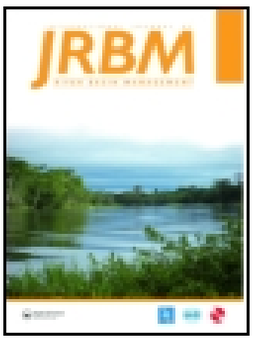

International Journal of River Basin Management

ISSN: 1571-5124 (Print) 1814-2060 (Online) Journal homepage: https://www.tandfonline.com/loi/trbm20

Multiscale modelling of monthly streamflows using MEMD-GP coupled approach

S. Adarsh & M. Janga Reddy

To cite this article: S. Adarsh & M. Janga Reddy (2019): Multiscale modelling of monthly streamflows using MEMD-GP coupled approach, International Journal of River Basin Management, DOI: 10.1080/15715124.2019.1656220

To link to this article: https://doi.org/10.1080/15715124.2019.1656220

Accepted author version posted online: 16 Aug 2019.
Published online: 27 Aug 2019.

INTL. J. RIVER BASIN MANAGEMENT
https://doi.org/10.1080/15715124.2019.1656220

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLES

Multiscale modelling of monthly streamflows using MEMD-GP coupled approach

S. Adarsh[a,b] and M. Janga Reddy [a]

[a]Department of Civil Engineering, Indian Institute of Technology Bombay, Mumbai, India; [b]Department of Civil Engineering, TKM College of Engineering, Kollam, India

ABSTRACT

This study presented multiscale characterization of monthly streamflow time series using Multivariate Empirical Mode Decomposition (MEMD) and developed an innovative approach for streamflow prediction by coupling MEMD with Genetic Programming (GP). Firstly, the possible hydro-climatic teleconnection of monthly streamflows of Mahanadi river basin in India with two large-scale climate oscillations of ElNiño Southern Oscillation (ENSO) and Equatorial Indian Ocean Oscillation (EQUINOO) is investigated by applying MEMD based Time-Dependent Intrinsic Correlation (TDIC) analysis. The TDIC analysis showed that the association between large-scale climate oscillations and streamflows is not unique always, but both the nature and strength of the association varies with time scales and over the time domain. Based on this finding, the study proposed MEMD-GP coupled approach for streamflow prediction, in which different modes corresponding to different process scales obtained by the MEMD are predicted separately using GP; and summation of these predicted modes provides the monthly streamflow at the station. A statistical performance evaluation based on multiple criteria showed that the proposed approach performs better than the multiple linear regression, M5 model tree and GP models for monthly streamflow prediction including extreme low and high flows, due to its unique capability to include the significant predictors at different time scales.

ARTICLE HISTORY
Received 22 June 2018
Accepted 4 August 2019

KEYWORDS
Streamflow; correlation; multiscale; MEMD; TDIC

Introduction

Characterization and feature extraction of hydrological time series is of great concern among the hydrologists as it may help for better understanding the process involved and their teleconnections with large scale climate oscillations (or variables), which may eventually help for improved simulation and prediction of hydro-climatic variables. Hydro-climatic teleconnection refers to the association of hydrologic variables with large-scale atmospheric/oceanic oscillations from different parts of the world (Maity *et al.* 2007). The teleconnection analysis can help in improved understanding of the liable mechanisms and improved prediction of hydrologic variables through the inclusion of the most appropriate predictor variables. In the Indian context, the history of such studies starts from the times of Walker (1923). The large-scale anomalous warming of sea surface temperature (SST) over the central and eastern Pacific Ocean with associated change in pressure field referred as ElNiño Southern Oscillation (ENSO) is inextricably linked with Indian Summer Monsoon Rainfall (ISMR) (Rasmusson and Carpenter 1983). Different researchers gave different physical explanations on the teleconnection mechanism between ISMR and ENSO events (Rasmusson and Carpenter 1983, Kane 1998, Krishna Kumar *et al.* 1999, Li *et al.* 2001). Majority of the studies recognized a negative association between the ISMR and ENSO while unanticipated experiences like above-average rainfall of 1997 and anomalous low rainfall of 2002 lead to the investigations on the concurrent effect of ENSO and other climatic oscillations on ISMR. After the discovery of the role of Indian Ocean Dipole Mode (IOD) on ISMR (Saji *et al.* 1999), the mutual influence of Equatorial Indian Ocean Oscillation (EQUINOO) along with ENSO (Gadgil *et al.* 2004, Maity and Nagesh Kumar 2006a, 2006b) on ISMR was unraveled. The association of ISMR with EQUINOO is due to the association of the monsoon rainfall over the Indian region with the northward propagation of convective system generated over the Indian Ocean region (Gadgil *et al.* 2004). Later on, it was found that the hydro-climatic teleconnections could be extended to large spatial scales such as river basins or subdivisions and also for investigating the climatic linkages of streamflow (Maity and Nagesh Kumar 2008, 2009). However, such studies were performed by using statistical correlation analysis between the respective time series and owing to the non-stationary, non-linear and multiscale behaviour of the candidate series, it is more logical to investigate such linkages in different process scales. The advanced spectral analysis techniques such as wavelet transform can decompose the data into multiple process scales and allows the modeller to investigate such associations in different process scales.

Wavelet transforms perhaps is the most popular multiscale spectral analysis tool, in which the original time series is usually decomposed into a set of sub-signals by continuous or discrete wavelet method, called data pre-processing. Each sub-signal plays a different role in the original series, and the behaviour of each sub-signal is distinct (Sang *et al.* 2016). The discrete wavelets have been widely used as a data pre-processing tool, which de-noises the signal and can be hybridized with artificial intelligence methods to use as a robust tool for improved predictions of hydrological variables (Nourani *et al.* 2014). It can also be used for characteristic analysis like trend analysis of hydrologic time series (Partal and Küçük 2006, Nalley *et al.* 2012, Adarsh and Janga Reddy 2015), and for characterization and feature

CONTACT S. Adarsh adarsh_lce@yahoo.co.in

extraction of hydrological series (Smith *et al.* 1998). Few researchers successfully used the continuous variant of wavelet transform (so-called Continuous Wavelet Transform, CWT) method for investigating the teleconnections between hydrological variable and climatic variables (Coulibali and Burn 2004, Anctil and Coulibaly 2004, Massei *et al.* 2007, 2011). It was also used by few studies for investigating the association between hydrological variables such as streamflow and suspended sediment, by which the effect of human interventions on streamflow variability can be captured successfully (Zhang *et al.* 2008, Rossi *et al.* 2009). The advancement of wavelets such as wavelet coherence and wavelet cross-spectrum helped to give better insights for the teleconnection studies (Grinsted *et al.* 2004, Zhang *et al.* 2007, Carey *et al.* 2013; Araghi *et al.* 2017). The selection of appropriate mother wavelet function and fixing the appropriate number of decomposition levels are two main challenges in the application of wavelet analysis. Also, the continuous nature of periodic scales induces vagueness in the separation of process scales. Even though the discrete wavelet can solve the second issue, it works only in dyadic periodic scales and the selection of an appropriate decomposition level is a complicated process (Sang *et al.* 2016). To overcome this issue to a certain extent, Huang *et al.* (1998) proposed a data-adaptive decomposition method called Empirical Mode Decomposition (EMD), which does not demand any *apriori* selection of a basis function. Later, the EMD extended as a spectral analysis technique by introducing a second phase, in which the obtained orthogonal modes are subjected to Hilbert transform to result in time–frequency spectra. The spectral analysis technique so-called Hilbert Huang Transform (HHT) has been catching popularity in the past decade (Huang *et al.* 2009, Kuai and Tsai 2012, Massei and Fournier 2012, Janga Reddy and Adarsh 2016). The HHT provides much flexibility in performing correlation analysis on multiple time scales. Chen *et al.* (2010) proposed a running correlation analysis method namely Time-Dependent Intrinsic Correlation (TDIC), in which the challenging problem of fixing the size of scaling window is solved by coupling the HHT method with the standard running correlation analysis procedure. TDIC analysis is gaining popularity in multiscale correlation analysis of hydro-meteorological and environmental time series (Huang and Schmitt 2014, Ismail *et al.* 2015, Adarsh and Janga Reddy 2016a, 2016b, 2018).

As most of the hydrologic processes are complex, the soft computing methods and their hybrid variants are widely used for hydrological forecasting (Nourani *et al.* 2014). In hybrid variants, first decomposition methods like Wavelet Transform or EMD are used for the decomposition of datasets (Karthikeyan and Nagesh Kumar 2013, Zhu *et al.* 2016). Then separate regression models are developed to estimate the response variables for different process scales using soft computing methods, and the estimations were finally recombined to get the overall prediction of the target variable. However, the decomposition of different time series of concern into multiple time scales using the traditional EMD may result in 'mode-misalignment' and lead to misinterpretations in teleconnection studies; and pose difficulties in developing hybrid predictive models. These issues restricted the researchers from considering the influence of multiple causal variables, and therefore, most of the past studies followed the decomposition of the target variable alone and not the predictor datasets. The multivariate extension of EMD (MEMD) proposed by Rehman and Mandic (2010) is an intuitive method that captures the common scales inherent in different time series of a multivariate dataset during decomposition and results in the same number of modes for different variables, and hence it can overcome the problem of mode misalignment. Therefore, this study presents an effective approach involving MEMD based hybrid model for the prediction of monthly streamflows. The specific objectives of the paper are: (i) to propose MEMD-TDIC approach for investigating the teleconnections of monthly streamflow in Mahanadi river basin, with two prominent climatic oscillations ENSO and EQUINOO; (ii) to demonstrate the usefulness of MEMD-Genetic Programming coupled framework for prediction of monthly streamflows of Mahanadi River.

Materials and methods

Investigating multiscale teleconnections of basin scale streamflows

In the past, to investigate the hydro-climatic teleconnections, few studies performed multiscale decomposition of the time series pair and compared the periodicity of their oscillatory modes, and other studies performed simple statistical correlation analysis (Iyengar and Raghu Kanth 2005, Maity and Nagesh Kumar 2008). This study presents a practical approach that performs a running correlation of oscillatory modes of the time series pair using TDIC analysis. For a multiscale investigation of hydro-climatic teleconnections, first correlation analysis performed between the oscillatory components of hydrologic variables and different climatic indices. Subsequently, the comparison of their residues is made in the time domain to examine the link between them in a non-stationarity perspective.

