



Drought prediction using a wavelet based approach to model the temporal consequences of different types of droughts



Rajib Maity*, Mayank Suman, Nitesh Kumar Verma

Department of Civil Engineering, Indian Institute of Technology Kharagpur, Kharagpur 721302, West Bengal, India

ARTICLE INFO

Article history:

Received 23 February 2016
 Received in revised form 27 April 2016
 Accepted 21 May 2016
 Available online 27 May 2016
 This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Alon Rimmer, Associate Editor

Keywords:

Drought prediction
 Standardized Precipitation Index (SPI)
 Standardized Soil Moisture Index (SSMI)
 Standardized Stream Flow Index (SSFI)
 Wavelet Transform (WT)

SUMMARY

Droughts are expected to propagate from one type to another – meteorological to agricultural to hydrological to socio-economic. However, they do not possess a universal, straightforward temporal dependence. Rather, assessment of one type of drought (successor) from another (predecessor) is a complex problem depending on the basin's physiographic and climatic characteristics, such as, spatial extent, topography, land use, land cover and climate regime. In this paper, a wavelet decomposition based approach is proposed to model the temporal dependence between different types of droughts. The idea behind is to separate the rapidly and slowly moving components of drought indices. It is shown that the temporal dependence of predecessor (say meteorological drought) on the successor (say hydrological drought) can be better captured at its constituting components level. Such components are obtained through wavelet decomposition retaining its temporal correspondence. Thus, in the proposed approach, predictand drought index is predicted using the decomposed components of predecessor drought. Several alternative models are investigated to arrive at the best possible model structure for predicting different types of drought. The proposed approach is found to be very useful for foreseeing the agricultural or hydrological droughts knowing the meteorological drought status, offering the scope for better management of drought consequences. The mathematical framework of the proposed approach is general in nature and can be applied to different basins. However, the limitation is the requirement of region/catchment specific calibration of some parameters before using the proposed model, which is not very difficult and uncommon though.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Drought is a hydrological extreme phenomenon of prolonged water deficit. However, unlike the other extreme hydrological phenomena, such as flood, it is slow initiating and long lasting phenomenon leading to huge economic losses. As per the [American Meteorological Society \(1997\)](#), droughts are of four types, namely meteorological, agricultural, hydrological, and socioeconomic. The deficit in precipitation, soil moisture and stream flow/reservoir storage leads to meteorological, agricultural and hydrological drought respectively.

Since the hydrological cycle is a continuous transport of water, the occurrence of meteorological drought is expected to propagate to other types of droughts ([Maybank et al., 1995](#)). For example, prolonged period of meteorological drought and high evaporation loss may evolve into soil moisture deficit, which causes agricultural drought. Again, hydrological droughts may evolve as a

consequence of prolonged periods of agricultural drought. However, the temporal transition of droughts may be speculated easily, but it is difficult to model, because the correlation between them is affected by various climatological, topographical and geographical characteristics. Sometimes temporal lag in correlation is also expected if the basin size is large ([Peters, 2003](#)). Moreover, the measurement of precipitation in a catchment is much easier, economical and comparatively more accurate than the measurement of soil moisture or the runoff components. Hence, an understanding of temporal transition between the different types of drought will make their management easier and economical.

For studying and characterizing droughts, several indices are proposed by different investigators. These indices can be broadly divided into three categories depending on the type of drought. For quantifying meteorological drought, the Standardized Precipitation Index (SPI) is one of the widely used indices. It is a normalized index, so its value is unaffected by the local climatic and geographical condition. Hence, the intensity of droughts occurring at different places having different climatic conditions can be compared using SPI ([McKee et al., 1993](#)). Similar standardized indices,

* Corresponding author. Tel.: +91 3222 283442; fax: +91 3222 282254.

E-mail address: rajib@civil.iitkgp.ernet.in (R. Maity).

which are mathematically consistent, are also used for soil moisture and stream flow, named as Standardized Soil Moisture Index (SSMI) and Standardized Stream Flow Index (SSFI) respectively (Mo, 2008; Shukla et al., 2011).

Wavelet Transform (WT) is a mathematical tool effectively utilized to identify and separate the slowly and rapidly varying components of any time series. Wavelet is a finite disturbance of limited duration with a mean of zero. WT expresses the signal as the coefficients of shifted and scaled version of the original wavelet, known as mother wavelet. Wavelet can be varied in both frequency and time domain by changing the scale and shift parameter respectively (Daubechies, 1988; Burrus et al., 1998). Hence, WT represents the signal in frequency–time domain. Unlike the Fourier transform (Duhamel and Vetterli, 1990; Fleming et al., 2002), WT saves the time information available in signal also. Thus, the WT is better for analyzing non-stationary signals (Mallat and Zhang, 1993; Cao et al., 1995; Burrus et al., 1998; Soltani, 2002). The mathematical details of wavelet transform is presented in Appendix-A including Haar wavelet, which is used in this study.

A special type of wavelet, discrete wavelets is formed when shift and scaling parameters are taken discrete instead of continuous variables in wavelet function. In the Discrete Wavelet Transform (DWT), signal is convolved with the pair of low pass and high pass filters followed by subband down sampling producing two components. The component, which is obtained by passing the signal through low pass filter is called approximation component and the other component is called detailed component. The approximated component shows the trend of the original drought index series and it possesses slow dynamics. On the other hand, the detailed component shows the local details of the series and it possesses fast dynamics (Burrus et al., 1998; Mallat, 1999). The subband down sampling is done as per the Nyquist–Shannon sampling theorem (Shannon, 1949) to reduce data redundancy by removing the component values falling on even positions, but subband down sampling introduces time-invariance in the DWT components. As a result minor changes in the drought series may result in large changes in DWT components. Moreover, as in DWT decomposition, the components series and input drought series do not have the same length, it becomes difficult for further analysis towards prediction. To overcome these problems, Stationary Wavelet Transform (SWT) is designed which avoid the subband down sampling. The SWT results in components having data redundancy, but have the same length as that of the original drought series. Multi-Resolution Analysis (MRA) is a procedure in which SWT is applied multiple time to the approximate component to get the component at even lower frequency ranges. Thus, Multi-Resolution Stationary Wavelet Transform (MRSWT) may be very effective for the decomposition of the drought indices series. Mathematical details of MRA are provided in Appendix-B.

General applications of wavelet based approaches are being increasingly used in studying hydrological processes, such as time series analysis (Pan et al., 2005; Westra and Sharma, 2006; Karthikeyan and Nagesh Kumar, 2013; Sang et al., 2016), water quality in agricultural watershed (Kang and Lin, 2007), stream flow prediction (Smith et al., 1998; Bayazit et al., 2001), rainfall-runoff relation (Labat et al., 2000), drought forecasting (Ozger et al., 2011; Özger et al., 2012), trend analysis of evapotranspiration and its relation to draught (Madhu et al., 2015), prediction of meteorological drought (Kim and Valdés, 2003), spatial and temporal variability of drought (Ujeneza and Abiodun, 2014; Wang et al., 2015), variability in monsoon rainfall in West Africa (Dieppois et al., 2013) etc.

Most of the existing studies focus on prediction of a particular type of drought considering its key variable(s). For instance, Kim and Valdés, (2003) utilized a conjecture model of wavelet transform and ANN for predicting meteorological drought using the past

value of the respective drought index. Ozger et al., (2011) utilized wavelet fuzzy logic model for forecasting meteorological drought using temperature, precipitation and climate indices like (El Nino 3.4 and Pacific Decadal Oscillation). Özger et al., (2012) utilized wavelet fuzzy logic model for forecasting meteorological drought by using El Nino 3.4 and persistence of the drought index series. However, these studies do not take the inter-relation and propagation of different types of droughts into account. The objective of this study is to identify the temporal transition from one type of drought to another considering its time-varying characteristics at constituent wave components of their standardized indices. To achieve this objective, the rainfall, soil moisture and streamflow data is converted to their standardized drought indices. These standardized indices are mathematically consistent and independent of basin geomorphological characteristics. Three time series (one each for each drought type being studied) are first investigated for their pair-wise lagged correlation usually resulting from lagged response of basin under study. The drought series are then decomposed using Wavelet Transform. For modeling drought inter-relation, seven different models (model 1 to 7) based on the information obtained by analysis of lagged correlation and wavelet coefficients of the drought indices are then established to predict a successor drought from a predecessor one.

2. Study area and data

The Upper Mahanadi Basin, a part of the Mahanadi river basin, is considered as the study area to demonstrate the proposed methodology. The study basin is mostly located in the state of Chhattisgarh in India as shown in Fig. 1. The area of the study watershed is 29,645 km². The approximate location of the study area is 20°N to 23°N latitude and 80.5°E to 82.5°E longitude. Daily rainfall data and monthly soil-moisture data for the study area are obtained for the period of 1971 to 2005 from the India Meteorological Department (IMD) (Rajeevan and Bhate, 2008) and Climate Prediction Centre (CPC) of the National Oceanic and Atmospheric Administration (NOAA) (CPC, 2014; Fan et al., 2004) respectively. These data are available at a spatial resolution of 0.5° latitude × 0.5° longitude and the data are taken from grid point lying within the study basin as shown in Fig. 1. Daily rainfall data at each grid point is converted to monthly rainfall depth by accumulating it over the month. Moreover, the soil moisture data obtained from CPC, NOAA, is produced using leaky bucket method (Huang et al., 1996; van den Dool, 2003). This method estimates the soil moisture using water balance equation and hence, the data is free from any human intervention.

