Predictability of Hydrological Systems Using the Wavelet Transformation: Application to Drought Prediction

Rajib Maity and Mayank Suman

1 Introduction

Hydrological processes are complex and associated with multiple hydroclimatic factors. As a consequence, the hydrologic time series are continuously evolving over time and exhibit nonstationary nature [1, 2]. Generally, the hydrologic time series is presented in the time domain and this representation is useful when the temporal changes in different statistical properties are attempted. However, this representation is not adequate in some cases as it hides important information about frequency content of the time series and its temporal evolution (if any). Information on constituting frequency of a time series may be extracted using mathematical transforms like Fourier Transform, Wavelet Transform, etc.

Fourier Transform (FT) is a mathematical tool that is used to separate frequency component of time series. The basis function used in the FT is circular functions (sine and cosine functions). FT is based on the fact that any continuous periodic time series can be constructed by using adequate number of appropriate sine or cosine waves. FT transforms a time series from the time–amplitude domain to frequency–amplitude domain. FT has been used in hydrology by many investigators. For instance, FT was used by Kirchner et al. [3] for studying contaminant transport in catchment. Şen [4] studied the FT of periodic-stochastic hydrologic sequences in general.

The outcome of FT can point out the frequencies of sine or cosine waves in the given time series, however, it cannot provide the information about the temporal evolution of amplitude of these frequencies. Rather, it provides the mean amplitude or power of the different frequencies present in the time series. This drawback can

R. Maity (✉)
Department of Civil Engineering, Indian Institute of Technology, Kharagpur, India
e-mail: rajib@civil.iitkgp.ac.in; rajibmaity@gmail.com

M. Suman
School of Water Resources, Indian Institute of Technology, Kharagpur, India
e-mail: mayanksuman@protonmail.com

S. K. Singh and C. T. Dhanya (eds.), *Hydrology in a Changing World*, Springer Water,
https://doi.org/10.1007/978-3-030-02197-9_5

be partially overcome by using short-term Fourier transform (also called Windowed Fourier Transform), in which the transformation instead of operating on the whole of time series at once operates on some selected length of the time series called a window. However, this approach can only be applied when one is confident about window size. If the window size changes too often for a time series, this methodology does not yield satisfactory results. Hence, the FT is best suited for stationary time series.

The Wavelet Transform (WT) is another mathematical tool extensively used for analysis of time series in hydrology. Unlike FT, WT helps in getting temporal information about different frequencies in the time series also, which may prove useful while analyzing nonstationary time series. WT is being widely used for hydrological time series prediction [5–7]. Smith et al. [8] used WT for streamflow prediction. Özger et al. [9] and Maity et al. [10] utilized WT for drought forecasting and its evolution. Labat et al. [11] modeled the rainfall–runoff relation using WT. These studies highlight the appropriateness and effectiveness of WT based methodologies to model relationship between hydrological series.

This chapter aims at exploring the potential of wavelet transform for prediction of hydrological systems. In this regard, the mathematical framework of wavelet transform and multi-resolution analysis using wavelet functions are discussed in the subsequent sections. An example problem of predicting drought using multi-resolution wavelet is also provided for showing the effectiveness of wavelet transform for hydrologic prediction.

2 Wavelet Function

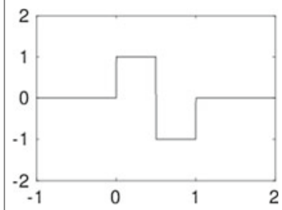
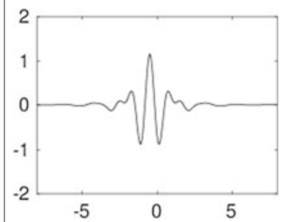
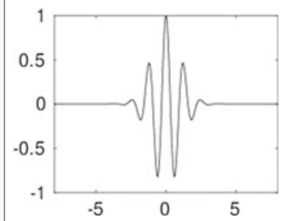
Wavelet is a finite disturbance of zero mean amplitude. Wavelet function has unit energy and its integration over the real number line is zero. Details of a few well-known wavelet functions like Haar, Morlet, etc., are shown in Table 1. Wavelet functions are localized in both time and frequency space. Many different wavelet functions can be derived from one wavelet function by shifting it temporally and/or scaling, without changing any functional form [12]. The original wavelet function is called mother wavelet and all the derived wavelet functions are called daughter wavelets.

For a mother wavelet function $\Psi(t)$, the daughter wavelet functions (denoted by $\Psi_{a,b}(t)$) can be obtained as

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}}\Psi\left(\frac{t-b}{a}\right) \tag{1}$$

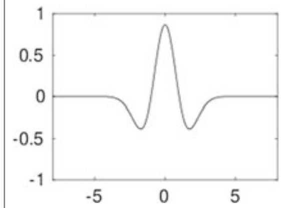
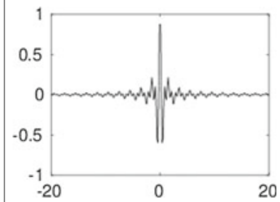
where a, b and t are scaling parameter, shifting parameter, and time step, respectively. The scaling parameter helps in varying the frequency of mother wavelet as it is inversely related to wave frequency and the shifting parameter helps in shifting the mother wavelet with respect to time. Scaling as a mathematical operation

Table 1 Details of some mother wavelets

Name	Mother wavelet function	Graphical representation
Haar or Daubechies 1	$\Psi(t) = \begin{cases} 1 & 0 \le t \le 0.5 \\ -1 & 0.5 \le t \le 1 \\ 0 & \text{otherwise} \end{cases}$	
Meyer	In frequency domain $\Psi(\omega) = \begin{cases} \frac{1}{\sqrt{2\pi}} \sin\left(\frac{\pi}{2} \nu\left(\frac{3\vert\omega\vert}{2\pi} - 1\right)\right) e^{j\omega/2} & \text{if } \frac{2\pi}{3} < \vert\omega\vert < \frac{4\omega}{3} \\ \frac{1}{\sqrt{2\pi}} \sin\left(\frac{\pi}{2} \nu\left(\frac{3\vert\omega\vert}{2\pi} - 1\right)\right) e^{j\omega/2} & \text{if } \frac{4\pi}{3} < \vert\omega\vert < \frac{8\omega}{3} \\ 0 & otherwise \end{cases}$ where, $\nu(x) = \begin{cases} 0 & x \le 0 \\ x & 0 < x < 1 \\ 1 & x \ge 1 \end{cases}$	
Morlet	$\Psi(t) = c_\sigma \pi^{(-1/4)} e^{(-1/2)t^2} \left(e^{i\sigma t} - k_\sigma\right)$ where, $k_\sigma = e^{-1/2\sigma^2}$ and $c_\sigma = \left(1 + e^{-\sigma^2} - 2e^{-3/4\sigma^2}\right)^{-1/2}$	

(continued)

Table 1 (continued)

Name	Mother wavelet function	Graphical representation
Ricker or Mexican Hat	$\Psi(t) = \frac{2}{\sqrt{3\sigma}\pi^{1/4}}\left(1 - \left(\frac{t}{\sigma}\right)^2\right)e^{\frac{-t^2}{2\sigma^2}}$	
Complex Shannon 1-1	$\Psi(t) = \sqrt{F_b} sinc(F_b x) e^{(2i\pi\ F_c x)}$ The wavelet is named as $F_b - F_c$ For Figure $F_b = F_c = 1$	