The different steps involved in the procedure are:

1. Decompose the time series of a hydrological variable (streamflow) and climate indices using the EMD method or its variants. If multiple variables are involved in the teleconnection study, MEMD is more suitable for the decomposition of the time-series signals. With MEMD simultaneous decomposition of multiple signals of concern results in a set of rotational modes of specific periodic properties and distinct characteristics.
2. Perform a correlation analysis by finding the correlation coefficient between the IMFs of hydrologic time series with that of a particular climatic index time-series to draw proper inferences regarding the association, in terms of periodicity.
3. Rescale the residue of hydrologic time-series about its mean.
4. Compare the zero-mean residue of climatic index time series with the zero-mean residue of hydrological time series.
5. If the zero crossings of both series are nearly at the same time instant, vital conclusions can be drawn regarding the association of the two signals in terms of non-stationarity.
6. Perform TDIC analysis between the respective IMFs (of comparable periodicity) to draw useful inferences on the association between the two in different time scales.

In this study, MEMD is used for the decomposition of multivariate datasets, and a flowchart of the proposed

methodology is given in Figure 1. Theoretical background and other details of the MEMD and TDIC algorithms are given in Appendix 1 and 2, respectively.

MEMD-GP coupled approach for hydrologic prediction

In the present study, MEMD is invoked for the simultaneous decomposition of multivariate datasets required for developing the regression model. The proposed methodology involves the following steps:

1. Decompose the multivariate data set comprising discharge at current time step and the lagged inputs using MEMD to get different orthogonal modes (subseries) called intrinsic mode functions (IMFs), each with a specific time scale of variability
2. Build genetic programming models to predict each mode as a function of the corresponding mode of different input variables
3. Predict the modes of discharge (at the current time step) at different time scales by the refined models
4. Combine the predicted modes to get the predicted discharge at observation scale

Figure 2 shows the flowchart of the proposed MEMD-GP methodology.

Genetic Programming (GP) proposed by Koza (1992) uses a set of computer programmes to approximately solve different type of optimization or regression problems. The basic principle of GP lies in optimization, which can also be effectively applied for Genetic Symbolic Regression (GSR). GSR involves finding a mathematical expression in the symbolic form relating finite values of a set of independent variables (x_i) and a set of dependent variables (y_i). Like other evolutionary algorithms, the solution starts with a random set of population. The solution set in each iteration is collectively called as a generation. The subsequent generations created through genetic operations such as 'mutation' and 'crossover'. Each possible solution set can be visualized as a 'parse tree'. A parse tree is built up from a terminal set (the independent variables and constants in the problem) and a function set (the basic operators used to form the function are +, −, /, *, log, etc.). As the new solution evolves through basic genetic operations, these solution set replaces the previous set and are supposed to perform better, which is assessed by the 'fitness' (a measure of how closely the population solves the problem) based on the principle of 'survival of the fittest'.

The Model tree is a machine-learning technique proposed by Quinlan *et al.* (1992) which combines the features of classification and regression. The algorithm splits the parameter space into sub-areas and builds a linear regression model in each of them. So, they are analogous to piecewise linear functions. The splitting in the MT follows the idea of a decision tree, but instead of class labels, it has linear regression functions at the leaves. The splitting criterion is used to determine which attribute is the best to split the portion T (set of examples that reach the node) of the training data that reaches a particular node. It is based on treating the standard deviation of the class values in T as measure of the error at that node, and calculating the expected error

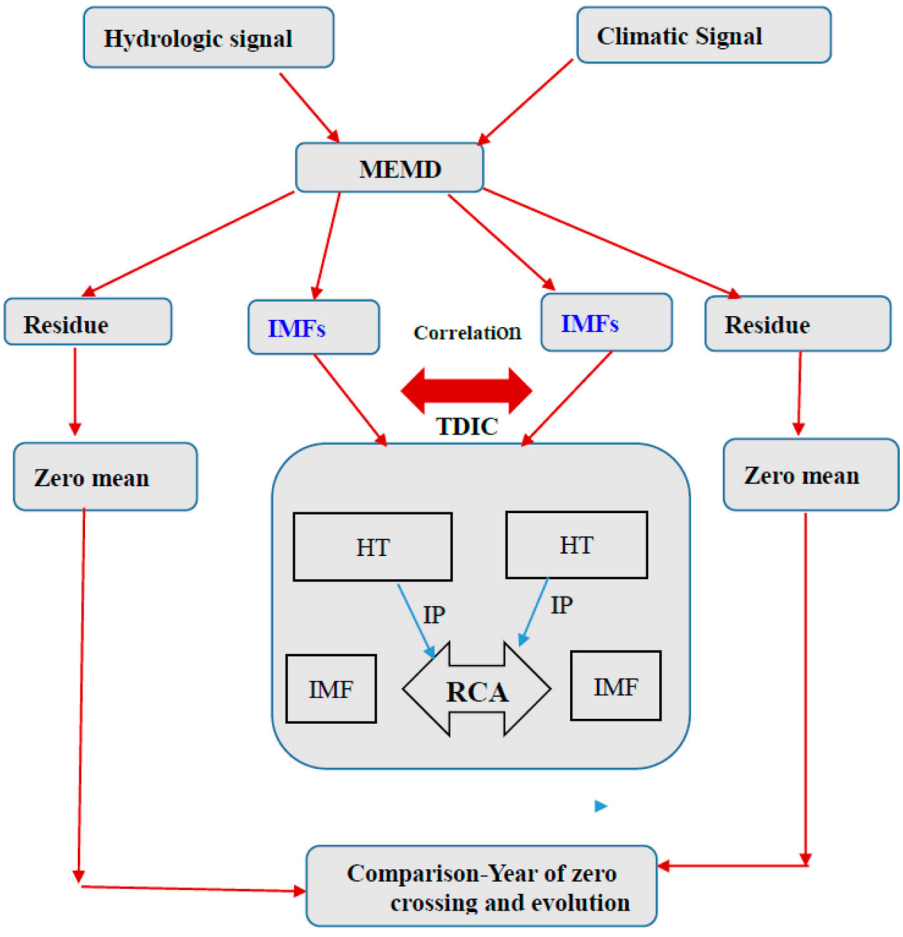

Figure 1. Flowchart of the methodology for hydro-climatic teleconnection analysis using TDIC approach. RCA refers to running correlation analysis; IP refers to instantaneous period.

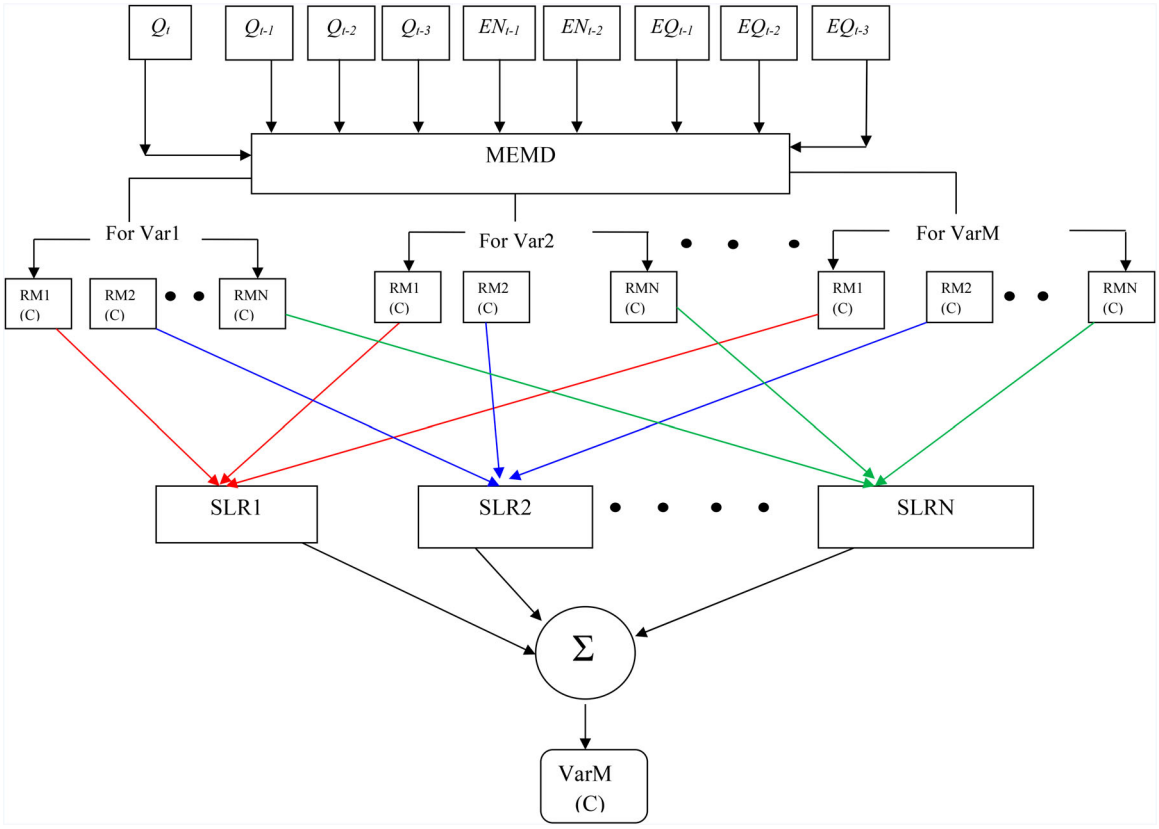

Figure 2. Flowchart of the MEMD-GP coupled methodology for streamflow prediction. Here OM refers to orthogonal mode; Var1, Var2 … VarM refers to M variables.

as a result of testing each attribute at that node. The formula to compute the standard deviation reduction (SDR) is

$$\mathrm{SDR} = sd(T) = \sum_i \frac{|T_i|}{|T|} sd(T_i)$$

Where T = set of examples that reach the node; T_i = subset of examples that have the i^{th} outcome of the potential set and sd = standard deviation. The variant of a Model tree which uses SDR as the splitting criteria is called as M5 Model tree. When a Model tree is used to predict the value for a test instance, the tree is followed down to a leaf in the usual way, using the instances attribute values to make the routing decisions at each node. The leaf contains a linear model based on some of the attribute values and is evaluated for the test instance to yield a raw predicted value. After the initial tree has grown, several steps have to be taken, such as calculation of error estimates, generation of linear models, simplification of linear models, and pruning and smoothing. Once the linear model is in place for each interior node, the tree is pruned back from the leaves so long as the expected estimated error decreases. Smoothing is a process used to compensate for the sharp discontinuities that may inevitably occur between adjacent linear models at the leaves of the pruned trees. More details on the smoothing and pruning operations of the algorithm can be found elsewhere (Witten *et al.* 2005).