Daily stream flow data at the outlet of the basin (Jondhra station) are collected for the period of June, 1979 to December, 2005 from the Water Resources Information System (India-WRIS version 4, 2014) in India. Being restricted by the stream flow data availability, the study period is considered as January 1980 to December 2005. The daily stream flow data is converted to monthly data by taking average over all the days. The basic statistical properties of the all the data are shown in Table 1.

3. Methodology

Methodology is based on the hypothesis that the water deficit is expected to propagate through the hydrological cycle with time and the time series of relevant drought indices can be considered as a wave or signal. Further, it is expected that the effect of individual factors (like climatological, topological, etc.) affecting precipitation, soil moisture or streamflow should get manifested as various constituent waves in the respective drought index time series. So, we tried to study the drought indices time series at component

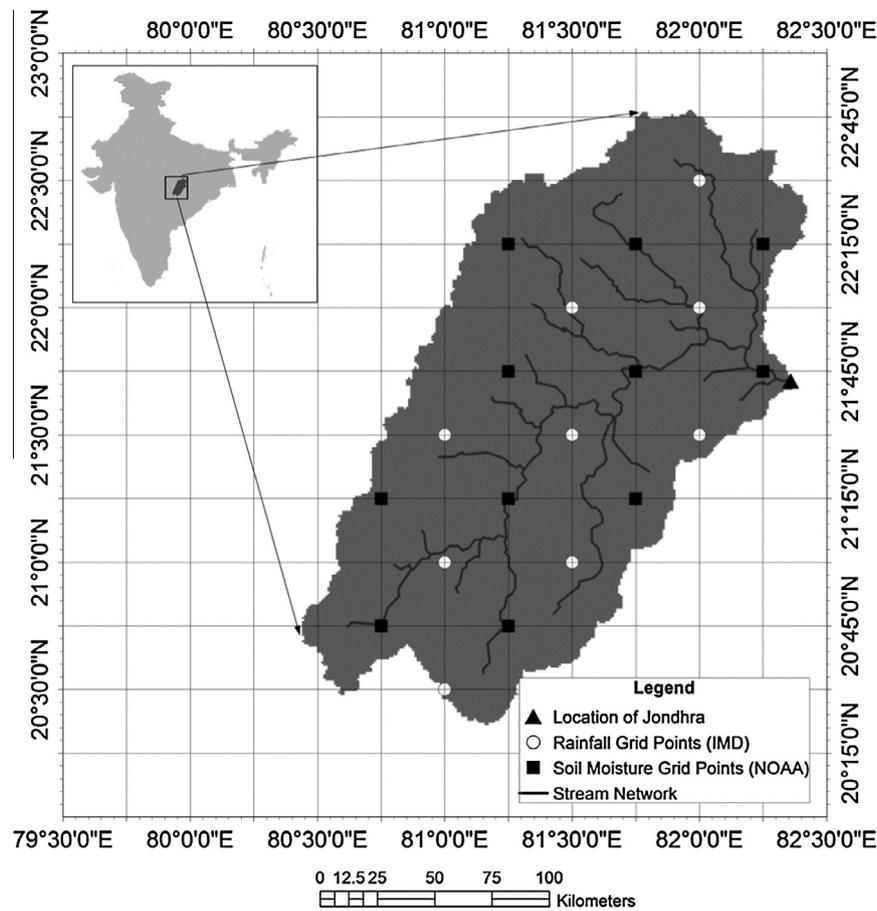


Fig. 1. Location map of upper Mahanadi River basin (Study Area) with rainfall grid points and soil moisture grid points.

Table 1
Statistical details of data used.

Variables	Value or range of values of different statistical parameters			
	Mean	Standard deviation	Skewness coefficient	Kurtosis ^a
Monthly total rainfall in mm (not spatially averaged)	86.62–117.81	125.49–183.55	1.43–1.88	4.10–6.81
Spatially averaged monthly total rainfall in mm	97.83	137.08	1.44	4.07
Monthly mean soil moisture in mm (not spatially averaged)	324.18–373.30	135.78–147.14	0.19–0.38	1.96–2.14
Spatially averaged monthly mean soil moisture in mm	343.82	141.51	0.29	2.04
Monthly mean streamflow in cumec	272.40	502.00	2.65	11.13

^a Kurtosis for normal distribution is 3.

level for revealing hidden interrelation between them. Different models are formulated using the information of time lag to predict the predictand drought index.

Overall methodology is broadly divided into three modules – (i) drought characterization and generation of its time series, (ii) Study of lagged correlation between the components of drought index to check whether delayed response of one drought index exist on the other, (iii) Formulation of different models considering the lagged information of predictor drought index, based on MRSWT and selection of most potential model type for prediction. The methodological overview is shown in Fig. 2 for a quick grasp. Details of these modules are presented in the following subsections.

3.1. Drought characterization through standardized indices

A wide range of drought indices is available in the literature. The suitability of a particular index depends on its application for

a particular problem (Keyantash and Dracup, 2002), and as such not a single index can be considered universal. Since this study deals with meteorological, agricultural and hydrological droughts, it is required to utilize mathematically similar indices for all these droughts. Keeping this point in mind, Standardized Precipitation Index (SPI), Standardized Soil Moisture Index (SSMI) and Standardized Stream Flow Index (SSFI) are used for meteorological, agricultural and hydrological droughts calculated using monthly precipitation, soil moisture and streamflow (at basin outlet) respectively. The concept of these drought indices is statistically similar to each other. SPI was first developed by McKee et al. (1993) for the Fort Collins, Colorado river basin in the USA. SPI can be defined as probability index of precipitation with respect to the standard deviation of precipitation for a given location and time period calculated from the historical precipitation data.

The computation of all the above indices (at a particular averaging time-scale, say 3-monthly) can be outlined in the following common steps –

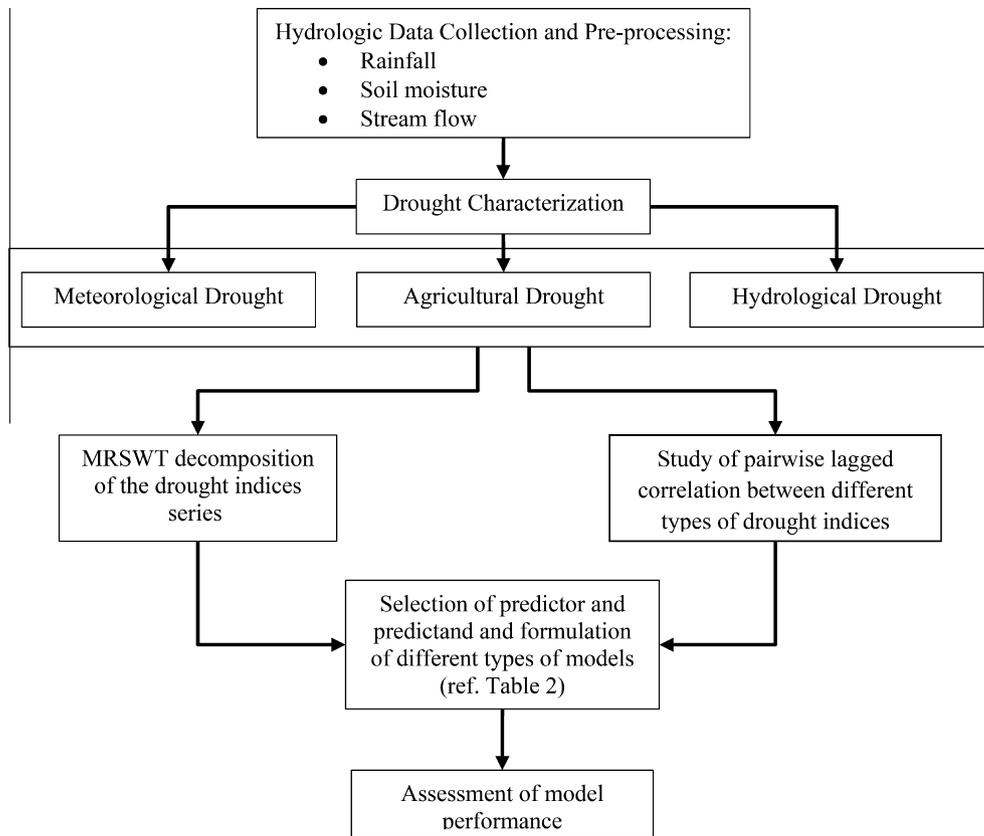


Fig. 2. Overview of methodological approach.

- (i) Time series of concerned variable is either accumulated or moving averaged over the desired temporal scale.
- (ii) A suitable probability density function (pdf) is fitted.
- (iii) From the fitted pdf, the Cumulative Probability Distribution (CDF) is obtained,
- (iv) Using the CDF, reduced variate of the concerned variable is computed,
- (v) The reduced variate is transformed to a standard normal variate (mean = 0 and standard deviation = 1). This is the desired standardized index.

Conceptually, all these indices represent the number of standard deviations away (above or below) that a particular value is from the mean. All these indices can have both positive and negative values, positive value showing a surplus and negative value showing a deficit or drought. Additionally, it should be noted that steps (ii) and (iv), during the index calculation, take out the climatology (hence seasonality) in the calculated drought index.