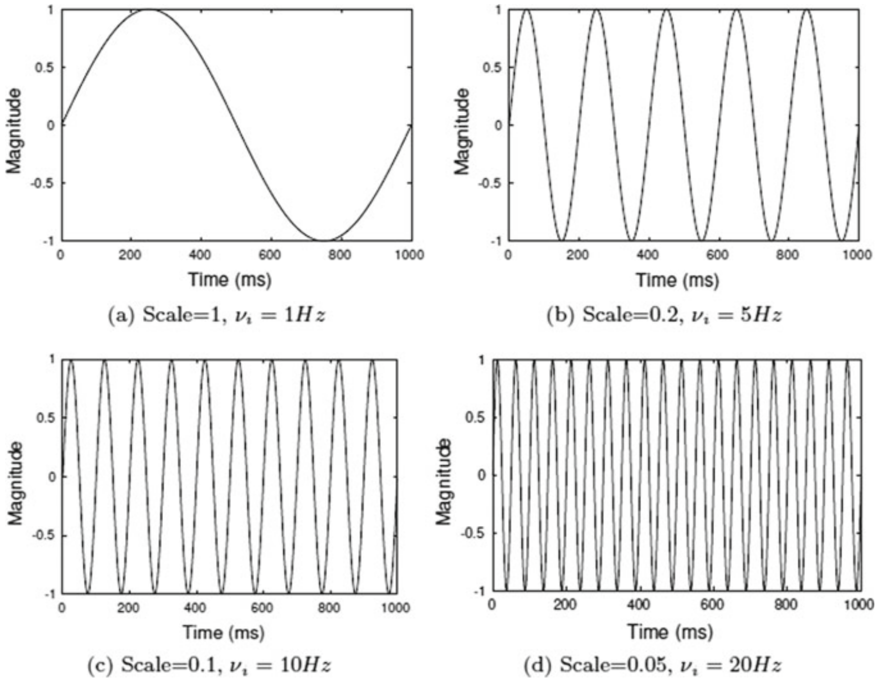


Fig. 1 Different scale/frequencies of unit amplitude sign wave (ν_i represents frequency)

either dilates or compresses a wavelet function, i.e., larger scales correspond to the dilated (or stretched out) daughter wavelet function (compared to mother wavelet) and smaller scales correspond to the compressed daughter wavelet function. For instance, in Fig. 1, different scales of sine wave with unit amplitude are shown. It can be observed from Fig. 1a and d that decrease in scale leads to contraction in signal and *vice versa*.

The importance of shifting and scaling operation on mother wavelet for wavelet transform is discussed in Sect. 3. It should be further noticed that, with the increase in scaling parameter, the frequency of the derived daughter wavelet decreases. Hence, with finite scaling factor the mother wavelet and all daughter wavelet cannot cover lower frequency range (in that case, scaling factor can become too high as scaling factor is inversely proportional to frequency). Another function called scaling function or Father wavelet function (denoted by $\phi(t)$) is used for covering the whole frequency range of the time series during discrete wavelet transform (discussed later). Father wavelet functions, like mother wavelet functions, are of finite duration and act as low-pass filter. In the next section, Haar wavelet is dealt in greater depth.

Haar Wavelet

Haar wavelet, proposed by Alfréd Haar in 1909, is a square-shaped wavelet, which is also the first member of the Daubechies class of wavelets and regarded as daubechies1

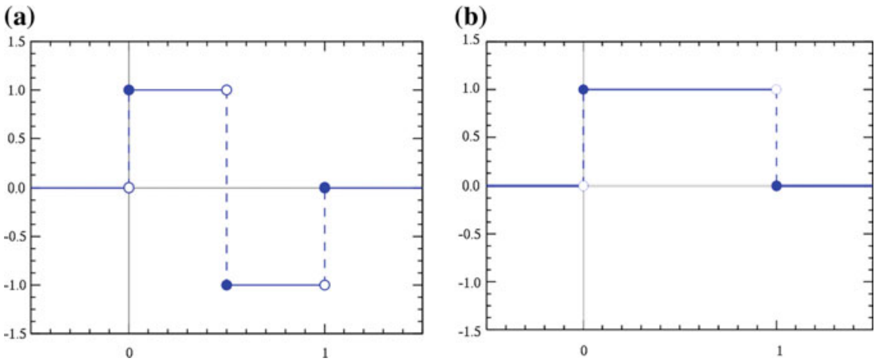


Fig. 2 Haar Wavelet function **a** Mother Wavelet **b** Scaling function

or db1. As per Maheswaran and Khosa [13], this wavelet function has better time localization capability, so, it is useful for short-term predictions. Haar wavelet (Fig. 2a) is defined as

$$H(t) = \begin{cases} 1 \; 0 \le t < 0.5 \\ -1 \; 0.5 \le t < 1 \\ 0 \; otherwise \end{cases} \tag{2}$$

The scaling function for Haar wavelet (Fig. 2b) is given by:

$$S(t) = \begin{cases} 1 \; 0 \le t < 1 \\ 0 \; otherwise \end{cases} \tag{3}$$

Haar wavelet and scaling functions are having the following properties:

(i) The Haar wavelet and its scaling function can be expressed as linear combination of scaling function of different scales.

$$S(t) = S(2t) + S(2t - 1) \tag{4}$$

$$H(t) = S(2t) - S(2t - 1) \tag{5}$$

(ii) Any continuous real function on [0, 1] can be approximated by linear combinations of dyadic Haar wavelet with different scales and shifts $(1, H(t + b_1), H(2t + b_2), H(4t + b_3), \ldots, H(2^n t + b_n), \ldots)$.

(iii) Similarly, any continuous real function with compact support can be approximated by a linear combination of scale functions with different scales and shifts $(S(t + b_1), S(2t + b_2), S(4t + b_3), \ldots, S(2^n t + b_n), \ldots)$.

3 Wavelet Transform

Wavelet transform aims to provide the state of different frequency/frequency band in the time series with time. For this purpose, the WT uses a family of daughter wavelets (Eq. 1) for transformation. Both the operations of shifting and scaling used during the derivation of different daughter wavelets have their significance with respect to wavelet transform. Shifting of wavelet function helps in capturing the state of different frequencies along the time. Scaling operation on the other hand changes the frequency of the mother wavelet function (Eq. 1). The scaling parameter is similar to the scale used in maps, i.e., high scale (thus low frequency) corresponds to non-detailed global view (of the time series), and low scale (high frequency) corresponds to detailed view. Scaling is required to capture the information regarding different frequency ranges in the time series as per the uncertainty principle of signal analysis which states

$$\Delta t \Delta\omega \geq \frac{1}{2} \tag{6}$$

where Δt represents time step and $\Delta\omega$ represents resolution in angular frequency ($\omega = 2\pi\nu$, where ν is wave frequency). Hence, the larger is time resolution chosen for the analysis, the smaller will be frequency resolution analyzed or *vice versa*. Hence, to analyze the time series at different frequency resolutions, the scaling of mother wavelet is required.

In a nutshell, WT transforms the time series into its constituents or components based on shifting and dilation or scaling of the mother wavelet $\Psi(t)$. During WT, the time series is convoluted with mother wavelet of different scales and shifts to obtain the wavelet components. It should be noted that despite having finite length, scaling and shifting of mother wavelet enable it to catch most of intermittent disturbances of different durations. By using daughter wavelet of higher scale, WT extracts the slow moving changes or global information in time series and by using daughter wavelet of lower scale, WT extracts the detailed information about local disturbances. This enables the wavelet transform to provide the time and frequency information or time–frequency representation of the time series, unlike, Fourier Transform. Fourier Transform loses the time information during transformation because it uses sinusoidal wave, a function with infinite support as basis function.

Based on the selection of scaling and shifting parameters and mode of application of wavelet transform, the wavelet transform can be of different types. Three of the most widely used wavelet transforms are as follows:

- Continuous Wavelet Transform (CWT)
- Discrete Wavelet Transform (DWT)
- Stationary Wavelet Transform (SWT)

These transforms are discussed in the following subsections.