Study area and dataset

Mahanadi river is one of the important intermittent rivers in India, which has a length of 850 km approximately rises in the highlands of Chattisgarh and flows through Chattisgarh and Orissa states in east-central India drains an area of 141,589 km^2 to discharge into the Bay of Bengal along the east coast of India. The major part of the upper reaches of the Mahanadi river catchment encompasses most of the areas of Chattisgarh state. Basantpur river gauging station in Mahanadi river operated by the Central Water Commission (CWC) of India is located a few kilometres upstream of the Hirakud dam; the most important and largest among different control structures of the basin and the upper Mahanadi cover 60% of the basin nearly 83,000 km^2 of the basin area. Tikrapara station is located in the central Mahanadi basin (in Orissa state between 82°E–86°E longitude and 19° N–22°N latitude having area of nearly 43,000 km^2). Figure 3 shows the location map of the study area, gives the details of the location of the Basantpur and Tikrapara stations and Hirakud dam.

The data of monthly streamflows of Mahanadi river at two sites of Basantpur and Tikrapara stations for the period 1973–2007 was collected from the web portal of water resources information system of India (http://www.india-wris.nrsc.gov.in/wris.html) and used in the study. The SST data corresponding to the Niño3.4 region (120°W–170°W, 5°S–5°N) called as Oceanic Niño Index (ONI), was obtained from NOAA National Weather Service Climate Prediction Center (http://www.cpc.ncep.noaa.gov/data/indices/) for the period 1973–2007 and used as the ElNiño Southern Oscillation (ENSO) index. The negative of the anomaly of the zonal component of surface wind in the equatorial Indian Ocean region (60–90°E, 2.5°S–2.5°N) collected from National Centre for Environmental Prediction (NCEP) (http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.html) for the

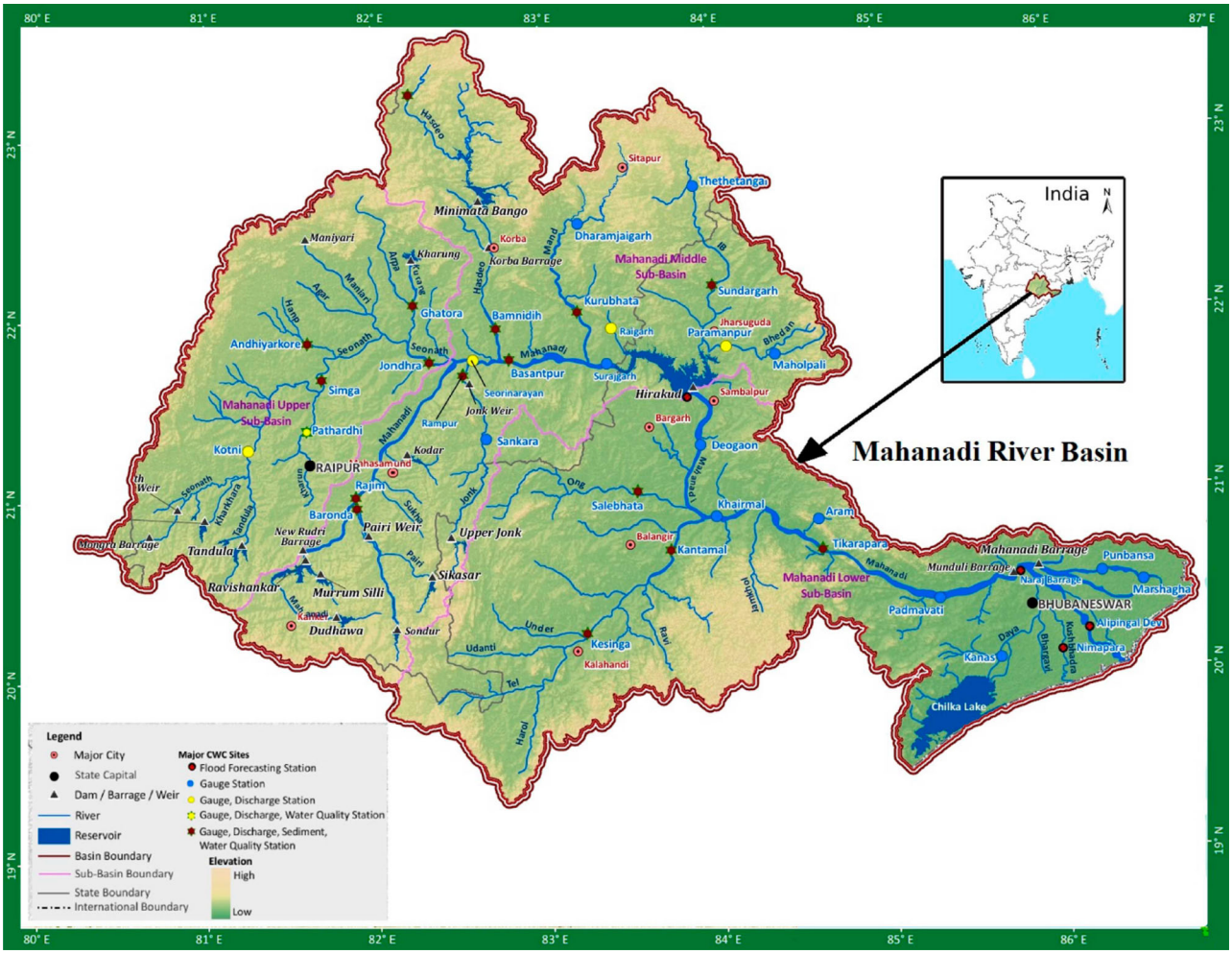

Figure 3. Map of the study area, Mahanadi basin and location of hydrologic stations (Reprinted Based on http://india-wris.nrsc.gov.in/wrpinfo/index.php?title=Mahanadi).

same period and referred as Equatorial Indian Ocean Oscillation (EQUINOO) index.

Results and discussion

First, the multivariate dataset comprising monthly streamflows from Basantpur (or Tikrapara) station of Mahanadi river basin for the period 1973–2007 and two potential large scale climate oscillations ENSO and EQUINOO are decomposed into multiple time scales using MEMD method. In the implementation of MEMD, a total of 64 direction vectors are chosen, maximum threshold, minimum threshold and the fraction of decomposition are set as 0.75, 0.075 and 0.5 respectively following the guidelines given in the past studies (Rilling *et al.* 2003, Hu and Si 2013). The decomposition resulted in 6 IMFs and one

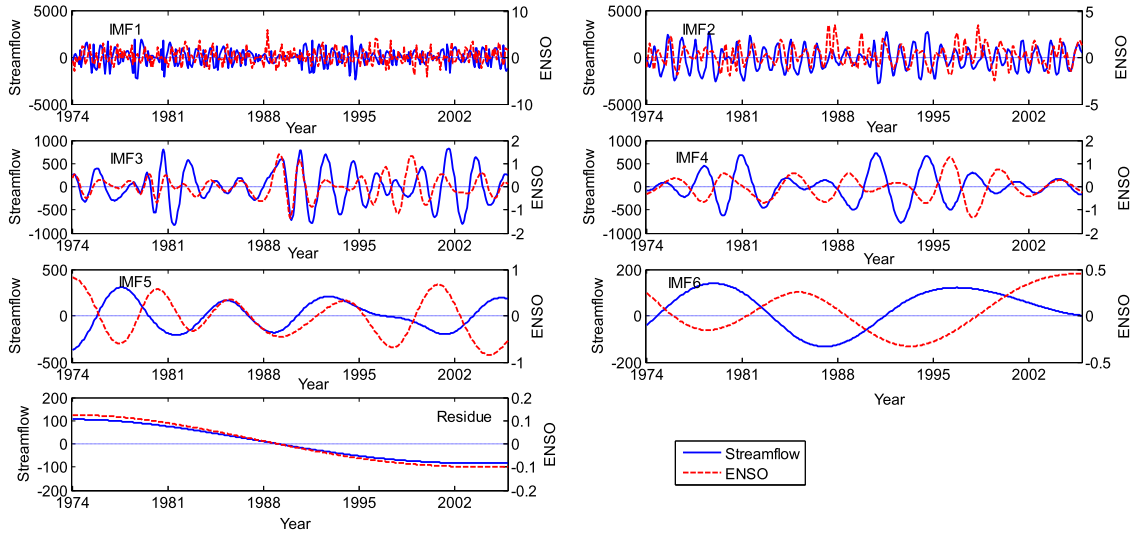

Figure 4. Comparison between the rotational modes of streamflow of the Basantpur station with that of ENSO.

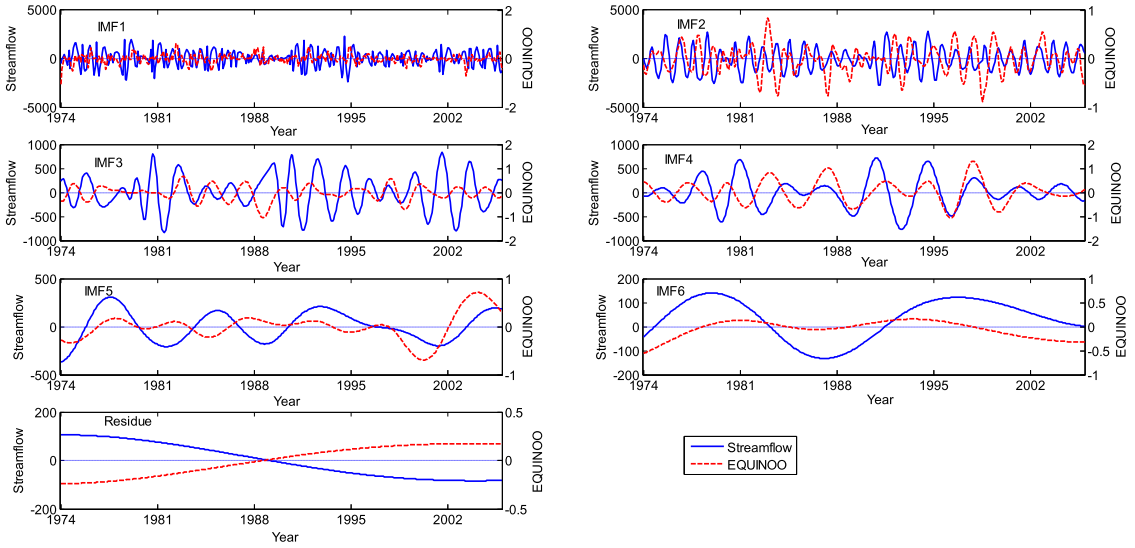

Figure 5. Comparison between the rotational modes of streamflow of the Basantpur station with that of EQUINOO.