Depending on the characteristics of the basin under study, some time lag is expected before effect of predecessor drought situation is felt over a successor one. The time lag can also originate due to nature of variable under study. For example, soil moisture is expected to have higher memory effect than precipitation, so agricultural drought will be comparatively slower to initiate and to end. To quantify the time lag in drought propagation, lagged correlations between different predictand–predictor drought indices are studied. The lag period having highest correlation is taken as measure of delay in response that predictor drought series has on the predictand drought series.

3.2. Modeling of drought indices interrelation

The modeling of the interrelation between the drought indices components can lead to prediction of successor drought from the

state of predecessor one. The modeling may be achieved through any traditional approach like, Multiple Linear Regression (MLR), Auto-Regressive Integrated Moving Average model with exogenous inputs (ARIMAX) or even soft computing approaches, such as Artificial Neural Network (ANN). In the present study, models are formulated in two versions (keeping input and output variable same) – one using feed-forward ANN with single hidden layer and other using MLR. The independent drought indices are decomposed into components using MRSWT upto level 2. The mathematical details of MRSWT are presented in Appendix-B. One important issue needs to be mentioned here. By using MRSWT based decomposition, the prediction of droughts leads to the problem of predicting the slow and fast dynamic components separately. This approach is more fruitful as prediction of slow dynamic or approximate component can be done with more confidence because variations are expected to be smaller and less abrupt compared to fast dynamic or detailed signal component. Prediction of the fast dynamic component is challenging as the model has to learn the fast dynamic and reduce noise simultaneously. The challenge can be solved by overfit/underfit tradeoff. Learning fast dynamic can lead to under fitting but learning to predict noise cause over fitting (Soltani, 2002).

The decomposition through MRSWT results into three components (d_1 , d_2 and a_2) for each of the drought indices. Models are formulated on the assumption that a dependent drought index or its components are affected by all the decomposed components of the independent drought index simultaneously with some delay. The information about the delay in response is given due consideration in formulation of models. The details of seven different models so formulated are presented in Table 2. It should also be noted that minimum lead period also depend upon the level of decomposition being used to avoid the use of future information during the prediction. Since MRSWT with level 2 is used in this study, minimum

Table 2

Details of different types of models (No. 1 to 7). The function f is either of multiple linear regression or feed forward ANN function with single hidden layer and the function g represents wavelet reconstruction function. Subscripts a_2 , d_2 and d_1 represent the decomposed components of the respective drought index series at level 2. T_1 is equal to 2^D , where D is the level of decomposition, i.e. $T_1 = 2^2 = 4$ and $T_{n+1} = T_n + 1$ for $n = 1, 2, \dots$

Model No.	Model description
1	$SSMI(t) = f\left(\begin{matrix} SPI_{a_2}(t - T_1), SPI_{d_2}(t - T_1), SPI_{d_1}(t - T_1), \\ SPI_{a_2}(t - T_2), SPI_{d_2}(t - T_2), SPI_{d_1}(t - T_2) \end{matrix}\right)$
2	$SSMI(t) = f\left(\begin{matrix} SPI_{a_2}(t - T_1), SPI_{d_2}(t - T_1), SPI_{d_1}(t - T_1), \\ SPI_{a_2}(t - T_2), SPI_{d_2}(t - T_2), SPI_{d_1}(t - T_2), \\ SSMI_{a_2}(t - T_1), SSMI_{d_2}(t - T_1), SSMI_{d_1}(t - T_1) \end{matrix}\right)$
3	$SSMI_C(t) = f\left(\begin{matrix} SPI_{a_2}(t - T_1), SPI_{d_2}(t - T_1), SPI_{d_1}(t - T_1), \\ SPI_{a_2}(t - T_2), SPI_{d_2}(t - T_2), SPI_{d_1}(t - T_2), \\ SSMI_{a_2}(t - T_1), SSMI_{d_2}(t - T_1), SSMI_{d_1}(t - T_1) \end{matrix}\right)$ for $c \in d_1, d_2$ and a_2 $SSMI(t) = g(SSMI_{d_1}, SSMI_{d_2}, SSMI_{a_2})$
4	$SSFI(t) = f\left(\begin{matrix} SPI_{a_2}(t - T_1), SPI_{d_2}(t - T_1), SPI_{d_1}(t - T_1), \\ SPI_{a_2}(t - T_2), SPI_{d_2}(t - T_2), SPI_{d_1}(t - T_2) \end{matrix}\right)$
5	$SSFI(t) = f\left(\begin{matrix} SPI_{a_2}(t - T_1), SPI_{d_2}(t - T_1), SPI_{d_1}(t - T_1), \\ SPI_{a_2}(t - T_2), SPI_{d_2}(t - T_2), SPI_{d_1}(t - T_2), \\ SSMI_{a_2}(t - T_1), SSMI_{d_2}(t - T_1), SSMI_{d_1}(t - T_1) \end{matrix}\right)$
6	$SSFI(t) = f\left(\begin{matrix} SPI_{a_2}(t - T_1), SPI_{d_2}(t - T_1), SPI_{d_1}(t - T_1), \\ SPI_{a_2}(t - T_2), SPI_{d_2}(t - T_2), SPI_{d_1}(t - T_2), \\ SSMI_{a_2}(t - T_1), SSMI_{d_2}(t - T_1), SSMI_{d_1}(t - T_1), \\ SSFI_{a_2}(t - T_1), SSFI_{d_2}(t - T_1), SSFI_{d_1}(t - T_1) \end{matrix}\right)$
7	$SSFI_C(t) = f\left(\begin{matrix} SPI_{a_2}(t - T_1), SPI_{d_2}(t - T_1), SPI_{d_1}(t - T_1), \\ SPI_{a_2}(t - T_2), SPI_{d_2}(t - T_2), SPI_{d_1}(t - T_2), \\ SSMI_{a_2}(t - T_1), SSMI_{d_2}(t - T_1), SSMI_{d_1}(t - T_1), \\ SSFI_{a_2}(t - T_1), SSFI_{d_2}(t - T_1), SSFI_{d_1}(t - T_1) \end{matrix}\right)$ for $c \in d_1, d_2$ and a_2 $SSFI(t) = g(SSFI_{d_1}, SSFI_{d_2}, SSFI_{a_2})$

lead period for prediction is 2^2 i.e. 4. The model that predict dependent drought index component instead of drought index time series (such as model 3 and 7) need the wavelet reconstruction to be done on the predicted components. This procedure returns the components from frequency–time domain into amplitude–time domain.

All the proposed models, except those based on ANN approach are tested using two different validation schemes – I and II. ANN based models are validated with scheme I only. Details of these schemes are discussed below. These validation schemes are also illustrated in Fig. 3.

- i. Scheme I – Fixed Development and Testing Period: In this scheme everything besides the development period is considered testing period and the data set remain stationary in one model calibration–prediction run. The parameters of the model are estimated during the development period. All of the testing period data are predicted in the next model run and compared. So, in this validation scheme, a model runs only two times, one for calibration in development period and other for prediction of testing data set.
- ii. Scheme II – Moving Window Approach: In this scheme, testing period data length is same as that of development period, but these data periods are moving over time from one iteration to another. The model is first developed with the development period data set and for prediction, the window is shifted by one time step and the data from this new time step is considered in the testing period pool. Hence, though there is overlap between the development and testing period datasets, only one time step of the time series is considered as predicted in each iteration. For the next iteration, both development and testing periods are shifted by one time step and the process is continued until the prediction of whole remaining time series is complete. This scheme is

useful to update the model parameters to capture any slow moving changes in the time series, particularly in the context of climate change. Similar kind of methodology has been used to predict components of wavelet transform (Mabrouk et al., 2008), analyse the sensitivity of rainfall–runoff model (Massmann et al., 2014), assess impact of climate change on streamflow in Weihe River (Jiang et al., 2015).

3.3. Model performance evaluation

Performances of different models are assessed based on four statistical measures, namely correlation coefficient (r), Refined Index of Agreement (D_r), Nash–Sutcliffe efficiency (NSE) and unbiased Root Mean Square Error (uRMSE). Expressions for r and NSE can be found elsewhere (Krause et al., 2005). The expression of D_r is given by (Willmott et al., 2012)

$$D_{r_frac} = \frac{\sum_{i=1}^n |Y_i - X_i|}{2 \sum_{i=1}^n |X_i - \bar{X}|} \tag{12a}$$

$$D_r = \begin{cases} 1 - D_{r_frac} & \text{for } D_{r_frac} \leq 1 \\ \frac{1}{D_{r_frac}} - 1 & \text{for } D_{r_frac} > 1 \end{cases} \tag{12b}$$

where X_i and Y_i are the i^{th} observed and predicted values and n is the total number of observations. uRMSE is the RMSE calculated between the deviations of observed and predicted values from their respective means. Lower the value of uRMSE, better the model performance.

4. Results and discussions

Monthly data of rainfall and soil moisture are distributed spatially and needed to be spatially averaged. However, one-way analysis of variance (ANOVA) and Bonferroni test are conducted to check the significance of spatial viability of the data at different grid points. For rainfall, only in one month (August) and only at one site, site mean differs from the overall monthly mean. In all other months, the overall monthly mean across all sites is not significantly different from the individual site means. In case of soil moisture, the overall spatial monthly mean is in good agreement with individual site means in all cases (all months and all sites).