3.1 Continuous Wavelet Transform (CWT)

If shifting and scaling factors are considered to be continuous over real number line while applying wavelet transform, the WT is called continuous wavelet transform (CWT). The CWT is computed by changing the scale of the analysis window, shifting the window in time, multiplying by the time series, and integrating over all times. In CWT, the transform is mathematically expressed as

$$W_f(a,b) = \frac{1}{\sqrt{C_\Psi}} \int X(t)\, \Psi^*_{a,\,b}(t)dt \tag{7}$$

where $\Psi^*(t)$ denotes complex conjugate of $\Psi(t)$, $C_\Psi = 2\pi \int \left|\hat{\Psi}(\omega)\right|^2 / \omega d\omega$ and $\hat{\Psi}(\omega)$ denotes the Fourier transform of $\Psi(t)$ given by

$$\hat{\Psi}(\omega) = \frac{1}{\sqrt{2\pi}} \int e^{i\omega t} \Psi(t)dt \tag{8}$$

If the mother wavelet ($\Psi(t)$) is orthogonal, then the inverse of wavelet transformation is given by

$$X(t) = \frac{1}{\sqrt{C_\Psi}} \iint \frac{W_f(a,b)\, \Psi_{(a,b)}(t)}{a^2} dadb \tag{9}$$

3.2 Discrete Wavelet Transform (DWT)

Discrete class of wavelets is formed when shifting and scaling parameters are considered discrete instead of continuous variables while applying wavelet transform. If the discrete wavelet is sampled over dyadic space-time grid, the resulting wavelets are called dyadic discrete wavelet [14]. The dyadic daughter wavelets are denoted by

$$\Psi_{j,b}(t) = \frac{1}{\sqrt{2^j}} \Psi\left(\frac{t}{2^j} - b\right) \tag{10}$$

The wavelet transform is given by

$$W_f(a,b) = \frac{1}{\sqrt{C_\Psi}} \sum X(t)\, \Psi^*_{a,b}(t) \tag{11}$$

where $\Psi^*(t)$ denotes complex conjugate. C_Ψ is as defined before. Discrete wavelet component is down-sampled or subband coded according to Nyquist–Shannon theorem [15]. The Nyquist–Shannon sampling theorem is a fundamental connection between continuous and discrete representation of time series or signal. This theorem is applicable to any signal having finite range of frequencies or in other words, signal having zero Fourier transform coefficient outside some finite range of frequencies. According to this theorem, if any signal is sampled two times, first with a sampling rate of N_1 at scale a_1, second at a sampling rate of N_2 at scale a_2, then the information contained in these two sampling procedures is equivalent, given

$$N_2 = \frac{a_1}{a_2} N_1 \tag{12}$$

As the frequency range of wavelet components (generated by Eq. 11) is decreased by half, hence, the components can therefore be subsampled by 2, by discarding every alternate sample or sample falling at even places from the beginning. As a result, each of the components has half the length that original time series or signal had. Hence, DWT halves the time resolution but doubles the frequency resolution. Since, the frequency band of the time series now spans only half the previous frequency band; it effectively reduces the uncertainty in the frequency by half. This procedure is also known as subband coding (or down-sampling). Subband coding, however, results in wavelet coefficients depending on their location. As a result, a small change in input signal causes large changes in wavelet coefficients. This is termed as translation-invariance of DWT and is considered a major drawback which limits its application in signal analysis [16].

It should be noted that a discrete mother wavelet acts as a band-pass filter and scaling it for each level (for dyadic space) effectively halves its bandwidth. This creates the problem that in order to cover the entire spectrum (till the frequency limiting to zero), an infinite number of scaling is required. Hence, to cover the complete spectrum another function associated with the mother wavelet, Father Wavelet is used. Further, dyadic wavelet functions are orthogonal so the inverse of wavelet transform is given by

$$X(t) = \frac{1}{\sqrt{C_\Psi}} \sum_{j,\,k \in Z} X(t)\,\Psi_{a,b}(t) \tag{13}$$

Alternatively, DWT can also be carried out by using a pair of filters—a high-pass and a low-pass filter. In DWT, the component obtained after convolution of signal with low-pass filter followed by dyadic down-sampling is called approximate component and one obtained by using high-pass filter and dyadic down-sampling is called detailed component. Low-pass filter is derived from scaling function and high-pass filter is derived from mother wavelet function. The DWT filters for Haar mother wavelet (discussed in Sect. 2) are given by

$$h_{r,c} = \begin{cases} 1/\sqrt{2} & c \in \{r, (r+1)\, mod\, n\} \\ 0 & otherwise \end{cases} \tag{14}$$

$$g_{r,c} = \begin{cases} (-1)^{r-c}/\sqrt{2} & c \in \{r, (r+1)\ mod\, n\} \\ 0 & otherwise \end{cases} \tag{15}$$

where $h_{r,c}$ and $g_{r,c}$ are the elements of matrix H and G respectively, r and c represent the row and column of filter matrix. H and G are low-pass and high-pass filter matrix, respectively. Here, "mod" represents a module function. $k\, mod\, n = n$ if $k = n$, otherwise $k\, mod\, n =$ remainder of k divided by n. On closer observation, the low-pass filter is 2 term moving average operation and the high-pass filter is first-order differencing operation normalized with a factor of $1 \big/ \sqrt{2}$. When the time series is multiplied with these filters followed by dyadic down-sampling (ignoring every other value), two components are obtained. The component obtained after multiplication with high-pass filter is called detailed DWT component (denoted by d) and component obtained after multiplication with low-pass filter is termed approximate DWT component (denote by a).

3.3 Stationary Wavelet Transform (SWT)

Stationary Wavelet Transform (SWT) is specially designed to avoid the translation-invariance of DWT. SWT components are not down-sampled (as per Nyquist–Shannon sampling theorem) and the filter coefficients are up-sampled by a factor of $2^{(j-1)}$ in the jth level of algorithm. Hence, the SWT unlike DWT does not change the time resolution at any stage. But lack of subband coding results in redundancies in components as SWT components have twice the number of elements needed as per Nyquist–Shannon Theorem. However, SWT reduces the complexity of signal analysis as both input signal and its components have equal length. For obtaining Haar SWT components, time series can be multiplied with the filters given by Eqs. 14 and 15 without dyadic down-sampling.

4 Multi-resolution Analysis

Multi-Resolution Analysis (MRA) provides the detailed and approximate components at even lower levels by using low-pass filter component (approximate component) from higher level as input to wavelet transform at each subsequent level. Each application of WT reduces the frequency band of component into half and it helps in getting slow and fast dynamic component at different levels, which may enhance the accuracy of prediction. The MRA is named on the basis of the wavelet transform algorithm being used repeatedly, like Multi-Resolution Discrete Wavelet Transform (MRDWT) or Multi-Resolution Stationary Wavelet Transform (MRSWT). Irrespective of wavelet transformation used, after application of MRA a time series $X(t)$ is represented as

$$X(t) = \sum_{k} a_{0,k}\varphi_{0,k}(t) + \sum_{j=0}^{\infty}\sum_{k} d_{j,k}\Psi_{j,k}(t) \tag{16}$$

where $\varphi_{0,k}(t)$ and $\Psi_{j,k}(t)$ represent scaling function and mother wavelet function, respectively. The subscript pair j and k represent scale and shift parameters of mother wavelet or scaling function. The approximate component ($a_{0,k}$) and detailed component ($d_{j,k}$) are expressed as

$$a_{0,k} = \sum X(t)\,\varphi_{0,k}(t-k) \tag{17}$$

$$d_{j,k} = \sum X(t)\,2^{-j}\Psi_{j,k}\left(2^{-j}t-k\right) \tag{18}$$

If maximum level of decomposition is L, $a_{0,k}$ series is also represented as a_L. Similarly, $d_{j,k}$ series are also represented as d_j, where $j \in 1, 2, \ldots, L$. In form of filters, the components a_L and d_j are expressed as

$$a_L = G_L G_{L-1} \ldots G_1 X \tag{19}$$

$$d_j = H_j G_{j-1} G_{j-2} \ldots G_1 X = H_j a_{j-1} \text{ for } j \in \{1, 2, \ldots, L\} \tag{20}$$