Table 1. Mean period of the rotational modes obtained by multivariate datasets comprising streamflow, ENSO and EQUINOO series.

	Mean period (Months)			
	Station			
Mode Number	Basantpur	Tikrapara	ENSO	EQUINOO
IMF1	5.49	5.27	3.66	3.73
IMF2	11.29	10.69	10.67	8.93
IMF3	21.33	24.33	24.00	20.21
IMF4	38.40	42.18	48.00	48.00
IMF5	96.00	92.6	64.00	96.00
IMF6	192.00	194.0	192.00	384.00
Residue	LT	LT	LT	LT

residue for all the three series. The rotational modes obtained (using MEMD based decomposition) for streamflow of Basantpur station along with that of ENSO and EQUINOO are presented in Figures 4 and 5, respectively. Table 1 presents the details of periodicities obtained for different modes.

Similarly, the rotational modes plots for streamflow of Tikarapara station along with that of ENSO and EQUINOO are presented in Figures 6 and 7, respectively. Table 2 gives the details of correlations between different modes (considering the complete data length).

Figures 4–7 show the comparison of modes at different process scales. It is noticed that at some of the time spells the association of modes is alike, while at some other spells, there may be even opposing behaviour. Also, for low-frequency modes, only fewer cycles are present, and the relation between the climate signal and hydrologic signal are not alike. In the residue components, there exists a strong positive association in the ENSO-streamflow link and strong negative association in the EQUINOO-streamflow link for both the gauging stations. The cross- correlation analysis also (Table 2) confirms this long-term association, whereas the magnitude of the correlation is minimal at other time scales. This may be due to the fact that, for a given pair of oscillatory modes, at some shorter time spells the association may be strongly positive, while at some other spells the relation may turn into strongly negative, and in effect, the correlation gets neutralized. Performing accurate dynamical modelling by incorporating the information on local processes and topography may help in identifying such differences. Also, such models are to be validated with selected local precipitation events occurred in the past, which displayed a distinct character. The TDIC analysis employed to quantify such linkages (correlations) statistically. The results of TDIC analysis

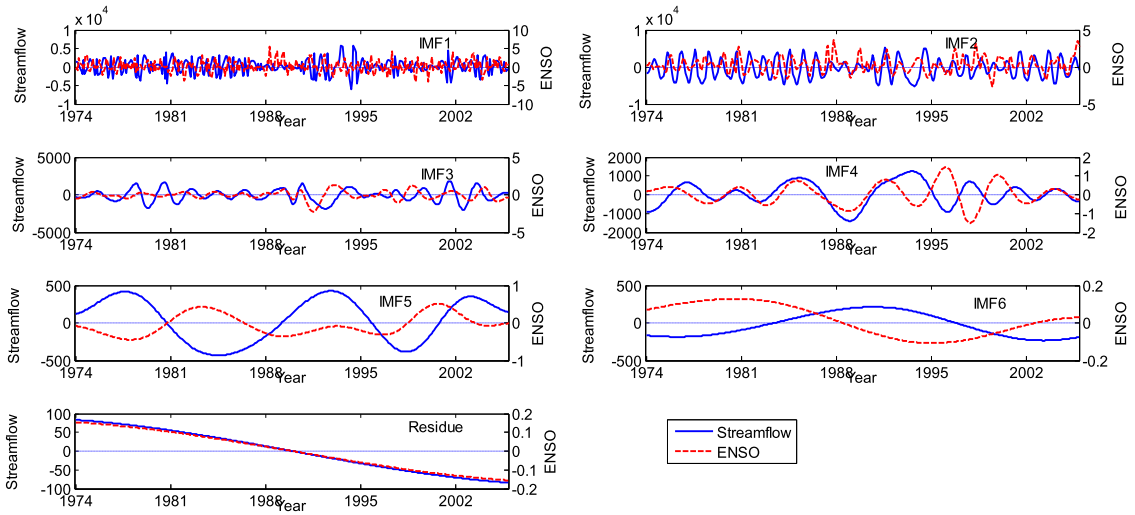

Figure 6. Comparison between the rotational modes of streamflow of the Tikrapara station with that of ENSO.

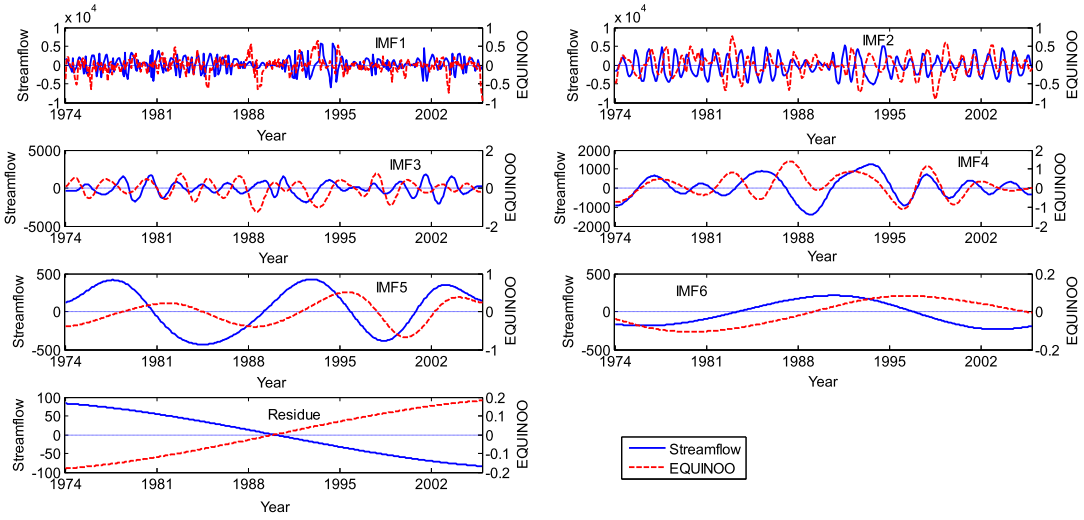

Figure 7. Comparison between the rotational modes of streamflow of the Tikrapara station with that of EQUINOO.

for ENSO-streamflow links and EQUINOO-streamflow links for Basantpur station are presented in Figures 8 and 9, respectively.

Similarly, the plots of TDIC analysis for ENSO-streamflow links and EQUINOO-streamflow links for Tikrapara station are presented in Figures 10 and 11, respectively. There is dominancy of negative associations in the first three IMFs of ENSO-streamflow links of Basantpur station (Figure 8), but there exist frequent reversals in the nature of the association between the two over the time domain. The bottom contour of the triangular plots depicts the instantaneous frequency and hence a shift of the plots to larger time scales can be noticed in higher-order IMFs (i.e. low-frequency modes).

Figure 9 shows that there is a dominance of anti-correlation in EQUINOO-streamflow relationships. From Figure 9, it is clear that there exists a long-range negative correlation at process scales of modes 1, 4, and 5. In IMF4, there is a localized positive association during ~1982–85, but the association is weak. At the process scale of IMF3 (like annual), there is a dominancy of positive association, but during ~1995–2000 the association is strongly negative. There is rich dynamics in the nature/strength of association between the two for IMF2. Similar behaviour is noticed in the plots of Tikarapara station (Figures 10 and 11).

The influence of large scales circulation patterns such as ENSO, Indian Ocean dipole (IOD) mode and EQUINOO

Table 2. Results of cross-correlation between the rotational modes of streamflow with that of ENSO and EQUINOO. The values in italics show that the correlation is significant at 5% level.

Mode of streamflow	Mode of climatic oscillation						
	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	Residue
				Streamflow-ENSO (Basantpur)			
IMF1	*−0.288*	*−0.161*	0.011	0.018	0.047	0.036	0.008
IMF2	0.035	*0.161*	−0.047	0.023	−0.011	0.013	−0.059
IMF3	−0.025	*−0.169*	*0.367*	0.052	0.003	0.003	−0.009
IMF4	−0.037	0.113	−0.006	*−0.421*	0.032	0.008	0.010
IMF5	0.075	0.000	−0.025	0.016	*−0.364*	*−0.278*	−0.017
IMF6	−0.003	−0.091	0.033	−0.004	0.035	*−0.362*	−0.089
Residue	−0.011	−0.052	0.009	−0.022	*0.252*	*−0.194*	*0.999*
				Streamflow-EQUINOO (Basantpur)			
IMF1	−0.041	0.091	0.028	−0.025	−0.023	−0.013	−0.008
IMF2	0.089	−0.110	−0.032	−0.001	0.018	−0.008	0.059
IMF3	0.007	−0.061	*−0.232*	−0.061	−0.027	−0.007	0.010
IMF4	−0.078	0.001	0.002	*0.211*	−0.028	0.001	−0.012
IMF5	0.049	0.037	0.150	0.017	*0.378*	*0.216*	0.019
IMF6	−0.006	0.098	0.179	−0.050	*−0.162*	*0.155*	0.067
Residue	−0.007	0.018	0.039	0.010	−0.109	0.044	*−0.999*
				Streamflow-ENSO (Tikrapara)			
IMF1	*−0.242*	*−0.167*	0.015	−0.016	0.000	0.027	0.018
IMF2	0.058	0.111	−0.036	0.039	0.027	−0.007	−0.031
IMF3	0.026	−0.090	0.105	−0.124	0.008	0.036	−0.005
IMF4	−0.083	−0.022	−0.035	0.065	*0.177*	−0.077	−0.042
IMF5	0.084	−0.067	−0.090	0.051	*−0.480*	−0.078	−0.017
IMF6	−0.003	−0.033	0.004	−0.021	*−0.236*	*−0.412*	0.061
Residue	0.002	−0.050	−0.005	−0.004	*−0.262*	*0.734*	*0.999*
				Streamflow-EQUINOO (Tikrapara)			
IMF1	−0.053	0.092	−0.004	−0.007	0.029	−0.026	−0.011
IMF2	0.051	−0.129	0.040	−0.026	−0.019	0.015	0.021
IMF3	*−0.159*	0.021	−0.102	−0.108	0.013	−0.029	0.006
IMF4	*0.140*	−0.034	0.031	*0.292*	0.242	0.057	0.045
IMF5	0.041	0.066	−0.017	0.004	*0.209*	0.069	0.019
IMF6	0.143	−0.079	*−0.211*	*0.349*	−0.008	*0.223*	−0.059
Residue	−0.003	0.035	0.083	0.053	*−0.161*	*−0.840*	*−0.999*