Taking monthly rainfall depth, soil moisture time series and stream flow series as input SPI, SSMI and SSFI respectively are calculated using a mixed distribution – Gamma distribution for non-zero values with probability mass at zero. For monthly rainfall depth accumulation was done during SPI calculation but for all other variables moving average are calculated during index calculation. Notations of SPI-1, SSMI-1 and SSFI-1 are used for 1-month time scale. Similarly, SPI-3, SSMI-3 and SSFI-3 are used for 3-month time scale. SPI-3, SSMI-3, and SSFI-3 time series are shown in Fig. 4. By visual inspection of the figure, it can be inferred that indices does not possess seasonality. For studying the interrelation and propagation of different types of droughts, possible predecessor–successor or predictor–predictand pairs are selected. For instance, SPI is taken as predictor for SSMI and SSFI; SSMI is considered a predictor for SSFI. These relations are deemed valid irrespective of averaging period. The study period chosen is January, 1980 to December, 2005, so all drought index series are having 312 data points. First 160 data points are considered for the initial scrutiny and model development. The rest of the data are used for model testing. For initial scrutiny, the pairwise correlation coefficients (r) and the refined index of agreement (D_r) between the indices are computed and the results are tabulated in Table 3. From Table 3, the correlation coefficient and refined index of agreement is higher for 3-month time scale indices. It is due to higher average period

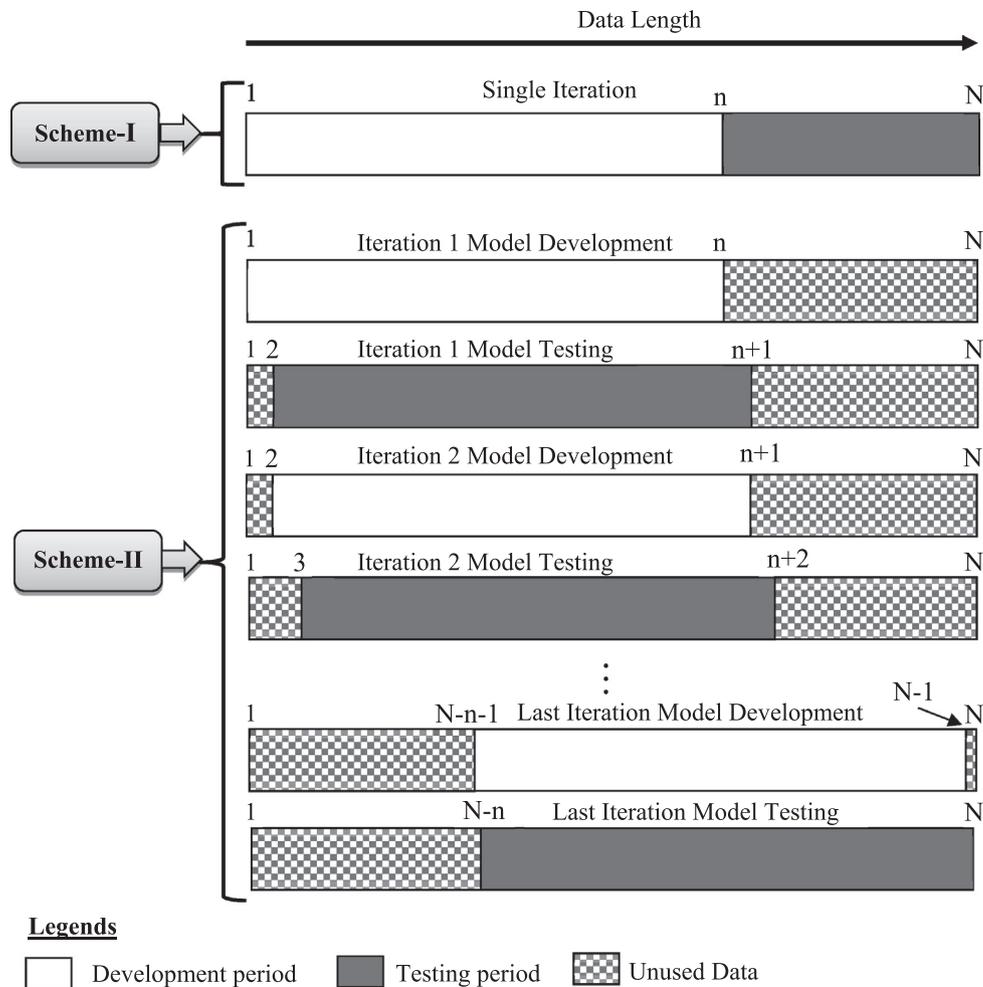


Fig. 3. Schematic diagram of two types of validation schemes. In scheme II, at any model testing iteration only the last value is recorded for performance assessment though the testing period overlaps the model development period of the same iteration.

used to calculate the indices, which lead to more smoothening in data. Moreover, it is found that linear association between SSFI and SPI and SSFI and SSMI are similar and statistically significant. This observation supports the fact that both precipitation and soil moisture have their influence on streamflow variation. Direct runoff due to precipitation events in catchment may affect the streamflow immediately, whereas the soil moisture is expected to affect streamflow by delayed subsurface flow. This suggests to incorporate the combination of different predictors (say, SPI and SSMI) with suitable lag to achieve possible better performance in predicting target drought index (say SSFI). It should also be noted that, so far the lagged information is not considered from any of the predictor. The values in Table 3 are used as a reference for comparing the performance of different models as mentioned in the methodology. Any model that can exhibit better performance compared to these values can be considered as efficient and improvement over these reference values can be quantified.

The lagged correlation between all possible predictor–predictand drought index pairs is then calculated. The results are shown in Fig. 5. It is noticed that the correlation coefficient between SSMI-3 and SPI-3 with lag 1, is the highest. This result suggests that SSMI has higher memory and changes slowly as compared to SPI. Thus, utilization of lagged values from predictor time series may enhance the prediction performance. In case of SSFI-3 and SSMI-3 as well as SSFI-3 and SPI-3, the correlation coefficient is highest without any lag. These observations suggest that SSFI is

affected by both SPI and SSMI; utilization of values from these two predictors combined should enhance the prediction performance. In all predictor–predictand pairs, the value of correlation coefficients decreases gradually with the further increase in lag. Moreover, the lag considered in modeling of inter-relation of drought indices should be either equal or greater than the averaging period and minimum lead period requirement as discussed in Section 3.2. To reiterate, minimum required lead period for prediction is 4 since MRSWT with level 2 is used in this study. In case of drought indices calculated using 3 month accumulation, SPI-3 with lag 4 and 5 may be considered while predicting SSMI-3. Similarly, for SSFI-3, SPI-3 with lag 4, 5 and SSMI-3 with lag 4 may be important. Different combinations of the predictors, lead to seven different models denoted as model 1 to model 7 as detailed in Table 2. Model 1, 2 and 3 are used for predicting SSMI and model 4 to 7 is used for predicting the SSFI. During the application of models, the predictor drought time series is first decomposed into its components using MRSWT. For example, the components of SPI-3 are shown in Fig. 6. The predictand drought index or its components are predicted using the components of predictor drought indices. The model performances during the development period and testing period are tabulated in Tables 4 and 5 respectively. It should be noted that for ANN based model, each model is trained 200 times and best trained model is selected for prediction. During development period, the model performance is found to improve (Tables 4a and 4b) as compared to Table 3. For instance, corresponding to one month