The low- and high-pass filters for Haar mother wavelet at any level l are given by

$$h_{l,r,c} = \begin{cases} 1/\sqrt{2} & c \in \{r, \left(r + 2^{(l-1)}\right) mod\, n\} \\ 0 & otherwise. \end{cases} \tag{21}$$

$$g_{l,r,c} = \begin{cases} (-1)^{r-c}/\sqrt{2} & c \in \{r, \left(r + 2^{(l-1)}\right) mod\, n\} \\ 0 & otherwise. \end{cases} \tag{22}$$

where $h_{l,r,c} \in H_l$, $g_{l,r,c} \in G_l$, H_l and G_l are low-pass and high-pass filter at level l. r and c represent row and column, respectively. It should be noted that for $l = 1$ the above equations are same as Eqs. 14 and 15.

5 Illustrative Example on Drought Prediction

Drought is a hydrological extreme of prolonged water deficit. It is slow initiating but long lasting phenomenon leading to huge economic losses. As per the American Meteorological Society [17], 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. This illustrative example is on the drought prediction over one small and another medium size watersheds from central part of India. The methodology and the results are mostly borrowed from Maity et al. [10] and Suman and Maity [18].

Since the hydrologic cycle is a continuous transport of water, the occurrence of meteorological drought is expected to propagate to other types of droughts [19]. Hence, it can be hypothesized that prolonged period of meteorological drought along with high evaporation loss may lead to soil moisture deficit, resulting in agricultural drought. Further, in the same way, intense agricultural drought may turn into hydrological drought given the long duration. This precedence or temporal consequences of different types of drought are easy to speculate, but it is difficult to model as a number of factors (of climatological, topographical and geographical characteristics) affect this precedence order. If basin size is relatively large, the lag in transition of drought is also expected [20]. Study of temporal transition of drought also has added advantage—the measurement of precipitation is more accurate and economical compared to measurement of soil moisture and streamflow (which may require specialized structure and may not be economical for large streams/catchment), hence, with the information of temporal transition of different types of drought, the drought prediction will be economical. Further, it may also lead to better drought preparedness and thus better mitigation strategy for the community. The following subsection briefly discusses the methodology of study followed by subsection for details of study area and results.

5.1 Methodology

Overall methodology is broadly divided into two modules—(i) Drought characterization using drought indices and generation of its time series. Further, the study of lagged correlation between the drought indices to check whether there is any delayed response of one drought index exists on the other, (ii) Formulation of different models considering the lagged information of predictor drought index, based on MRSWT components of drought indices. As stated above, the selected wavelet function is Haar wavelet, as this wavelet function is having better time localization capability, which renders it good for short lead period prediction. Further, the most potential model structure/type is selected for prediction. It should be noted that selected model structure may differ for different basins. The methodological overview is shown in Fig. 3. Details of these modules are presented in the following subsections.

5.1.1 Drought Characterization Through Standardized Indices

For drought characterization, many different drought indices are available in the literature like Palmer Drought Severity Index (PDSI), Keetch–Byram drought index (KBDI), Standardized Precipitation Index (SPI), etc. However, no single drought index is considered universal, rather, their suitability depends on its application for a particular problem [21]. For analyzing the interrelation of different kinds of drought

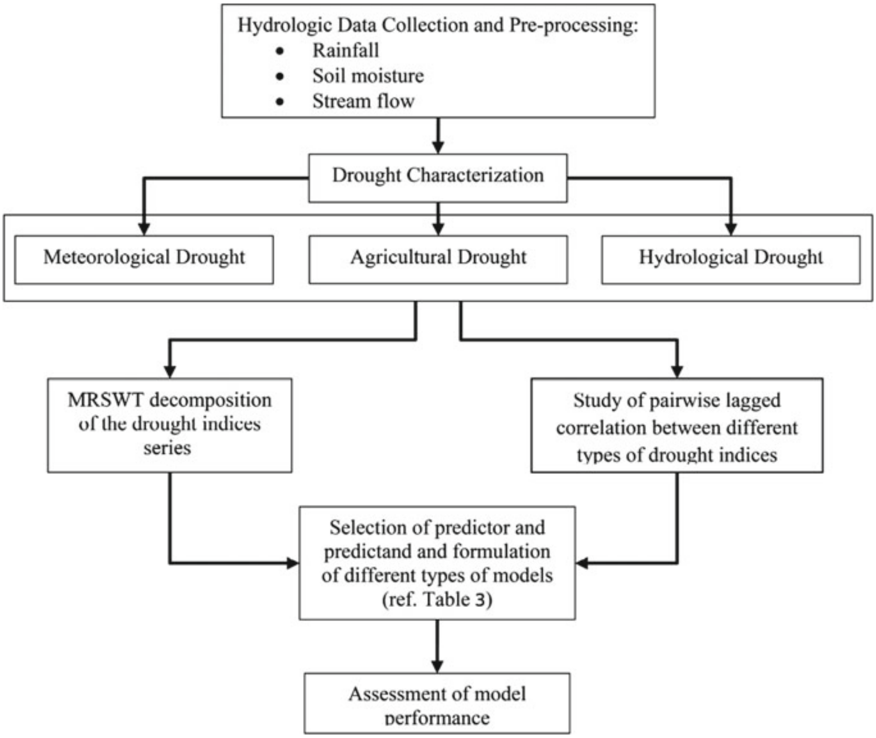


Fig. 3 Methodological overview (*Source* Maity et al. [10])

such as meteorological, agricultural, and hydrological droughts, a mathematical consistent drought index is needed for each of this drought type. Keeping this in mind, Standardized Precipitation Index (SPI), Standardized Soil Moisture Index (SSMI), and Standardized Stream Flow Index (SSFI) are used for characterization of meteorological, agricultural, and hydrological droughts, respectively. SPI, SSMI, and SSFI are 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. [22] for the Fort Collins, Colorado river basin in the USA. SPI can be defined as standard normal variate of precipitation with respect to the standard deviation of precipitation for a given location and time period calculated from the historical precipitation data. SSMI and SSFI have similar conceptualization.

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

(i) Time series of concerned variable is either accumulated or moving averaged for the desired averaging temporal scale.

(ii) A suitable Probability Density Function (pdf) is fitted (Gamma distribution in this example) and corresponding Cumulative Distribution Function (CDF) is obtained.
(iii) Using the fitted CDF, reduced variate of the concerned variable is computed.
(iv) The reduced variate is transformed to a standard normal variate (mean $= 0$ and standard deviation $= 1$) to obtain the desired standardized index.

All these indices can have both positive and negative values, positive value showing a surplus and negative value showing a deficit. Prolonged and severe period of deficit may indicate a drought.

Depending on the characteristics of the study basin, sometime lag may be expected before effect of predecessor drought situation is observed over a successor one. The time lag may also originate due to nature of variable being studied. On the basis of expected precedence order, different predictor and predictant drought index relationships are considered. For instance, SPI is taken as predictor for SSMI and SSFI; SSMI is considered a predictor for SSFI. To quantify the time lag in drought propagation, lagged correlations between different predictand–predictor drought indices are studied. The lag with highest correlation is considered as the measure of delay in response that predictor drought series has on the predictand drought series.

