, Cane MA (1999) On the weakening relationship between the Indian monsoon and ENSO. Science 284:2156–2159

Krishnamurthy V, Kinter JL (2003) The Indian monsoon and its relation to global climate variability. In: Rodó X, Comín FA (eds) Global climate. Springer, Berlin, Heidelberg, pp 186–236

Krishnaswamy J, Vaidyanathan S, Rajagopalan B, Bonell M, Sankaran M, Bhalla RS, Badiger S (2015) Non-stationary and non-linear influence of ENSO and Indian Ocean dipole on the variability of Indian monsoon rainfall and extreme rain events. Clim Dyn 45(1–2):175–184

Kumar KK, Rajagopalan B, Hoerling M, Bates G, Cane M (2006) Un raveling the mystery of Indian monsoon failure during ElNiño. Science 314:115–119

Mini VK, Pushpa VL, Manoj KB (2016) Inter-annual and long term variability of rainfall in Kerala. Vayu Mandal 42(1):30–42

Pai ML, Pramod KV, Balchand AN (2014) Long range forecast on south west monsoon rainfall using artificial neural networks based on clustering approach. Int J Inf Technol Comput Sci 6(7):1–8

Parthasarathy B, Munot AA, Kothawale DR (1994) All-India monthly and seasonal rainfall series: 1871–1993. Theor Appl Climatol 49(4):217–224

Rao AR, Hsu EC (2008) Hilbert-Huang transform analysis of hydrological and environmental time series, vol 60. Springer Science & Business Media, Berlin

Sahai AK, Soman MK, Satyan V (2000) All India summer monsoon rainfall prediction using an artificial neural network. Clim Dyn 16:291–302

Saji NH, Goswami BN, Vinayachandran Yamagata T (1999) A dipole mode in the tropical Indian Ocean. Nature 401:360–363

Singh P, Borah B (2013) Indian summer monsoon rainfall prediction using artificial neural network. Stoch Environ Res Risk Assess 27(7):1585–1599

Varikoden H, Preethi B (2013) Wet and dry years of Indian summer monsoon and its relation with Indo-Pacific sea surface temperatures. Int J Climatol 33:1761–1771

Walker GT (1923) Correlation in seasonal variations of weather VIII, a preliminary study of world-weather. Mem Indian Meteorol Dep 24(4):75–131

Wang B, Xiang B, Li J, Webster PJ, Rajeevan MN, Liu J, Kuung-Ja Ha (2015) Rethinking Indian monsoon rainfall prediction in the context of recent global warming. Nat Commun. https://doi.org/10.1038/ncomms8154 **(Article number: 7154)**

Wu Z, Huang NE (2005) Ensemble empirical mode decomposition: a noise-assisted data analysis method. Centre for ocean-land-atmospheric studies Tech. Rep. 193, Center for Ocean-Land-Atmosphere Studies, Calverton, MD, pp 1–51. ftp://grads.iges.org/pub/ctr/ctr_193.pdf