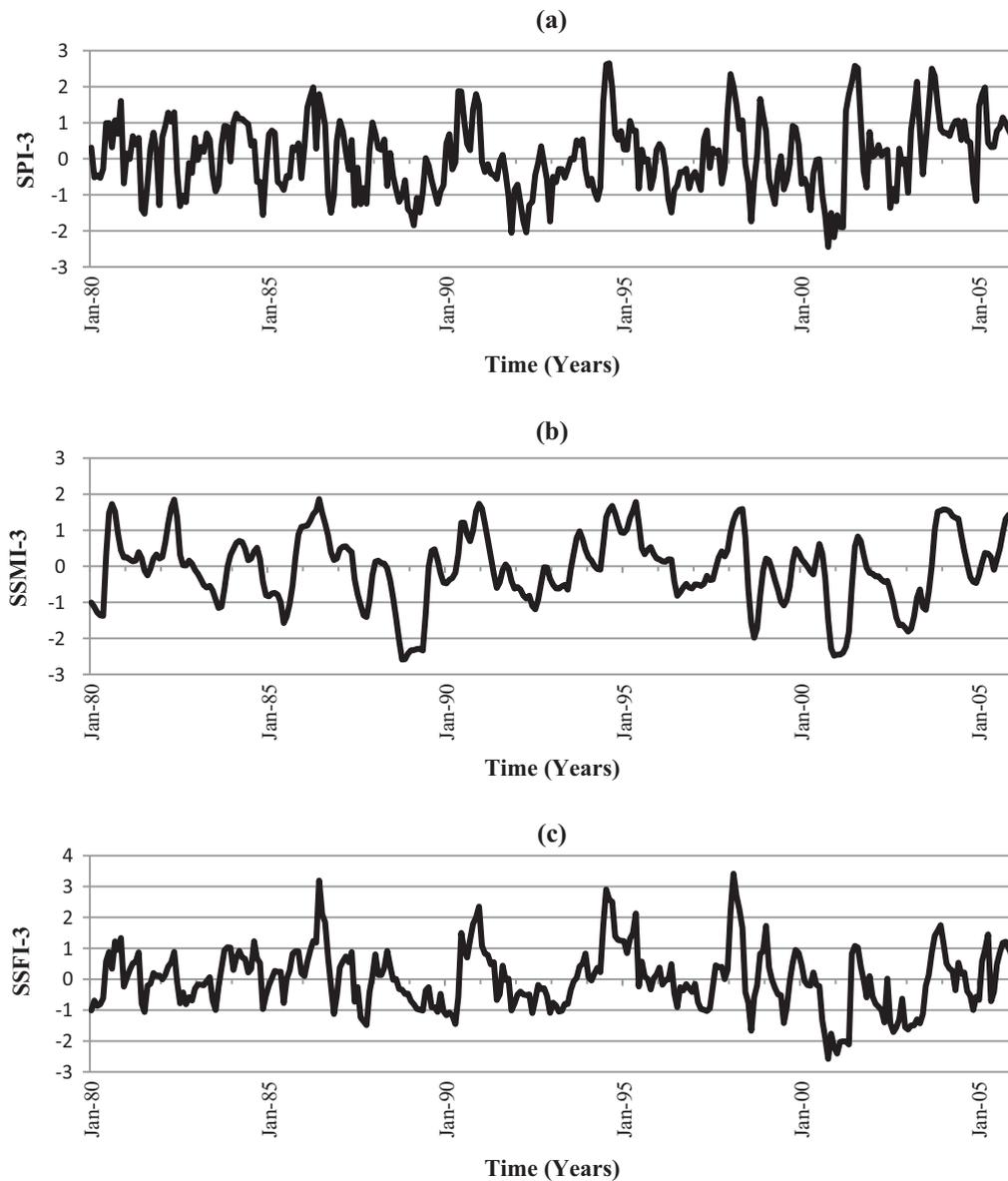


Fig. 4. Time series of (a) SPI-3 b) SSMI-3 (c) SSFI-3 for the period January, 1980 – December, 2005.

Table 3
Correlation coefficient (r) and refined index of agreement (D_r) for different drought indices pairs during development period.

Averaging period (in months)	Performance statistics	Predictand drought index	Predictor drought index	
			SPI	SSMI
1	r	SSMI	0.401	1.000
		SSFI	0.588	0.607
	D_r	SSMI	0.436	1.000
		SSFI	0.516	0.502
3	r	SSMI	0.590	1.000
		SSFI	0.682	0.661
	D_r	SSMI	0.543	1.000
		SSFI	0.585	0.564

averaging period, the coefficient of correlation for MLR version of model 2 during development period is 0.831 between observed and predicted SSMI-1 (Table 4a) which is higher than the coefficient of correlation (0.401) between observed SSMI-1 and SPI-1

(Table 3). Though it is apparent that model 2 (for SSMI) and 6 (for SSFI) are the best among other alternatives, it should be noted that the previous values of SSMI and SSFI are used in model 2 and 6 respectively. On the other hand, model 1 use only information of SPI (with lags) and model 5 uses only SPI and SSMI (with lags), not the previous values of predictand series. Thus, the merit of model 1 (in case of SSMI) and model 5 (in case of SSFI) should be duly credited. Moreover, model 3 and 7 (these models use inverse wavelet transform to generate the predicted drought time series) also show comparable performance to model 2 and 6. It is also noticed from Table 4b that ANN versions of models are performing better than MLR version in most of the cases during model development period. However, the difference in performance between MLR and ANN is found to decrease when the averaging period is higher, i.e., 3. For example, the correlation coefficient between observed and predicted SSMI-1 for MLR version and ANN version of model 2 are 0.831 and 0.873 respectively but for SSMI-3 it is 0.941 and 0.962 respectively. The performance of models predicting SSFI is in general, inferior compared to model predicting SSMI. The decrease in performance may be due to combined effect of

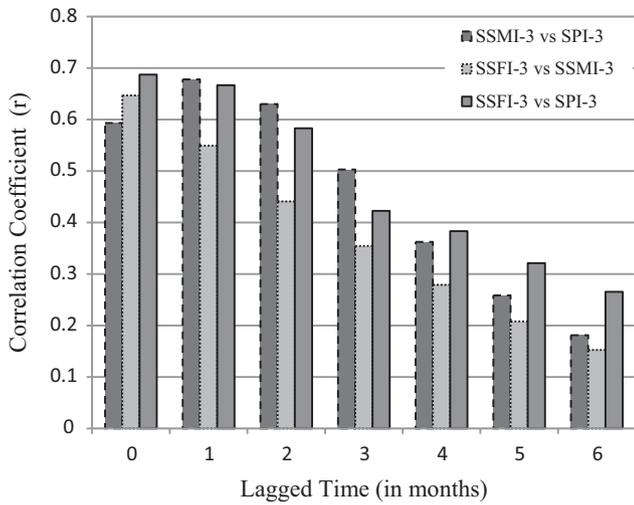


Fig. 5. Pairwise linear correspondence between SPI, SSMI and SSFI with lags during model development period. Lags are applicable for the second index as shown in the legends for different pairs.

higher memory of soil moisture and the fact that many factors that affect streamflow, like evapotranspiration, air temperature, etc., are not considered while predicting the SSFI.

Model performance during testing period is shown in Table 5. As mentioned earlier, two different validation schemes are followed for MLR version of models. For MLR version of models predicting SSMI, it is noticed that the model performance is either better or comparable with validation scheme I as compared to validation scheme II. Similarly, for MLR models predicting SSFI, model performance is either better or comparable with validation scheme

II as compared to validation scheme I. This observation suggests that streamflow perhaps has time varying correspondence or dynamic relationship with other drought indices i.e., its relationship with other variable has changed with time, so validation scheme II, which is more competent in modeling these dynamic relationships, produces better results. For example, with the validation scheme I and for predicting SSFI-3, the model 6 performance measures (r , D_r , NSE and uRMSE) are 0.792, 0.693, 0.610 and 0.698 respectively, whereas the same with validation scheme II are 0.801, 0.712, 0.638 and 0.682 respectively. Thus, the validation scheme II may be considered as more suitable where the correspondence between predictor and predictand may get modified over time due the various reasons, including changing basin characteristics, climate regime, etc. Interestingly, during testing period, models using MLR version are found to perform comparable to ANN version in most of cases. For example for SSFI-3 and validation scheme I, MLR based model 6 performance measures (r , D_r , NSE and uRMSE) are 0.792, 0.693, 0.610 and 0.698 respectively and corresponding value for ANN based model are 0.698, 0.617, 0.450 and 0.822 respectively. This observation suggests that decomposed wavelet coefficient has linear relationship, so ANN version could not add much to the performance achieved by MLR version. The NSE value higher than 0.80 is observed in model 2 (in case of indices with 3 month time scale), so this model can be utilized in practical studies. Moreover, as stated earlier the performance of model predicting SSFI is inferior to model predicting SSMI in testing period too. The scatter plots for SSMI-3 and SSFI-3 modeled by MLR version of model 1 to 7 for validation scheme II is shown in Fig. 7.

Two types of sensitivity analysis, namely mother wavelet sensitivity and development data length sensitivity are carried out to understand the effect of mother wavelet and the length of development period on the model performance. Mother wavelet sensitivity