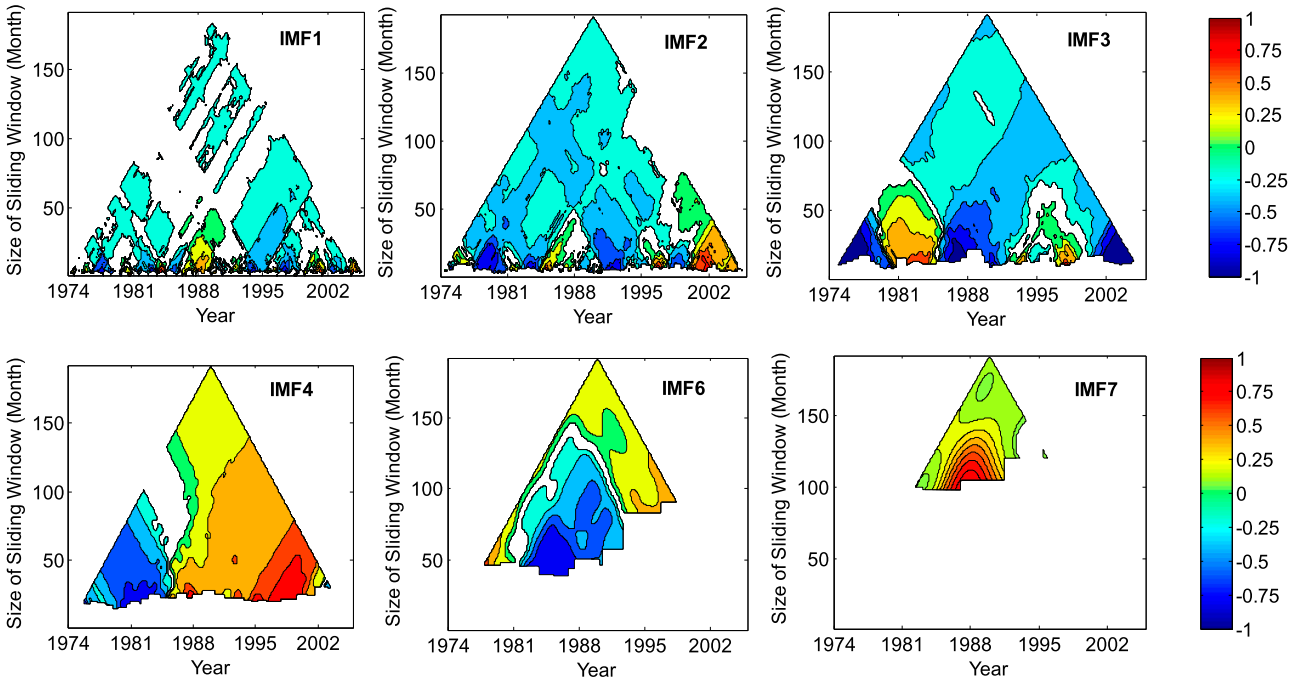

Figure 8. TDIC plot between the modes of streamflow of the Basantpur station and that of ENSO series. Only significant correlations (at significance level of 5%) are marked in the correlation plot.

on rainfall patterns in India were studied extensively in the past (Saji *et al.* 1999, Gadgil *et al.* 2004, Maity and Nagesh Kumar 2006a). The strength of the hydro-climatic teleconnections is more visible only at larger spatiotemporal scale, and few studies noted that such associations are prominent for larger river basins in India such as Mahahndi river basin (Maity *et al.* 2007, Maity and Nagesh Kumar 2008). The Basantpur station located a few kilometres upstream of the Hirakud reservoir (Figure 3) and its upstream catchment area lies mostly in the Chattisgarh state. Maity and Nagesh Kumar (2008) also noted the association of rainfall over Chattisgarh subdivision and streamflows of the basin with the prominent climatic predictors (such as ENSO and EQUINOO). Hence, the influence of these climatic oscillations on streamflow possibly exists on a monthly scale, as there exists a link between rainfall pattern and the climatic oscillations.

From the TDIC analysis, it is clear that the hydro-climatic associations of streamflows of Mahanadi river are not of unique character both with the process scales and over the time domain. Also, the strength of correlation varies with the process scales and along with the time domain.

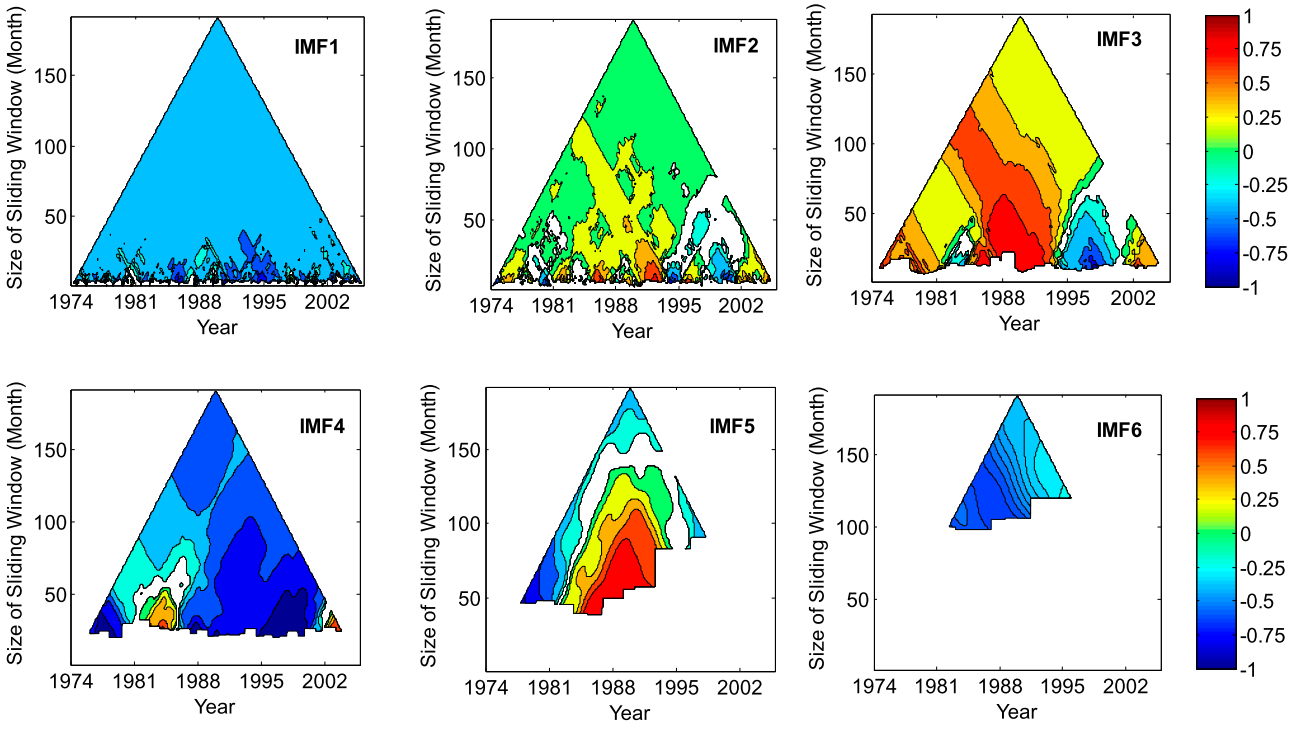

Figure 9. TDIC plot between the modes of streamflow of the Basantpur station and that of EQUINOO series. Only significant correlations (at significance level of 5%) are marked in the correlation plot.

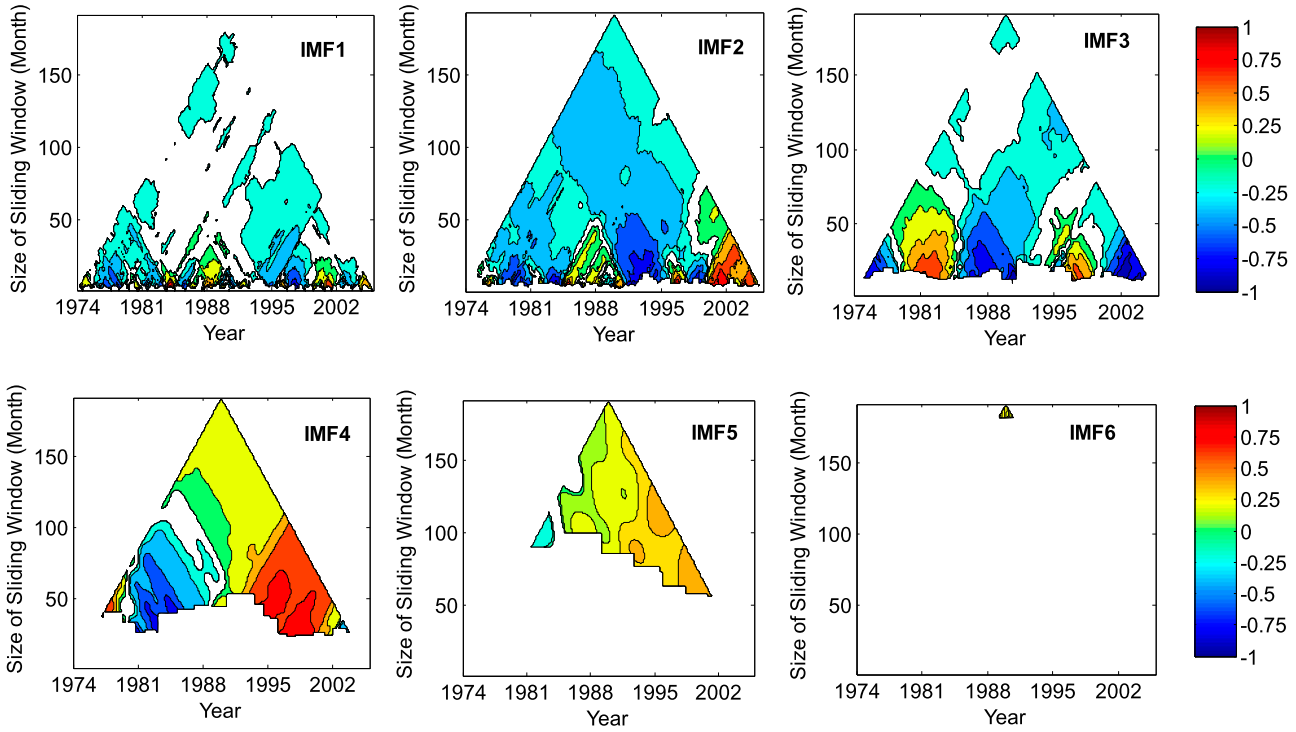

Figure 10. TDIC plot between the modes of streamflow of the Tikrapara station and that of ENSO series. Only significant correlations (at significance level of 5%) are marked in the correlation plot.