5.1.2 Modeling of Drought Indices Interrelation

The drought indices are decomposed into components using MRSWT up to level 2. The mathematical details of MRSWT are presented in Sect. 4. By using MRSWT, the prediction of drought indices leads to the problem of predicting the slow and fast dynamic components separately. This approach may be advantageous, 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 [23]. The decomposition through MRSWT results in three components (d_1, d_2, and a_2) for each of the drought indices.

The modeling of interrelation between the drought indices components may facilitate the prediction of successor drought from the state of predecessor one. Many approaches such as traditional (Multiple Linear Regression (MLR), Auto-Regressive Integrated Moving Average model with exogenous inputs (ARIMAX)) or even soft computing approaches (Artificial Neutral Network (ANN)) can be used for modeling. In this example, models are formulated in two versions (keeping input and output variables same)—one using feed-forward ANN with single hidden layer and other using MLR. 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. It should also be noted that minimum lead period also depends upon the level of decomposition being used to avoid the use of future information during the prediction. Since the maximum level for MRSWT is 2, minimum lead period for prediction is 2^2, i.e., 4.

5.1.3 Model Validation Scheme

All the proposed models, except those based on ANN approach, are tested using two different validation schemes—I and II. Details of these schemes can be found in Maity et al. [10] and also briefly explained hereafter. ANN-based models are validated with scheme I only. Details of these schemes are discussed below. These validation schemes are also illustrated in Fig. 4.

(i) Scheme I—Fixed Development and Testing Period: In this scheme, the whole data set is divided into development period and testing period. These periods remain stationary in one model calibration–prediction run. The parameters of the model are estimated during the development period. Complete testing period data set is predicted in the next model run. Hence, 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 the time series 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.

5.1.4 Model Performance Evaluation

Performances of different models are assessed based on four statistical measures, namely correlation coefficient (r), Refined Index of Agreement (D_r) and unbiased Root Mean Square Error ($uRMSE$). Expressions for r can be found elsewhere [24]. The expression of D_r is given by [25]

$$D_r = \begin{cases} 1 - D_{r\_frac} \; for \; D_{r\_frac} \le 1 \\ \frac{1}{D_{r\_frac} - 1} \; for \; D_{r\_frac} > 1 \end{cases} \tag{23a}$$

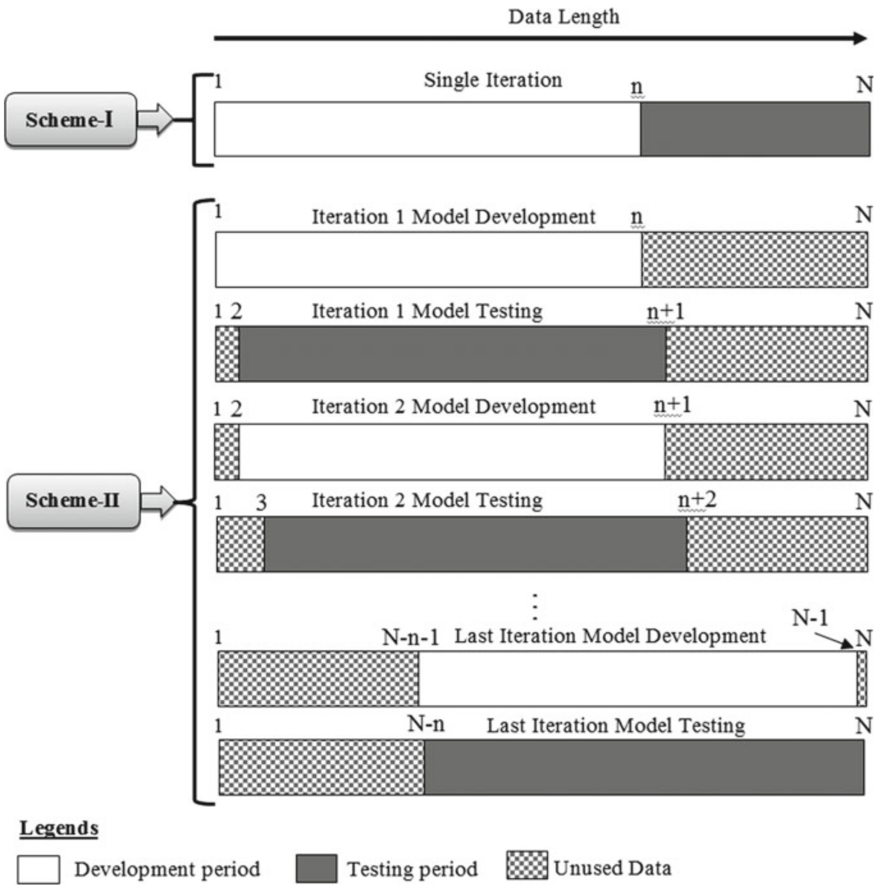


Fig. 4 Schematic diagram of two types of validation schemes [10]. In Scheme II, at any model testing iteration only the last value is recorded for performance assessment though the testing period overlaps the model development period of the same iteration

where D_{r_frac} is intermediate calculation step which is calculated as

$$D_{r\_frac} = \frac{\sum_{i=1}^{n} |Y_i - X_i|}{2 \sum_{i=1}^{n} \left| X_i - \bar{X} \right|} \tag{23b}$$

where X_i and Y_i are the i^{th} observed and predicted values, $\bar{X}$ is the mean of the observed values and n is the total number of observations.

The *uRMSE* is the RMSE calculated between the deviations of observed and predicted values from their respective means. It is expressed as

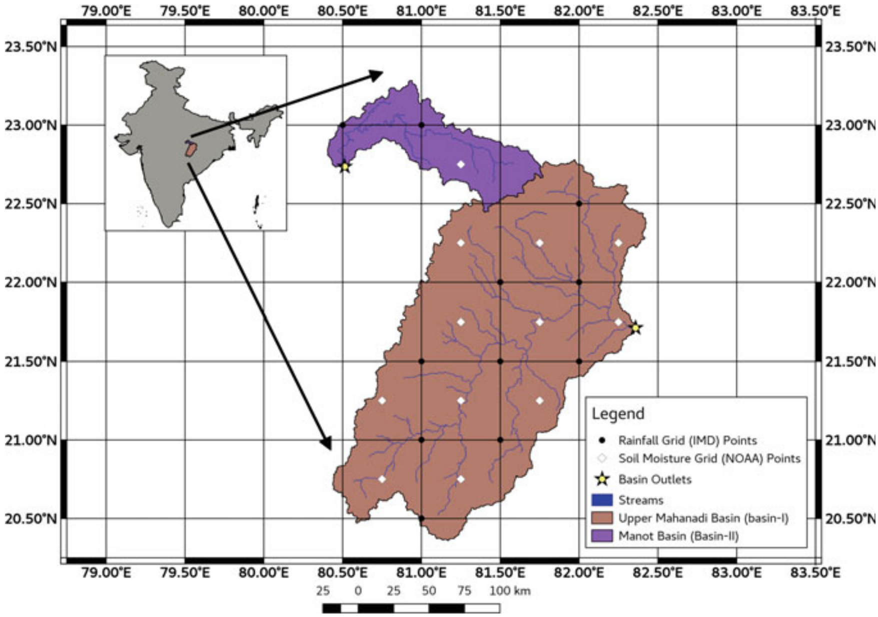


Fig. 5 Study basins—Upper Mahanadi Basin (Basin-I) and Upper Narmada Basin (Basin-II)

$$uRMSE = \sqrt{\frac{\sum_{i=1}^{n} \left\{ \left(X_i - \overline{X} \right) - \left(Y_i - \overline{Y} \right) \right\}^2}{n}} \tag{24}$$

where X_i, Y_i, $\bar{X}$ and n are as defined before, $\bar{Y}$ is the mean of the predicted values. The lower the value of uRMSE, the better the model performance. The uRMSE removes the mean bias between observed and predicted time series (unlike RMSE). Hence, uRMSE is better model performance measure (compared to RMSE) in the presence of mean bias [26].