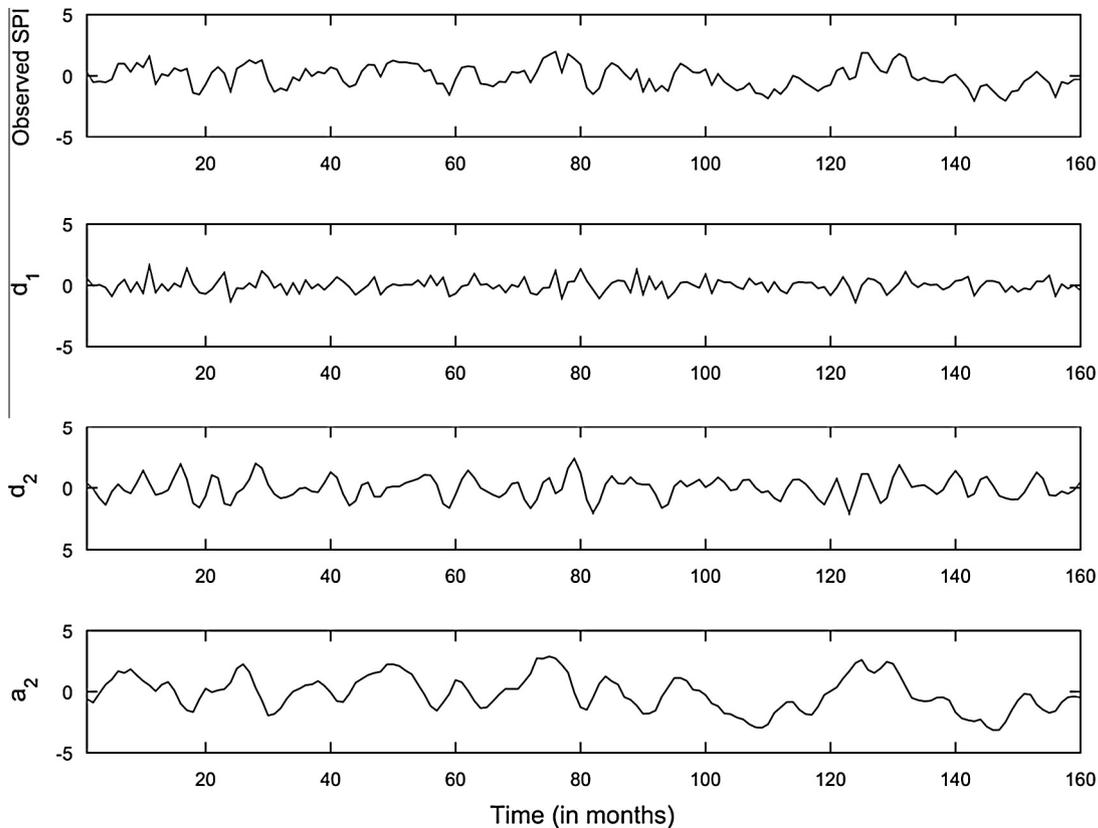


Fig. 6. Observed SPI-3 and its decomposed components up to level 2, i.e., a_2 , d_2 and d_1 , using Haar MRSWT. Figure shows first 160 data points of decomposed series, i.e. development period for models. Such decomposed series for SSMI-3 and SSFI-3 are also obtained (not shown).

Table 4a
Performance of model no. 1 to 7 during development period using MLR.

Averaging period (in months)	Performance measures	Model No.						
		1	2	3	4	5	6	7
1	<i>r</i>	0.657	0.831	0.766	0.581	0.589	0.607	0.590
	<i>D_r</i>	0.612	0.733	0.645	0.620	0.620	0.626	0.610
	NSE	0.431	0.690	0.519	0.338	0.347	0.369	0.336
	uRMSE	0.730	0.538	0.671	0.660	0.655	0.644	0.661
3	<i>r</i>	0.748	0.941	0.873	0.736	0.743	0.792	0.723
	<i>D_r</i>	0.666	0.837	0.723	0.687	0.690	0.719	0.661
	NSE	0.559	0.886	0.709	0.541	0.552	0.628	0.489
	uRMSE	0.632	0.322	0.513	0.559	0.552	0.503	0.590

Table 4b
Performance of model no. 1 to 7 during development period using ANN.

Averaging period (in months)	Performance measures	Model No.						
		1	2	3	4	5	6	7
1	<i>r</i>	0.681	0.873	0.798	0.563	0.428	0.567	0.590
	<i>D_r</i>	0.629	0.758	0.662	0.599	0.570	0.420	0.610
	NSE	0.456	0.753	0.567	0.277	0.178	-0.432	0.314
	uRMSE	0.710	0.472	0.637	0.675	0.735	0.869	0.655
3	<i>r</i>	0.579	0.962	0.874	0.619	0.818	0.888	0.689
	<i>D_r</i>	0.550	0.867	0.724	0.612	0.725	0.783	0.624
	NSE	0.152	0.925	0.711	0.336	0.647	0.777	0.384
	uRMSE	0.779	0.261	0.511	0.669	0.476	0.383	0.648

Table 5a
Performance for model no. 1 to 7 during model testing period using MLR (with both validation schemes I and II).

Averaging period (in months)	Validation scheme	Performance measures	Model No.						
			1	2	3	4	5	6	7
1	I	<i>r</i>	0.671	0.871	0.753	0.446	0.427	0.550	0.286
		<i>D_r</i>	0.628	0.761	0.647	0.544	0.538	0.580	0.507
		NSE	0.423	0.756	0.502	0.155	0.134	0.278	0.031
		uRMSE	0.766	0.507	0.724	1.018	1.031	0.949	1.098
	II	<i>r</i>	0.652	0.862	0.733	0.496	0.571	0.642	0.480
		<i>D_r</i>	0.621	0.754	0.648	0.566	0.594	0.627	0.571
		NSE	0.405	0.742	0.492	0.202	0.309	0.409	0.225
		uRMSE	0.782	0.522	0.733	0.990	0.934	0.873	0.998
3	I	<i>r</i>	0.720	0.954	0.864	0.636	0.625	0.792	0.587
		<i>D_r</i>	0.634	0.850	0.725	0.611	0.607	0.693	0.581
		NSE	0.470	0.908	0.689	0.360	0.343	0.610	0.307
		uRMSE	0.709	0.304	0.562	0.879	0.889	0.698	0.931
	II	<i>r</i>	0.709	0.947	0.845	0.646	0.711	0.801	0.670
		<i>D_r</i>	0.633	0.843	0.721	0.630	0.659	0.712	0.614
		NSE	0.467	0.896	0.675	0.371	0.488	0.638	0.411
		uRMSE	0.716	0.324	0.575	0.869	0.801	0.682	0.871

Table 5b
Performance for model no. 1 to 7 during model testing period using ANN (only for validation scheme I).

Averaging period (in months)	Performance measures	Model No.						
		1	2	3	4	5	6	7
1	<i>r</i>	0.619	0.817	0.739	0.504	0.423	0.523	0.372
	<i>D_r</i>	0.592	0.712	0.643	0.533	0.544	0.517	0.507
	NSE	0.292	0.659	0.495	0.149	0.159	0.030	0.051
	uRMSE	0.827	0.596	0.730	0.990	1.038	1.100	1.070
3	<i>r</i>	0.638	0.884	0.815	0.596	0.566	0.698	0.536
	<i>D_r</i>	0.538	0.778	0.687	0.592	0.515	0.617	0.561
	NSE	0.138	0.779	0.623	0.293	-0.028	0.450	0.228
	uRMSE	0.789	0.473	0.618	0.939	1.059	0.822	1.001

analysis on MLR version of the model was carried out using 160 development period data and with three mother wavelets namely Haar, Biorthogonal 1.1 and Reverse Biorthogonal 1.1. The model

performances are found to be mostly insensitive to mother wavelet. Development period data length sensitivity is carried out on the MLR version of the models for development period data length

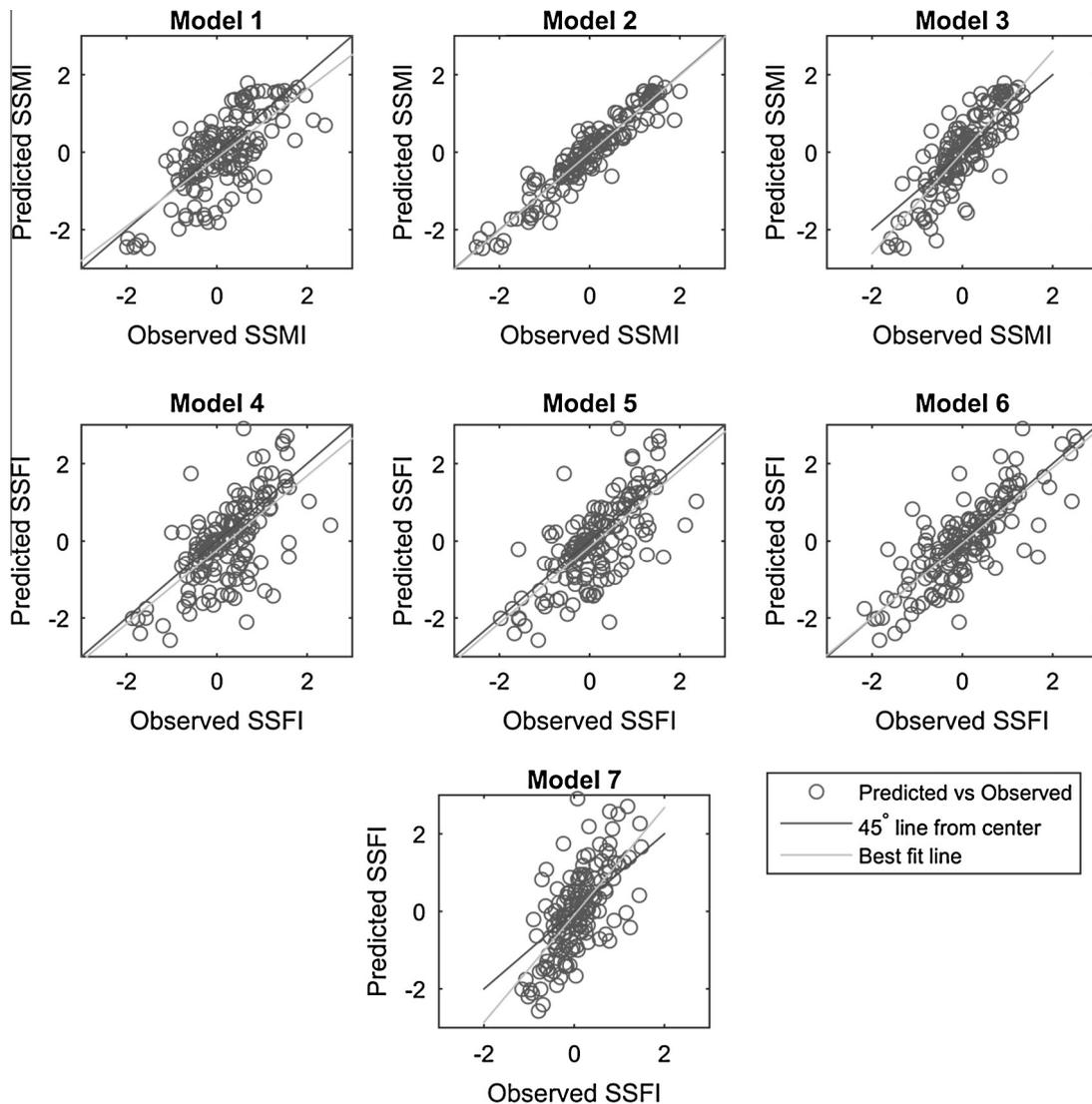


Fig. 7. Scatter plot between observed and predicted SSMI-3 and SSFI-3 by MLR version of models 1 to 7 during the testing period with validation scheme II.

ranging from 16 to 192. Model performance is found to depend on the development data length, but its variation is very less beyond the length of 140 data points.