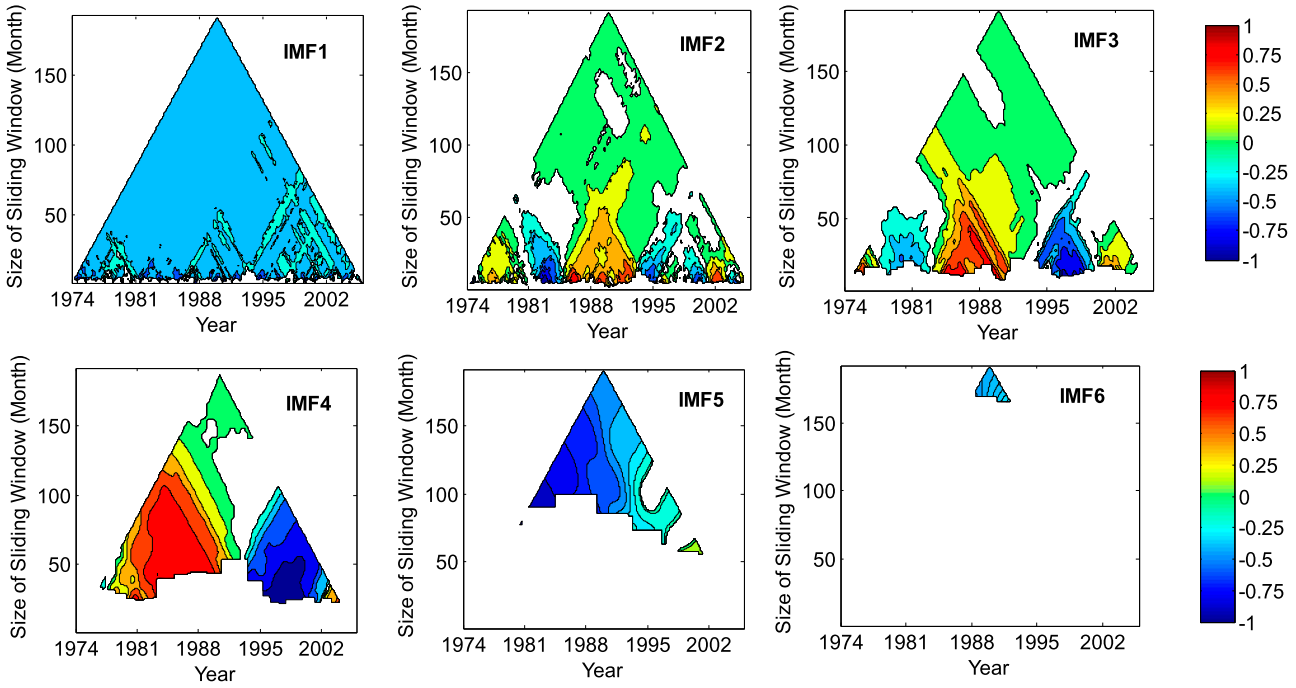

Figure 11. TDIC plot between the modes of streamflow of the Tikrapara station and that of EQUINOO series. Only significant correlations (at significance level of 5%) are marked in the correlation plot.

The reason for the negative association between streamflow and EQUINOO could be that the convective activity increases over the western part of the Indian Ocean. Being located in the eastern part of India, the rainfall might be lower-than-normal over the Mahanadi basin, which eventually causes in decreased streamflow at Basantpur station. Also, the negative associations of climatic oscillations with streamflows of Mahanadi river is more apparent in the data of Tikrapara station. The streamflow variability is controlled by both precipitation changes and human interventions like the construction of storage reservoirs and land-use changes (Zhang *et al.* 2008, 2009). Basantpur station is located at the upper reaches of Mahanadi basin, while the Tikrapara station is located at the downstream reach. Therefore, it is logical to believe that the direct influence of climatic variables on streamflow variability is more evident at Basantpur, whereas the human interventions also play a role in the streamflow variability of Tikarapara station. Because of this reason, in the subsequent section, the application of MEMD based hybrid model for monthly streamflow predictions is demonstrated with the data of Basantpur station.

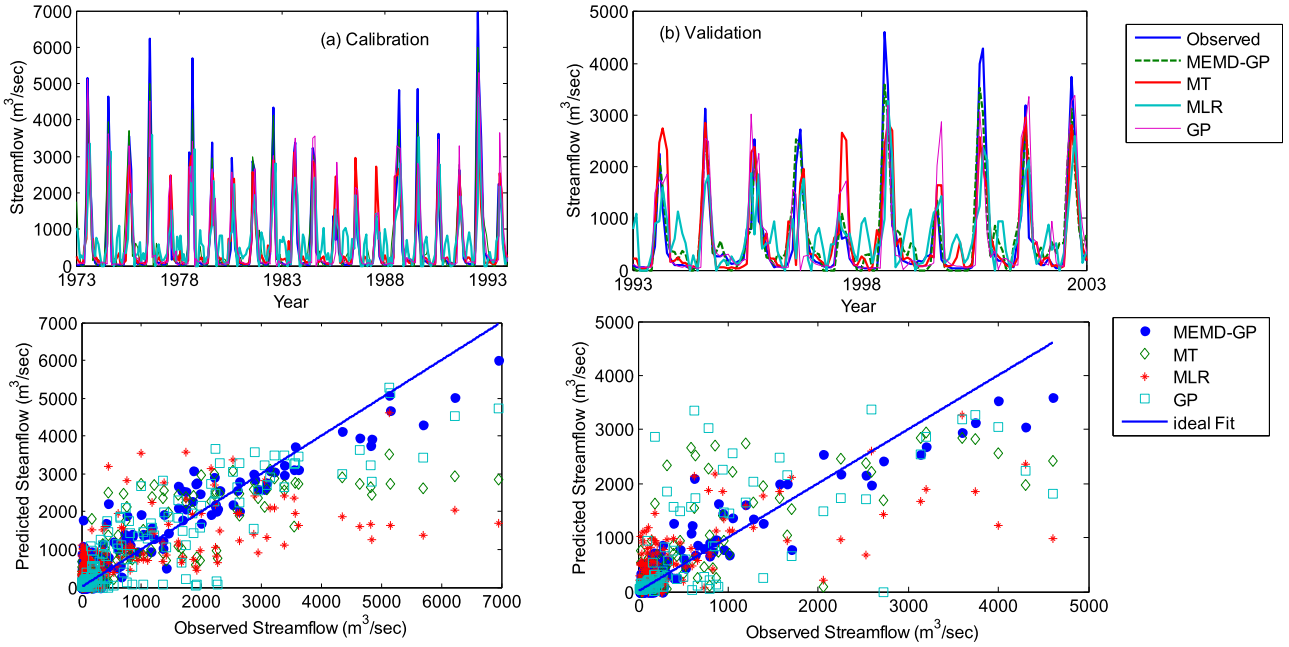

Figure 12. Comparison of streamflow predictions for the Basantpur station obtained using MEMD-GP, MT, MLR, and GP models with the observed data. Upper panels show time series plots and lower panels show scatter plots.

For the streamflow prediction of Basantpur station, first, the partial autocorrelation function (PACF) analysis is employed to find the number of lagged streamflow values to be considered as inputs. Also, the cross-correlation between the streamflow of generic month '*t*' with the climatic indices of past 12 months is performed to identify the climatic inputs. Thus three past lagged values of streamflow (Q_{t-1}, Q_{t-2}, Q_{t-3}), two lagged values of ENSO index (EN_{t-7}, EN_{t-8}) and two lagged values of EQUINOO (EQ_{t-3}, EQ_{t-4}). Thus the multivariate dataset comprises in total of 8 variables, during the calibration. The MEMD resulted in 8 modes and residue for all the variables. Now separate GP based regression models are developed for prediction of orthogonal modes of streamflow of time step '*t*'. Finally, the predicted modes are recombined to get the monthly streamflow of a generic month *t*.

To evaluate the efficacy of the proposed approach, the streamflow predictions are also made by GP, M5 Mode tree and Multiple Linear Regression (MLR) methods considering the above stated seven variables as input. In the application, first 70% of the data are used for training, and the remaining 30% of the data are used for validation. In the application of GP, a population size of 100, the initial mutation frequency of 0.9 and crossover frequency of 0.05, a function set comprising '*if-else*' rules along with basic arithmetic operators (*, +, −) are adopted. The time series and scatter plots of predictions are provided in Figure 12. The commonly used performance evaluation measures such as correlation coefficient (R), Nash-Sutcliffe Efficiency (NSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE) are used for comparing the performance of different models, and the results are summarized in Table 3. From Table 3, it can be noticed that least error statistics and highest correlation are found for the MEMD-GP hybrid model results; i.e. the hybrid model involving MEMD and GP performs remarkably well when compared with MLR, MT, and GP models for prediction of monthly streamflows.

Further, the performance of MEMD-GP methodology is examined for their efficacy in predicting extreme flows by considering three categories of flows, such as high flow, low flow, and medium flow. For this, first, the dataset is divided into three categories as follows. The data points that exceed mean plus standard deviation are considered as high flow, below mean are considered as low flows and in between are considered as medium flows. For such points, the predictions are shown in Figure 13, and the error and correlation statistics (in terms of RMSE and R) for validation data are summarized in Table 4. These results also, clearly show that MEMD-GP hybrid model is successful in capturing data points of different ranges and provides improved prediction of extremes flows. This superiority may be attributed to two facts that (i) additional variable(s) can be included in the modelling process with much ease with the use of MEMD model, instead of EMD. It also avoids mode misalignment and cleaner separation of process scales; (ii) the proposed method is capable of capturing the relevant information from different process scales. The GP models developed at different process scales may retain only the relevant information at a specific process scale. In short, the MEMD based TDIC analysis deciphers the association between streamflow and large scale climate oscillations at different process scales and providing better insights into such linkages; and

Table 3. Performance evaluation of different models for monthly streamflow predictions of Basantpur station. PEC-Performance Evaluation Criteria; C-Clibration; V-Validation.