5.2 Study Areas

Two different basins are selected—upper Mahanadi basin up to Jondra (henceforth, basin-I) and upper Narmada basin up to Manot (henceforth, basin-II). Basin-I is mostly located in the state of Chhattisgarh in India as shown in Fig. 5. The area of the basin is 29645 km^2 and it is approximately bounded by 20° N to 23° N latitude and 80.5° E to 82.5° E longitude. Basin-II is located in state of Madhya Pradesh in India as shown in Fig. 5. It has an area of 4667 km^2 and it is approximately bounded by 22.5° N to 23.5° N latitude and 80° E to 82° E longitude.

Daily rainfall data and monthly soil moisture data for the study basins are obtained for the period of 1971 to 2005 from the India Meteorological Department (IMD) [27] and Climate Prediction Center (CPC) of the National Oceanic and Atmospheric Administration (NOAA) [28], 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 respective study basin as shown in Fig. 5. Daily rainfall data at each grid point is converted to monthly rainfall depth by accumulating it over the month. Daily stream flow data at the outlet of the basins (Jondhra station for Basin-I and Manot station for Basin-II) are procured from the Water Resources Information System [29] in India. For basin-I streamflow record of June, 1979 to December, 2005 is available, so the study period for basin-I is considered as January 1980 to December 2005. However, for basin-II the streamflow record for June, 1978 to December, 2005 is available, hence, the study period is taken as January, 1979 to December, 2005 for this basin. The daily stream flow data is converted to monthly data.

5.3 Results and Discussions

Taking monthly rainfall depth, soil moisture time series, and stream flow series as input SPI, SSMI, and SSFI, respectively, for different basins are calculated using a mixed distribution—Gamma distribution for nonzero values with probability mass at zero. For monthly rainfall depth, accumulation over averaging timescale was done during SPI calculation but for all other variables moving average is calculated during index calculation. Notations of SPI-1, SSMI-1, and SSFI-1 are used for 1-month timescale. Similarly, SPI-3, SSMI-3, and SSFI-3 are used for 3-month timescale. SPI-3, SSMI-3, and SSFI-3 time series are shown in Fig. 6. From Fig. 6, 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. SPI is taken as predictor for SSMI and SSFI; SSMI is considered a predictor for SSFI. The relationships are deemed valid regardless of selected averaging period and basin. As stated earlier, January, 1980 to December, 2005 is chosen as study period for basin-I, so all drought index series are having 312 data points. First 160 data points are considered for the initial scrutiny and model development. For basin-II, the study period is selected as January, 1979 to December, 2005. Being a small basin, the response of one variable over the other is expected to be fast and more dynamic, hence, a longer development length of 204 is selected. It should be further noted that 10% of the data length after development length is used for validation in case of ANN-based models. 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 according to their predictand–predictor relationship and the results are tabulated in Table 2. From Table 2, the correlation coefficient and refined index of agreement are higher for 3-month timescale indices. It is due to higher average period used to calculate the indices, which lead to more smoothening. The coefficient of

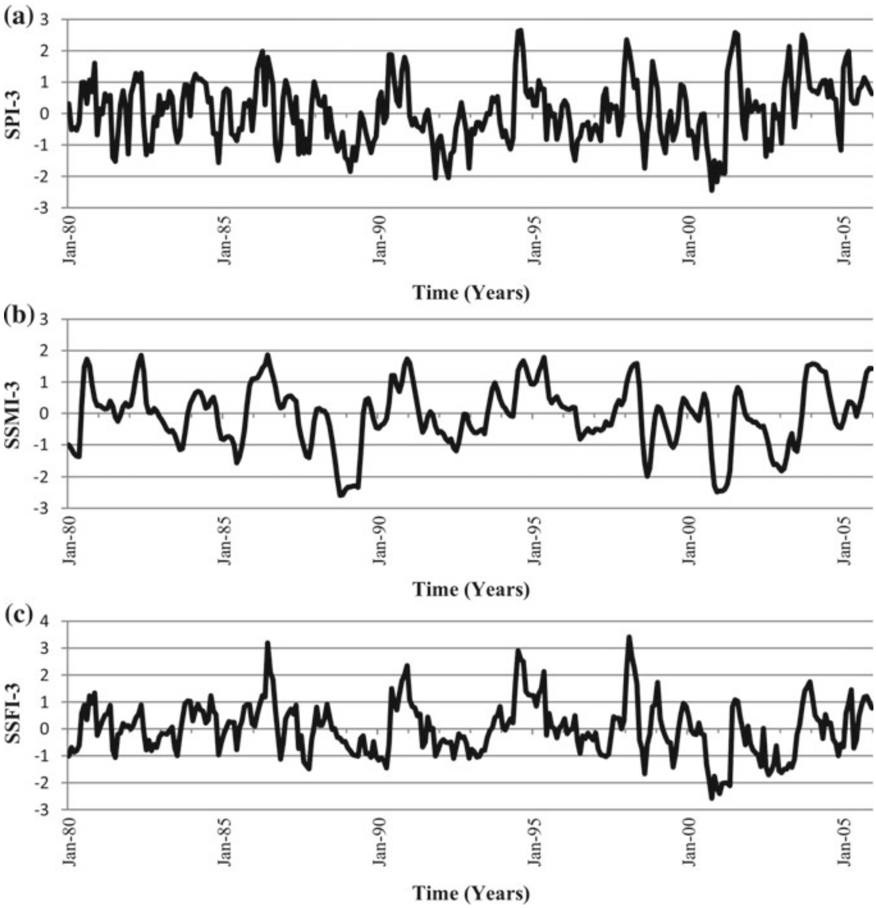


Fig. 6 Time series of **a** SPI-3 **b** SSMI-3 **c** SSFI-3 (January, 1980–December, 2005) for basin-I

correlation is found significant for all the cases, reaffirming the hypothesis that SSFI should be affected by both SPI and SSMI. 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 perdictors (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 2 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–predictant drought index pairs is then calculated for both basins. The results for basin-I are shown in Fig. 7. It

Table 2 Correlation coefficients (r) and the refined index of agreement (D_r) for different drought indices pairs during development period

Basin	Averaging period (in months)	Performance statistics	Predictand drought index	Predictor drought index	
				SPI	SSMI
Basin-I	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
Basin-II	1	r	SSMI	0.303	1.000
			SSFI	0.613	0.517
		D_r	SSMI	0.433	1.000
			SSFI	0.541	0.496
	3	r	SSMI	0.494	1.000
			SSFI	0.618	0.479
		D_r	SSMI	0.502	1.000
			SSFI	0.564	0.447

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–predictant pairs, the value of correlation coefficients decreases gradually with the further increase in lag. For basin-II, the correlation coefficient for zero lag is found to be the highest for all predictand–predictor relationships. The correlation coefficient is found to decrease gradually with increase in lag. However, the lag considered in modeling of interrelation of drought indices should be either equal or greater than the averaging period and minimum lead period requirement as discussed in Sect. 5.1.2. To reiterate, minimum required lead period for prediction is 4, since MRSWT with level 2 is used. In case of drought indices calculated using 3-month accumulation, SPI-3 with lag 4 and/or 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.