Overall methodological framework of the proposed models is general in nature. These models can easily be applied to other basins. However, the methodology heavily depends on the availability of historical data and the models need to be calibrated each time the region/catchment and associated variable change. Although very few contributions (Das and Ghosh, 2014; Das and Maity, 2015) are noted so far on spatial transferability of the knowledge from one basin to another, the research on this aspect remains a challenging task and kept as the future scope of this study.

5. Summary and concluding remarks

Originating from precipitation deficit, different types of droughts propagate from one type to another – meteorological to agricultural to hydrological. This study focuses on the modeling of a successor drought using the input from its predecessor, using the potential of wavelet transforms. Following major conclusions are drawn from this study –

1. Given the complexity of the drought prediction, proposed wavelet based approach is established to be highly effective for foreseeing the agricultural or hydrological droughts knowing the meteorological drought status.
2. Drought series are better predicted at its constituent wave levels and MRSWT is an effective tool for decomposing the drought series. While considering three different mother wavelets namely Haar, Biorthogonal 1.1 and Reverse Biorthogonal 1.1, the model performances are found to be mostly insensitive to the choice of mother wavelet.
3. Model performance depends on the development data length, but its variation dies down beyond the data length of 140. However, moving window approach of validation scheme is found to be more competent in modeling the dynamic/time-varying association between different drought indices as compared to the scheme with fixed development and testing period.

The findings of this study help to assess the drought propagation in future with promising accuracy knowing the current and the previous SPI status. However, richer the information pool (say SSMI is also available along with SPI) the higher the accuracy in future assessment of SSFI is warranted. The results are expected to be useful for water resources managers for drought preparedness.

Acknowledgements

This work was partially supported by the Ministry of Earth Science (MoES), Government of India, through sponsored Project No. MoES/PAMC/H&C/30/2013-PC-II and Indian Space Research Organization (ISRO) through sponsored Project No. IIT/SRIC/CE/PMA/2013-14/5.

Appendix A. Wavelet Transform (WT)

Wavelet transformation, transforms any arbitrary signal into its constituent waves based on shifting and dilation of mother wavelet $\psi(t)$. The scale and shifted version of wavelet $\psi_{a,b}(t)$ can be represented by

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (\text{A-1})$$

where $a, b, t \in \mathbb{R}$, $a > 0$. $\psi(t)$ is a mother wavelet prototype, and a, b are scaling and shifting parameters respectively. Integration of $\psi(t)$ over real field is zero and usually its energy is kept to be unity.

Wavelet transform for a signal $f(t)$ is given by

$$W_f(a, b) = \frac{1}{\sqrt{C_\psi}} \int f(t) \psi_{a,b}^*(t) dt \quad (\text{A-2})$$

where $\psi^*(t)$ denote complex conjugate, $C_\psi = 2 \int |F(\psi(\omega))|^2 / \omega d\omega$, where F denote the Fourier transform given by $F(\psi(\omega)) = \int e^{-i\omega t} \psi(t) dt / \sqrt{2\pi}$.

If the basis wavelet or mother wavelet $\psi(t)$ is orthogonal then the inverse of wavelet transformation is given by

$$f(t) = \frac{1}{\sqrt{C_\psi}} \iint \frac{W_f(a, b) \psi_{(a,b)}(t)}{a^2} da db \quad (\text{A-3})$$

As $a, b \in \mathbb{R}$, they can be continuously varied leading to Continuous Wavelet Transform but sometimes these variables are sampled over discrete space time grid. The wavelet family so generated is termed discrete wavelet (Labat et al., 2000). If discrete wavelet is sampled over dyadic space time grid then they are called dyadic discrete wavelets (Cao et al., 1995). These wavelets are denoted by

$$\psi_{j,k}(t) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t}{2^j} - k\right) \quad (\text{A-4})$$

where $j, k \in \mathbb{Z}$, So its wavelet transform is given by

$$W_f(j, k) = \frac{1}{\sqrt{C_\psi}} \sum f(t) \psi_{j,k}^*(t) dt \quad (\text{A-5})$$

In a traditional DWT as per Nyquist–Shannon theorem (Shannon, 1949) subband coding or dyadic down sampling is done, but if subband coding is avoided the resulting procedure is called Stationary Wavelet Transform (SWT).

Dyadic discrete wavelet basis function has been proved orthogonal (Daubechies, 1988). The inverse of the wavelet transform is given by

$$f(t) = \frac{1}{\sqrt{C_\psi}} \sum_{j,k \in \mathbb{Z}} W_f(j, k) \psi_{j,k}(t) \quad (\text{A-6})$$

One of the popular mother wavelet functions is Haar, which is also known as daubechies1 or db1 (Burrus et al., 1998). Haar wavelet is defined as

$$H(t) = \begin{cases} 1 & 0 < t < 0.5 \\ -1 & 0.5 < t < 1 \\ 0 & \text{otherwise} \end{cases} \quad (\text{A-7})$$

Computationally, WT can be done using a pair of low pass and high pass filters. In DWT, signal convolution with low pass filter followed by dyadic down sampling give an approximate coefficient

and one obtained by using high pass filter and dyadic down sampling is called detailed coefficients (Burrus et al., 1998). For transforming a ‘one dimensional’ signal of ‘ n ’ (an even number) length, the high pass and low pass SWT filter have a dimension of $(n \times n)$. A Haar SWT high pass filter G and low pass filter H and can be constructed by following rule:

$$h_{ij} = \begin{cases} 1/\sqrt{2} & j \in \{i, (i+1) \bmod n\} \\ 0 & \text{Otherwise} \end{cases} \quad (\text{A-8})$$

$$g_{ij} = \begin{cases} (-1)^{i-j}/\sqrt{2} & j \in \{i, (i+1) \bmod n\} \\ 0 & \text{Otherwise} \end{cases} \quad (\text{A-9})$$

where $h_{ij} \in H$ and $g_{ij} \in G$. i is the number of row and j is the number of column. So, matrixes H and G are multiplied with $n \times 1$ dimension signal to calculate approximate and detailed coefficient respectively in SWT. In DWT as stated above subband coding is done by neglecting every second or component values falling on even positions to get the components. It should be noted that components in DWT have half the length than parent signal due to subband coding.

Appendix B. Multi Resolution Analysis (MRA)

Multi-Resolution Wavelet Transform (MRWT), also known as Multi-Resolution Analysis (MRA), can be performed by using low pass filter component as input to wavelet transform at each subsequent level (Labat et al., 2000). So, MRA helps in analysis of function at smaller frequency ranges, and hence it helps in increasing the accuracy of prediction. Multi Resolution analysis of signal can be done with both type of WT i.e. SWT or DWT. By using MRA a function in $L^2(\mathbb{R})$ can be represented as

$$f(t) = \sum_k a_{0,k} \varphi_{0,k}(t) + \sum_{j=0}^{\infty} \sum_k d_{j,k} \psi_{j,k}(t) \quad (\text{B-1})$$

where $a_{0,k}$ and $d_{j,k}$ is called coarse or approximating coefficient and detailed or wavelet coefficient respectively. $\varphi_{0,k}(t)$ is called a scaling function with shift k . The scaling function is associated with $\psi(t)$ wavelet function (Burrus et al., 1998). The scaling function for Haar wavelet is given by

$$\varphi(t) = \begin{cases} 1 & 0 < t < 1 \\ 0 & \text{Otherwise} \end{cases} \quad (\text{B-2})$$

$$\varphi_{0,k}(t) = \varphi(t - k) \quad (\text{B-3})$$

The coefficients involved in above Eq. (B-1) are calculated as follows:

$$a_{0,k} = \sum f(t) \varphi(t - k) \quad (\text{B-4})$$

$$d_{j,k} = \sum f(t) 2^{-j} \psi(2^{-j}t - k) \quad (\text{B-5})$$

As MRA is repeated application of selected WT on signal in first go and on approximate component on subsequent go, it can also be using pair of low pass (associated with scaling function) and high pass (associated with mother wavelet function) filters like the WT.