	MEMD-GP		GP		MT		MLR	
PEC	C	V	C	V	C	V	C	V
R	0.950	0.937	0.845	0.768	0.652	0.647	0.896	0.715
NSE	0.897	0.867	0.705	0.571	0.424	0.412	0.802	0.451
RMSE	400.400	355.865	677.063	639.135	945.883	748.550	554.115	723.6
MAE	250.76	249.89	349.57	355.48	568.93	485.16	294.61	384.91

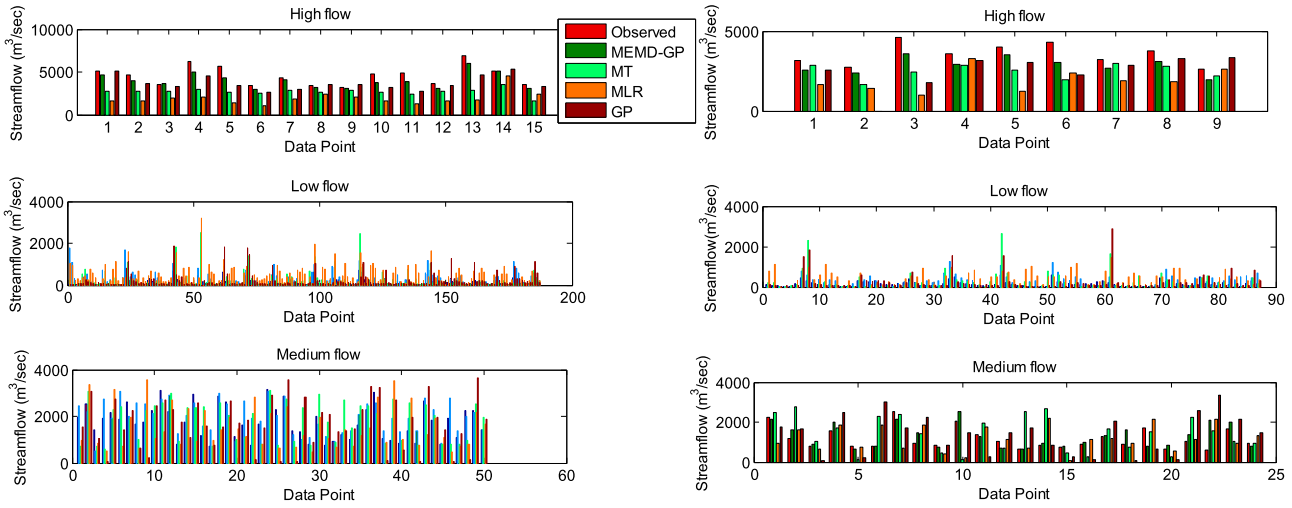

Figure 13. Comparison of streamflow predictions of the Basantpur station obtained using different models with the observed data for three flow states.

Table 4. Performance statistics of predictions by different methods under different flow states (validation data).

	R				RMSE			
Flow state	MEMD-GP	GP	MT	MLR	MEMD-GP	GP	MT	MLR
HF	0.922	0.108	−0.258	0.102	723.27	1289	1930.67	1554.95
LF	0.659	0.542	0.153	0.484	250.95	400.5	444.81	413.74
MF	0.686	0.252	−0.061	0.171	456.89	915.43	828.6	1044.38

MEMD-GP hybrid model is capable of capturing information from different process scales and results in better performance for streamflow prediction when compared to conventional data-driven regression methods.

Conclusions

This paper proposed MEMD-TDIC approach for investigating the hydro-climatic teleconnections and MEMD-GP hybrid framework for prediction of monthly streamflows in the Mahanadi river basin. From the results obtained for hydro-climatic teleconnection analysis and streamflow prediction in the Mahanadi river basin, the following conclusions are drawn from the study:

- The MEMD based TDIC analysis of monthly streamflows showed that the climatic teleconnections are dominantly positive at Basantpur station, while it is dominantly negative at Tikrapara station. This may be because the former station is located at upstream of the major control structure namely Hirkaud dam, while the latter one is located at far downstream of the basin, where the anthropogenic interventions are the significant drivers that influence the streamflow variability.
- The association between large-scale climate oscillations and streamflow of the basin is not unique always, but both the nature and strength of the association vary with time scales and over the time domain.
- Statistical performance evaluation based on multiple criteria showed that the proposed MEMD-GP hybrid approach displayed considerable improvement in the prediction of monthly streamflows of Basantpur station, including that of high flows when compared with multiple linear regression, M5 model tree, and GP methods. The proposed methodology using climate inputs is more suited for the prediction of natural flows in basins.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

M. Janga Reddy http://orcid.org/0000-0001-7003-8651

References

Adarsh, S., and Janga Reddy, M., 2015. Trend analysis of rainfall in four meteorological subdivisions in southern India using non-parametric methods and Discrete Wavelet Transforms. *International Journal of Climatology*, 35 (6), 1107–1124.

Adarsh, S., and Janga Reddy, M., 2016a. Analyzing the hydroclimatic teleconnections of summer monsoon rainfall in Kerala, India using Multivariate Empirical Mode Decomposition and Time Dependent Intrinsic Correlation. *IEEE Geoscience and Remote Sensing Letters*, 13 (9), 1221–1225.

Adarsh, S., and Janga Reddy, M., 2016b. Multiscale characterization of streamflow and suspended sediment concentration data using Hilbert Huang transform and time dependent intrinsic correlation analysis. *Modeling Earth Systems and Environment*, 2, 199. doi:10.1007/s40808-016-0254-z.

Adarsh, S., and Janga Reddy, M., 2018. Multiscale characterization and prediction of monsoon rainfall in India using Hilbert–huang transform and time-dependent intrinsic correlation analysis. *Meteorology and Atmospheric Physics*, 130 (6), 667–688.

Anctil, F., and Coulibaly, P., 2004. Wavelet analysis of the interannual variability in Southern Québec streamflow. *Journal of Climate*, 17, 163–173.

Araghi, A., *et al.*, 2017. Association between three prominent climatic teleconnections and precipitation in Iran using wavelet coherence. *International Journal of Climatology*, 37, 2809–2830.

Carey, S. K., *et al.*, 2013. Use of color maps and wavelet coherence to discern seasonal and interannual climate influences on streamflow variability in northern catchments. *Water Resources Research*, 49, 6194–6207.

Chen, X., Wu, Z., and Huang, N.E., 2010. The time-dependent intrinsic correlation based on the empirical mode decomposition. *Advances in Adaptive Data Analysis*, 2, 233–265.

Coulibali, P., and Burn, D.H., 2004. Wavelet analysis of variability in annual Canadian streamflows. *Water Resources Research*, 40. doi:10.1029/2003WR002667.

Gadgil, S., *et al.*, 2004. Extremes of the Indian summer monsoon rainfall, ENSO and equatorial Indian Ocean oscillation. *Geophysical Research Letters*, 31 (12), L12213. doi:10.1029/2004GL019733.

Grinsted, A., Moore, J.C., and Jevrejeva, S, 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics*, 11 (5), 561–566.

Hu, W., and Si, B-C., 2013. Soil water prediction based on its scale-specific control using multivariate empirical mode decomposition. *Geoderma*, 193-194, 180–188.

Huang, N.E., and Wu, Z., 2008. A review on Hilbert-Huang transform: method and its applications to geophysical studies. *Reviews of Geophysics*, 46 (2), 1–23. doi:10.1029/2007RG000228.

Huang, Y., and Schmitt, F.G., 2014. Time dependent intrinsic correlation analysis of temperature and dissolved oxygen time series using empirical mode decomposition. *Journal of Marine Systems*, 130, 90–100.

Huang, N.E., *et al.*, 1998. The empirical mode decomposition and the Hilbert spectrum for non-linear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454, 903–995.

Huang, Y., *et al.*, 2009. Analysis of daily river flow fluctuations using empirical mode decomposition and arbitrary order Hilbert spectral analysis. *Journal of Hydrology*, 373, 103–111.

Huang, G., *et al.*, 2016. Time-frequency analysis of non-stationary process based on multivariate empirical mode decomposition. *Journal of Engineering Mechanics*, 142 (1), 04015065. doi:10.1061/(ASCE)EM.1943-7889.0000975.

Ismail, D.K.B., Lazure, P., and Puillat, I., 2015. Advanced spectral analysis and cross correlation based on empirical mode decomposition: application to the environmental time series. *IEEE Geoscience and Remote Sensing Letters*, 12 (9), 1968–1972.

Iyengar, R.N., and Raghu Kanth, T.S.G., 2005. Intrinsic mode functions and a strategy for forecasting Indian monsoon rainfall. *Meteorology and Atmospheric Physics*, 90, 17–36.

Janga Reddy, M., and Adarsh, S., 2016. Time–frequency characterization of sub-divisional scale seasonal rainfall in India using the Hilbert–huang transform. *Stochastic Environmental Research and Risk Assessment*, 30 (4), 1063–1085.

Kane, R.P, 1998. Extremes of the ENSO phenomenon and Indian summer monsoon rainfall. *International Journal of Climatology*, 18, 775–791.

Karthikeyan, L., and Nagesh Kumar, D., 2013. Predictability of non-stationary time series using wavelet and EMD based ARMA models. *Journal of Hydrology*, 502, 103–119.

Koza, J.R, 1992. *Genetic programming: on the programming of computers by means of natural selection*. Cambridge, MA: The MIT Press.

Krishna Kumar, K., Rajagopalan, B., and Cane, M.A., 1999. On the weakening relationship between the Indian Monsoon and ENSO. *Science*, 284 (5423), 2156–2159. doi:10D1126/science.284D5423D2156.

Kuai, K.Z., and Tsai, C.W., 2012. Identification of varying time scales in sediment transport using the Hilbert–Huang Transform method. *Journal of Hydrology*, 420-421 (2012), 245–254.

Li, T., *et al.*, 2001. On the relationship between Indian Ocean sea surface temperature and Asian summer monsoon. *Geophysical Research Letters*, 28, 2843–2846.

Maity, R., and Nagesh Kumar, D., 2006a. 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*, 111, D07104. doi:10.1029/2005JD006539.

Maity, R., and Nagesh Kumar, D., 2006b. Hydroclimatic association of the monthly summer monsoon rainfall over India with large-scale atmospheric circulations from tropical Pacific Ocean and the Indian Ocean region. *Atmospheric Science Letters*, 7 (4), 101–107.

Maity, R., and Nagesh Kumar, D., 2008. Basin-scale streamflow forecasting using the information of large-scale atmospheric circulation phenomena. *Hydrological Processes*, 22 (5), 643–650.