Five different models as shown in Table 3 are framed. Models 1 and 2 are used for predicting SSMI and models 3 to 5 are used for predicting the SSFI. During the

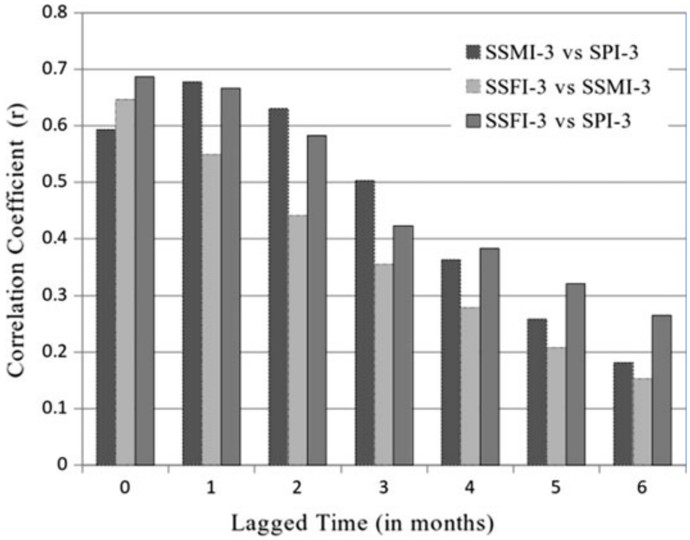


Fig. 7 Pairwise linear correspondence between SPI, SSMI and SSFI with lags during model development period in basin-I. Lags are applicable for the second index as shown in the legends for different pairs

application of models, the predictor drought time series is first decomposed into its components using MRSWT up to level 2. For instance, the components of SPI-3 for basin-I are shown in Fig. 8. The model performances during the development period and testing period are tabulated in Tables 4, 5 and 6, 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 (Table 4a and 4b) as compared to Table 2. The improvement in performance is more apparent in case of higher averaging period. For instance, in case of basin-I corresponding to 1 month 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 2). Though it is apparent that model 2 (for SSMI) and 5 (for SSFI) are best among other alternatives, it should be noted that the previous values of SSMI and SSFI are used in model 2 and 5, respectively. On the other hand, model 1 uses only information of SPI (with lags) and model 4 uses only SPI and SSMI (with lags), not the previous values of predictant series. Thus, the merit of model 1 (in case of SSMI) and model 4 (in case of SSFI) should be duly credited.

It is also noticed from the 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, in basin-I, the correlation

Table 3 Details of different types of models (No. 1 to 5). 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. $T_1 = 2^D$, where D is the level of decomposition, hence, $T_1 = 2^2 = 4$ and $T_{n+1} = T_n + 1$ for $n = 1,\ 2, \ldots$

Model no.	Model description
1	$SSMI(t) = f\begin{pmatrix} SPI_{a_2}(t-T_1), SPI_{d_2}(t-T_1), SPI_{d_1}(t-T_1), \\ SPI_{a_2}(t-T_2), SPI_{d_2}(t-T_2), SPI_{d_1}(t-T_2) \end{pmatrix}$
2	$SSMI(t) = f\begin{pmatrix} SPI_{a_2}(t-T_1), SPI_{d_2}(t-T_1), SPI_{d_1}(t-T_1), \\ SPI_{a_2}(t-T_2), SPI_{d_2}(t-T_2), SPI_{d_1}(t-T_2), \\ SSMI_{a_2}(t-T_1), SSMI_{d_2}(t-T_1), SSMI_{d_1}(t-T_1) \end{pmatrix}$
3	$SSFI(t) = f\begin{pmatrix} SPI_{a_2}(t-T_1), SPI_{d_2}(t-T_1), SPI_{d_1}(t-T_1), \\ SPI_{a_2}(t-T_2), SPI_{d_2}(t-T_2), SPI_{d_1}(t-T_2) \end{pmatrix}$
4	$SSFI(t) = f\begin{pmatrix} SPI_{a_2}(t-T_1), SPI_{d_2}(t-T_1), SPI_{d_1}(t-T_1), \\ SPI_{a_2}(t-T_2), SPI_{d_2}(t-T_2), SPI_{d_1}(t-T_2), \\ SSMI_{a_2}(t-T_1), SSMI_{d_2}(t-T_1), SSMI_{d_1}(t-T_1) \end{pmatrix}$
5	$SSFI(t) = f\begin{pmatrix} SPI_{a_2}(t-T_1), SPI_{d_2}(t-T_1), SPI_{d_1}(t-T_1), \\ SPI_{a_2}(t-T_2), SPI_{d_2}(t-T_2), SPI_{d_1}(t-T_2), \\ SSMI_{a_2}(t-T_1), SSMI_{d_2}(t-T_1), SSMI_{d_1}(t-T_1), \\ SSFI_{a_2}(t-T_1), SSFI_{d_2}(t-T_1), SSFI_{d_1}(t-T_1) \end{pmatrix}$

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

Model performance during testing period is shown in Tables 5 and 6. As mentioned earlier, two different validation schemes are followed for MLR version of models. For MLR versions of model 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 in case of basin-I, however, the opposite behavior is observed in case of basin-II. This observation suggests that in basin-I 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 compe-

Table 4 Performance of model no. 1 to 5 during development period

(a) using MLR

Basin	Averaging period (in months)	Performance measures	Model no. 1	2	3	4	5
Basin-I	1	r	0.657	0.831	0.581	0.589	0.607
		D_r	0.612	0.733	0.620	0.620	0.626
		uRMSE	0.730	0.538	0.660	0.655	0.644
	3	r	0.748	0.941	0.736	0.743	0.792
		D_r	0.666	0.837	0.687	0.690	0.719
		uRMSE	0.632	0.322	0.559	0.552	0.503
Basin-II	1	r	0.447	0.815	0.434	0.438	0.513
		D_r	0.564	0.733	0.570	0.571	0.693
		uRMSE	0.785	0.508	0.816	0.814	0.778
	3	r	0.558	0.935	0.672	0.674	0.767
		D_r	0.584	0.831	0.651	0.653	0.715
		uRMSE	0.725	0.309	0.694	0.692	0.601

(b) using ANN

Basin	Averaging period (in months)	Performance measures	Model no. 1	2	3	4	5
Basin-I	1	r	0.681	0.873	0.563	0.428	0.567
		D_r	0.629	0.758	0.599	0.570	0.420
		uRMSE	0.710	0.472	0.675	0.735	0.869
	3	r	0.579	0.962	0.619	0.818	0.888
		D_r	0.550	0.867	0.612	0.725	0.783
		uRMSE	0.779	0.261	0.669	0.476	0.383
Basin-II	1	r	0.446	0.789	0.374	0.387	0.278
		D_r	0.543	0.696	0.438	0.502	0.408
		uRMSE	0.803	0.540	0.840	0.865	1.138
	3	r	0.517	0.941	0.657	0.781	0.807
		D_r	0.557	0.837	0.645	0.723	0.709
		uRMSE	0.755	0.297	0.709	0.589	0.555