References

- American Meteorological Society, 1997. Policy statement: meteorological drought. Bull. Amer. Meteor. Soc. 78, 847–849.
- Bayazit, M., Onoz, B., Aksoy, H., 2001. Nonparametric streamflow simulation by wavelet or Fourier analysis. Hydrol. Sci. J. 46 (4), 623–634.
- Burrus, C.S., Gopinath, R.A., Guo, H., 1998. Introduction to Wavelets and Wavelet Transforms. Prentice Hall, New Jersey.
- Cao, L., Hong, Y., Fang, H., He, G., 1995. Predicting chaotic time series with wavelet networks. Physica D 85 (1), 225–238.

- CPC, 2014. Soil Moisture (V2). NOAA/OAR/ESRL PSD, Boulder, Colorado, USA. Available on <<http://www.esrl.noaa.gov/psd/data/gridded/data.cpcsoil.html>> (accessed in November, 2014).
- Das, M., Ghosh, S.K., 2014. A probabilistic approach for weather forecast using spatio-temporal inter-relationships among climate variables. In: Industrial and Information Systems (ICIIS), 2014, 9th International Conference on. IEEE, pp. 1–6.
- Das, S.K., Maity, R., 2015. A hydrometeorological approach for probabilistic simulation of monthly soil moisture under bare and crop land conditions. *Water Resour. Res.* 51, 2336–2355. <http://dx.doi.org/10.1002/2014WR016043>.
- Daubechies, I., 1988. Orthonormal bases of compactly supported wavelets. *Commun. Pure Appl. Math.* 41 (7), 909–996.
- Dieppois, B., Diedhiou, A., Durand, A., Fournier, M., Massei, N., Sebag, D., Xue, Y., Fontaine, B., 2013. Quasi-decadal signals of Sahel rainfall and West African monsoon since the mid-twentieth century. *J. Geophys. Res.: Atmos.* 118 (22), 12–587.
- Duhamel, P., Vetterli, M., 1990. Fast Fourier transforms: a tutorial review and a state of the art. *Signal Process.* 19 (4), 259–299.
- Fan, Y., van den Dool, H., 2004. Climate Prediction Center global monthly soil moisture data set at 0.5 degree resolution for 1948 to present. *J. Geophys. Res.* 109, D10102. <http://dx.doi.org/10.1029/2003JD004345>.
- Fleming, S.W., Lavenue, A.M., Aly, A.H., Adams, A., 2002. Practical applications of spectral analysis to hydrologic time series. *Hydrol. Process.* 16 (2), 565–574.
- Huang, J., van den Dool, H.M., Georgarakos, K.P., 1996. Analysis of model-calculated soil moisture over the United States (1931–1993) and applications to long-range temperature forecasts. *J. Clim.* 9, 1350–1362.
- India-WRIS version 4, 2014. Non Classified Hydro Observation data at Jondhra Station, India, Available on <<http://india-wris.nrsdc.gov.in/HydroObservationStationApp.html>> (accessed in November, 2014).
- Jiang, C., Xiong, L., Wang, D., Liu, P., Guo, S., Xu, C.-Y., 2015. Separating the impacts of climate change and human activities on runoff using the Budyko-type equations with time-varying parameters. *J. Hydrol.* 522, 326–338. <http://dx.doi.org/10.1016/j.jhydrol.2014.12.060>.
- Kang, S., Lin, H., 2007. Wavelet analysis of hydrological and water quality signals in an agricultural watershed. *J. Hydrol.* 338 (1), 1–14. <http://dx.doi.org/10.1016/j.jhydrol.2007.01.047>.
- Karthikeyan, L., Nagesh Kumar, D., 2013. Predictability of nonstationary time series using wavelet and EMD based ARMA models. *J. Hydrol.* 502, 103–119. <http://dx.doi.org/10.1016/j.jhydrol.2013.08.030>.
- Keyantash, J., Dracup, J.A., 2002. The quantification of drought: an evaluation of drought indices. *Bull. Amer. Meteor. Soc.* 83 (8), 1167–1180.
- Kim, T.W., Valdés, J.B., 2003. Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. *J. Hydrol. Eng.* 8 (6), 319–328.
- Krause, P., Boyle, D.P., Bäse, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. *Adv. Geosci.* 5, 89–97.
- Labat, D., Ababou, R., Mangin, A., 2000. Rainfall–runoff relations for karstic springs. Part II: continuous wavelet and discrete orthogonal multiresolution analyses. *J. Hydrol.* 238 (3), 149–178.
- Mabrouk, A.B., Abdallah, N.B., Dhifaoui, Z., 2008. Wavelet decomposition and autoregressive model for time series prediction. *Appl. Math. Comput.* 199, 334–340. <http://dx.doi.org/10.1016/j.amc.2007.09.067>.
- Madhu, S., Kumar, T.L., Barbosa, H., Rao, K.K., Bhaskar, V.V., 2015. Trend analysis of evapotranspiration and its response to droughts over India. *Theoret. Appl. Climatol.*, 41–51.
- Mallat, S.G., Zhang, Z., 1993. Matching pursuits with time–frequency dictionaries. *IEEE Trans. Signal Process.* 41 (12), 3397–3415.
- Mallat, S., 1999. *A Wavelet Tour of Signal Processing*. Academic Press, San Diego, CA.
- Massmann, C., Wagener, T., Holzmann, H., 2014. A new approach to visualizing time-varying sensitivity indices for environmental model diagnostics across evaluation time-scales. *Environ. Model. Softw.* 51, 190–194. <http://dx.doi.org/10.1016/j.envsoft.2013.09.033>.
- Maybank, J., Bonsai, B., Jones, K., Lawford, R., O'brien, E.G., Ripley, E.A., Wheaton, E., 1995. Drought as a natural disaster. *Atmos.–Ocean* 33 (2), 195–222.
- McKee, T.B., Doesken, N.J., Kleist, J., 1993. The relationship of drought frequency and duration to time scales. *Proceedings of the 8th Conference on Applied Climatology*, vol. 17(22). American Meteorological Society, Boston, MA, pp. 179–183.
- Mo, K.C., 2008. Model-based drought indices over the United States. *J. Hydrometeorol.* 9 (6), 1212–1230.
- Ozger, M., Mishra, A.K., Singh, V.P., 2011. Estimating Palmer Drought Severity Index using a wavelet fuzzy logic model based on meteorological variables. *Int. J. Climatol.* 31, 2021–2032. <http://dx.doi.org/10.1002/joc.2215>.
- Özger, M., Mishra, A.K., Singh, V.P., 2012. Long lead time drought forecasting using a wavelet and fuzzy logic combination model: a case study in Texas. *J. Hydrometeorol.* 13, 284–297. <http://dx.doi.org/10.1175/JHM-D-10-05007.1>.
- Pan, F., Smerdon, J.E., Stieglitz, M., Webster, P., 2005. A wavelet approach to reconstructing near-surface temperature time series from observations of subsurface temperatures at several meters depth. In: AGU Fall Meeting Abstracts.
- Peters, E., 2003. Propagation of drought through groundwater systems: illustrated in the Pang (UK) and Upper-Guadiana (ES) catchments, Wageningen Universiteit.
- Rajeevan, M., Bhat, J., 2008. A high resolution daily gridded rainfall Data Set (1971–2005) for Mesoscale Meteorological Studies, NCC research report, No. 9/2008. Available at <http://www.imdpune.gov.in/ncc_rept/RESEARCH%20REPORT%209.pdf> (accessed in November, 2014).
- Sang, Y., Singh, V., Sun, F., Chen, Y., Liu, Y., Yang, M., 2016. Wavelet-based hydrological time series forecasting. *J. Hydrol. Eng.*, 06016001 [http://dx.doi.org/10.1061/\(ASCE\)HE.1943-5584.0001347](http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0001347).
- Shannon, C.E., 1949. Communication in the presence of noise. *Proc. IRE* 37 (1), 10–21.
- Shukla, S., Steinemann, A.C., Lettenmaier, D.P., 2011. Drought monitoring for Washington State: indicators and applications. *J. Hydrometeorol.* 12 (1), 66–83.
- Smith, L.C., Turcotte, D.L., Isacks, B.L., 1998. Stream flow characterization and feature detection using a discrete wavelet transform. *Hydrol. Process.* 12 (2), 233–249.
- Soltani, S., 2002. On the use of the wavelet decomposition for time series prediction. *Neurocomputing* 48 (1), 267–277.
- Ujeneza, E.L., Abiodun, B.J., 2014. Drought regimes in Southern Africa and how well GCMs simulate them. *Clim. Dyn.* 44, 1595–1609. <http://dx.doi.org/10.1007/s00382-014-2325-z>.
- van den Dool, H., 2003. Performance and analysis of the constructed analogue method applied to U.S. soil moisture over 1981–2001. *J. Geophys. Res.* 108. <http://dx.doi.org/10.1029/2002JD003114>.
- Wang, H., Chen, Y., Pan, Y., Li, W., 2015. Spatial and temporal variability of drought in the arid region of China and its relationships to teleconnection indices. *J. Hydrol.* 523, 283–296. <http://dx.doi.org/10.1016/j.jhydrol.2015.01.055>.
- Westra, S., Sharma, A., 2006. Dominant modes of interannual variability in Australian rainfall analyzed using wavelets. *J. Geophys. Res.* 111 (D5). <http://dx.doi.org/10.1029/2005JD005996>.
- Willmott, C.J., Robeson, S.M., Matsuura, K., 2012. A refined index of model performance. *Int. J. Climatol.* 32 (13), 2088–2094.