Maity, R., and Nagesh Kumar, D., 2009. Hydroclimatic influence of large-scale circulation on the variability of reservoir inflow. *Hydrological Processes*, 23 (6), 934–942.

Maity, R., Nagesh Kumar, D., and Nanjundiah, R.S., 2007. Review of hydroclimatic teleconnection between hydrologic variables and large-scale atmospheric circulation patterns with Indian perspective. *ISH Journal of Hydraulic Engineering*, 13 (1), 77–92.

Massei, N., and Fournier, M., 2012. Assessing the expression of large-scale climatic fluctuations in the hydrological variability of daily Seine river flow (France) between 1950 and 2008 using Hilbert–huang transform. *Journal of Hydrology*, 448-449, 119–128.

Massei, N., *et al.*, 2007. Investigating possible links between the North Atlantic oscillation and rainfall variability in northwestern France over the past 35 years. *Journal of Geophysical Research*, 112, D09121. doi:10.1029/2005JD007000.

Massei, N., *et al.*, 2011. A wavelet approach to the short-term to pluridecennal variability of streamflow in the Mississippi River basin from 1934 to 1998. *International Journal of Climatology*, 31, 31–43.

Nalley, D., Adamowski, J., and Khalil, B, 2012. Using discrete wavelet transforms to analyze trends in streamflow and precipitation in Quebec and Ontario (1954–2008). *Journal of Hydrology*, 475, 204–228.

Nourani, V., *et al.*, 2014. Applications of hybrid wavelet–artificial intelligence models in hydrology: a review. *Journal of Hydrology*, 514 (6), 358–377.

Papadimitriou, S., Sun, J., and Yu, P.S., 2006. Local correlation tracking in time series. *Proceedings of IEEE Sixth International Conference on Data Mining*, 18–22 December, Hong Kong, 456–465.

Partal, T., and Küçük, M., 2006. Long-term trend analysis using discrete wavelet components of annual precipitations measurements in Marmara region (Turkey). *Physics and Chemistry of the Earth, Parts A/B/C*, 31, 1189–1200.

Quinlan, J.R., *et al.*, 1992. Learning with continuous classes. *In*: *Proceedings of Australian Joint Conference on Artificial Intelligence*. Singapore: World Scientific Press, 343–348.

Rasmusson, E.M, and Carpenter, T.H, 1983. The relationship between eastern equatorial Pacific sea surface temperature and rainfall over India and Sri Lanka. *Monthly Weather Review*, 111, 517–528.

Rehman, N., and Mandic, D.P., 2010. Multivariate empirical mode decomposition. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 466 (2117), 1291–1302.

Rilling, G., Flandrin, P., and Goncalves, P., 2003. On empirical mode decomposition and its algorithms. *In*: *Proceedings of IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP-03*, Grado (Italy), 8–11.

Rodo, X., and Rodriguez-Arias, M.A, 2006. A new method to detect transitory signatures and local time/space variability structures in the climate system: the scale-dependent correlation analysis. *Climate Dynamics*, 27, 441–458.

Rossi, A., *et al.*, 2009. The response of the Mississippi River to climate fluctuations and reservoir construction as indicated by wavelet analysis of streamflow and suspended-sediment load, 1950–1975. *Journal of Hydrology*, 377 (3), 237–244.

Saji, N.H, *et al.*, 1999. A dipole mode in the tropical Indian Ocean. *Nature*, 401, 360–363.

Sang, Y., *et al.*, 2016. Wavelet-based hydrological time series forecasting. *Journal of Hydrologic Engineering*. doi:10.1061/(ASCE)HE.1943-5584.0001347, 06016001.

Smith, L.C., Turcotte, D.L., and Isacks, B.L, 1998. Stream flow characterization and feature detection using a discrete wavelet transform. *Hydrological Processes*, 12, 233–249.

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.

Witten, I.H., Frank, E., and Hall, M.A., 2005. *Data mining: practical machine learning tools and techniques*. San Francisco, CA: Morgan Kaufmann.

Zhang, Q, *et al.*, 2007. Possible influence of ENSO on annual maximum streamflow of the Yangtze River, China. *Journal of Hydrology*, 333 (2-4), 265–274.

Zhang, Q., *et al.*, 2008. Periodicity of sediment load and runoff in the Yangtze river basin and possible impacts of climatic changes and human activities / périodicité de la charge sédimentaire et de l'écoulement dans le bassin du Fleuve Yangtze et impacts possibles des changements climatiques et des activités humaines. *Hydrological Sciences Journal*, 53 (2), 457–465.

Zhang, Q., *et al.*, 2009. Multiscale variability of sediment load and streamflow of the lower Yangtze river basin: possible causes and implications. *Journal of Hydrology*, 368, 96–104.

Zhu, S., *et al.*, 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.

Appendices

Appendix 1. Empirical mode decomposition and its multivariate extension

Empirical Mode Decomposition (EMD) is a data adaptive signal decomposition method proposed by Huang *et al.* (1998), which decomposes a time series signal into different oscillatory modes having specific periodicity.

The steps involved in EMD are:

(1) Identify all extrema (maxima and minima) of a given signal $X(t)$
(2) Connect the maxima points using interpolation function (like cubic spline) to form an upper envelope ($E_{max}(t)$); repeat the same for minima points to get a lower envelope ($E_{min}(t)$)
(3) Compute the mean of the upper and lower envelopes, $m(t)$
(4) Calculate the difference time series $d(t) = X(t)\text{-}m(t)$
(5) Let $d(t)$ be the new signal obtained and repeat steps (1) to (4) until $d(t)$ becomes a zero-mean series with no riding waves (i.e. there are no negative local maxima and positive local minima) with smoothened amplitudes. Such an oscillatory signal is called an Intrinsic Mode Function (IMF). To satisfy this step, an appropriate criterion is to be applied in the evolution of an IMF and the modified Cauchy type stopping criterion is applied in the study(Huang and Wu 2008).
(6) On satisfying the zero-mean condition, $d(t)$ can be designated as the first intrinsic mode function IMF1.
(7) Compute the residue $R_1(t)$ by subtracting IMF1 from original signal (i.e. $R1(t) = X(t)\text{-}IMF1(t)$) is used as new signal. The iterations (so called 'sifting') are to be repeated upon this signal to get IMF2.
(8) The higher oscillatory modes are obtained by treating the residue ($R_k(t)$) as the signal ($X(t)$), iteratively.

The k^{th} residue is defined as

$$R_k(t) = X(t) - \sum_{j=1}^{k} IMF_j(t)$$

The process will be continued till the resulting residue is a monotonic function or a function having only one extrema. The final component is called 'residue' which indicates the long term inherent trend within the time series.

Multivariate EMD method proposed by Rehman and Mandic (2010) 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. In this method, multiple envelops are produced by taking projections of multiple inputs along different directions in an m-dimensional space. Assuming $V(t) = \{v_1(t), v_2(t) \ldots . v_m(t)\}$ being the m vectors as a function of time t and $X^{\varphi_k} = \{x_1^k, x_2^k, \ldots ., x_m^k\}$ denoting the direction vector along different directions given by angles $\varphi_k = \{\varphi_1^k, \varphi_2^k, \ldots \varphi_{m-1}^k\}$ in a direction set X ($k = 1,2,3, \ldots .K$), 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. They are zero mean series with the number of extreme values in the overall data matches with the number of zero crossings or differ at the most by one.

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 $p^{\varphi_k}(t)$ of the datasets $V(t)$ along the direction vector X^{φ_k} for all k
3. Find temporal instants $t_i^{\varphi_k}$ corresponding to the maxima of projection for all k
4. Interpolate $[t_i^{\varphi_k}, V(t_i^{\varphi_k})]$ to obtain multivariate envelop curves $e^{\varphi_k(t)}$ for all k
5. The mean of envelope curves ($M(t)$) is calculated by $M(t) = \frac{1}{K}\sum_{k=1}^{K} e^{\varphi_k}(t)$
6. Extract the 'detail' $D(t)$ using $D(t) = V(t)\text{-}M(t)$. If $D(t)$ fulfils the stoppage criterion for a multivariate IMF, apply the above procedure from step (1) onwards upon the residue series (i.e. $V(t)\text{-}D(t)$). Otherwise, repeat the steps from (2) onwards upon the series $D(t)$.

The study adopted Hammersley sampling sequence for the generation of direction vectors and the Cauchy type stoppage criteria (Huang and Wu 2008, Huang *et al.* 2016) for implementation of MEMD.

Appendix 2. Time dependent intrinsic correlation analysis

Most of the hydrologic and meteorological time series possess multiscaling behaviour. Therefore it is more appropriate to use a scale dependent correlation analysis for hydroclimatic teleconnections studies (Papadimitriou *et al.* 2006, Rodo and Rodriguez-Arias 2006). The selection of most appropriate sliding window size is a major challenge to proceed with such analysis. Chen *et al.* (2010) proposed a method namely Time Dependent Intrinsic Correlation (TDIC) for determining scale dependent correlation between two candidate time series. The different steps involved in TDIC analysis are:

1. Use EMD (or its variants) to decompose the two time series of concern to get different Intrinsic Mode Functions (IMFs) of definite periodicity.
2. Find the instantaneous periods of both the IMFs (of a specific time scale) by HT
3. Find the minimum size of sliding window (t_d) as $t_d = \max(T_{1,i}(t_k), T_{2,i}(t_k))$ where $T_{1,i}$ and $T_{2,i}$ are instantaneous periods of the two candidate IMFs
4. Determine the sliding window size as $t_w^n = [t_k - 0.5nt_d{:}t_k + 0.5nt_d]$, where n is the multiplication factor for minimum sliding window size, a positive number.
5. Find the Pearson correlation between the segments of the two IMFs $R_i(t_k^n) = Corr(IMF_{1,i}(t_w^n), IMF_{2,i}(t_w^n))$ at any t_k, and perform student t-test to assess the statistically significance of these correlations.
6. Repeat steps 3–5 iteratively till the boundary of the sliding window crosses the end points of the time series

On completing the running correlation analysis, a correlation matrix is obtained, which is presented in the form of a triangular plot called as TDIC plot. In this plot, centre positions of the sliding windows will be in the x-axis and the size of the sliding window will be there in the y-axis. Appropriate color scheme is used to discern the correlations and the correlation at the apex point will be the correlation coefficient between the two time series, if we fix the maximum size of sliding window same as the data length (Chen *et al.* 2010).