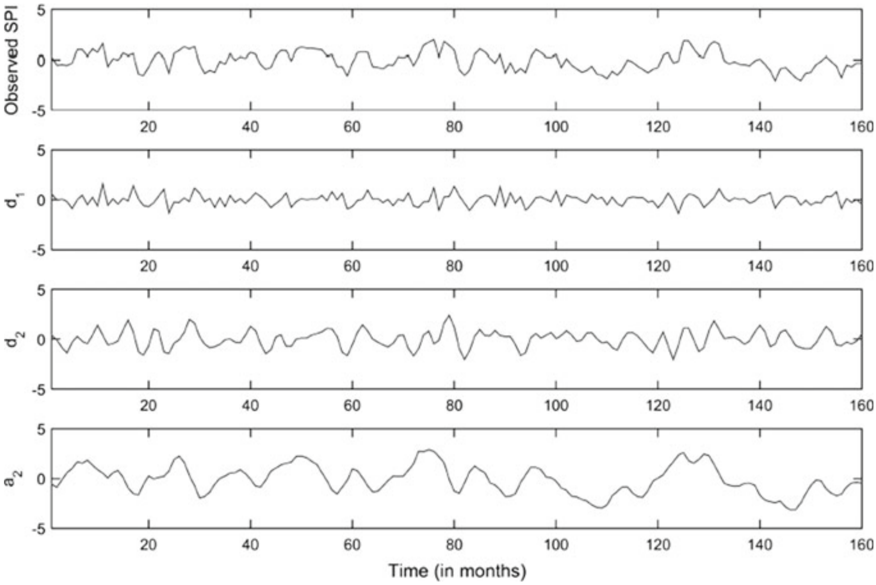


Fig. 8 Observed SPI-3 and its decomposed components up to level 2 for basin-I, i.e., a_2, d_2 and d_1, using Haar MRSWT. Figure shows the first 160 data points of decomposed series, i.e., development period for models. Such decomposed series for SSMI-3 and SSFI-3 are also obtained (not shown) for both basins

tent in modeling these dynamic relationships, produces better results. For example, with the validation scheme I and for predicting SSFI-3, the model 5 performance measures (r, D_r and uRMSE) are 0.792, 0.693, and 0.698, respectively, whereas the same with validation scheme II are 0.801, 0.712, and 0.682, respectively. Thus, the validation scheme II may be considered as more suitable where the correspondence between predictor and predictant may get modified over time due to the various reasons, including changing basin characteristics, climate regime, etc.

Interestingly, during testing, models using MLR version are found to perform comparable to ANN version in most of cases. This observation suggests that decomposed wavelet coefficient has linear relationship, so ANN version could not add much to the performance achieved by MLR version. Moreover, as stated earlier the per-

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

Basin	Averaging period (in months)	Validation scheme	Performance measures	Model no.				
				1	2	3	4	5
Basin-I	1	I	r	0.671	0.871	0.446	0.427	0.550
			D_r	0.628	0.761	0.544	0.538	0.580
			uRMSE	0.766	0.507	1.018	1.031	0.949
		II	r	0.652	0.862	0.496	0.571	0.642
			D_r	0.621	0.754	0.566	0.594	0.627
			uRMSE	0.782	0.522	0.990	0.934	0.873
	3	I	r	0.720	0.954	0.636	0.625	0.792
			D_r	0.634	0.850	0.611	0.607	0.693
			uRMSE	0.709	0.304	0.879	0.889	0.698
		II	r	0.709	0.947	0.646	0.711	0.801
			D_r	0.633	0.843	0.630	0.659	0.712
			uRMSE	0.716	0.324	0.869	0.801	0.682
Basin-II	1	I	r	0.652	0.841	0.465	0.491	0.560
			D_r	0.592	0.748	0.556	0.557	0.589
			uRMSE	0.954	0.641	0.987	0.973	0.922
		II	r	0.651	0.837	0.391	0.388	0.497
			D_r	0.623	0.750	0.545	0.535	0.566
			uRMSE	0.914	0.645	1.023	1.030	0.969
	3	I	r	0.735	0.935	0.735	0.749	0.791
			D_r	0.629	0.836	0.618	0.623	0.683
			uRMSE	0.863	0.415	0.744	0.727	0.666
		II	r	0.728	0.934	0.700	0.691	0.763
			D_r	0.658	0.836	0.646	0.638	0.677
			uRMSE	0.828	0.420	0.779	0.783	0.700

formance 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 5 for validation scheme II are shown in Fig. 9.

The models are checked for sensitivity for mother wavelet selection and development data length. Mother wavelet sensitivity 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 perfor-

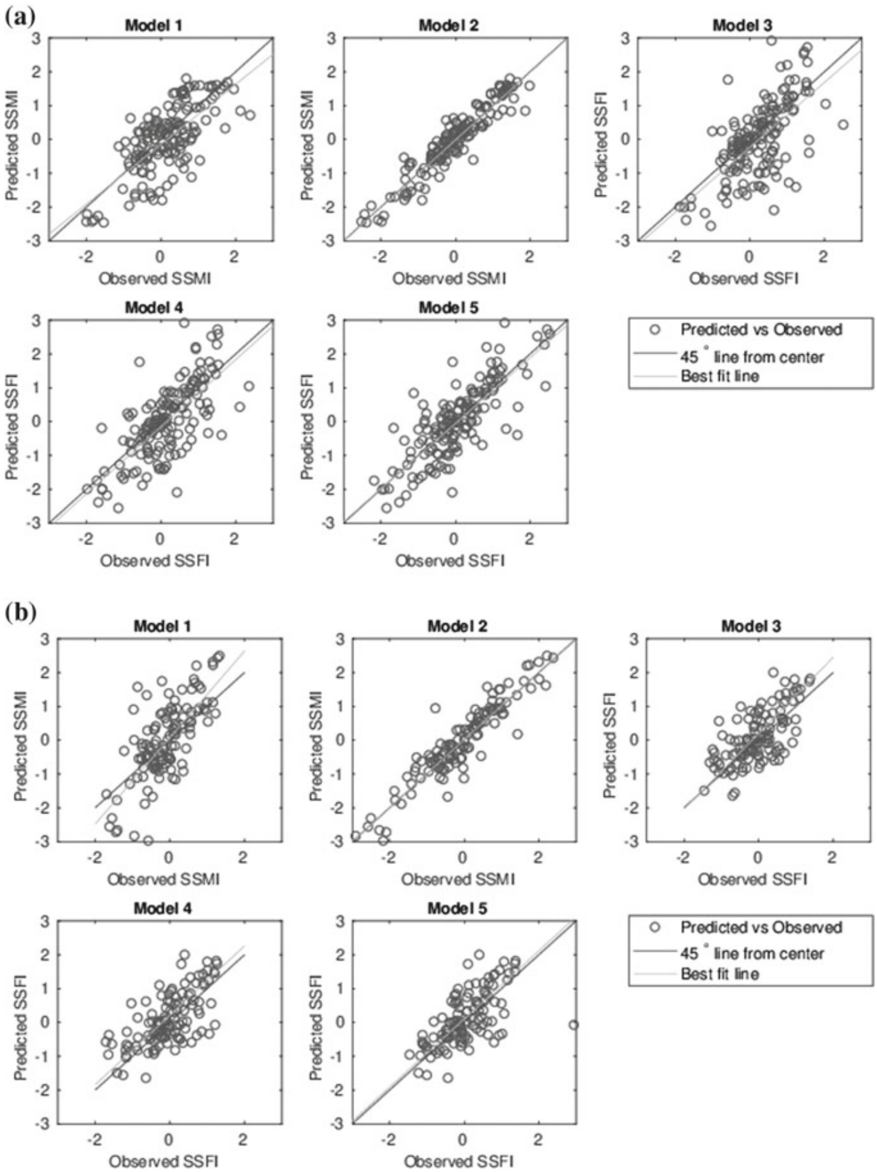


Fig. 9 Scatter plot between observed and predicted SSMI-3 and SSFI-3 by MLR version of models 1 to 5 during the testing period with validation scheme II for **a** Basin-I and **b** Basin-II

Table 6 Performance for model no. 1 to 5 during model testing period using ANN (only for validation scheme I)

Basin	Averaging period (in months)	Performance measures	Model no.				
			1	2	3	4	5
Basin-I	1	r	0.619	0.817	0.504	0.423	0.523
		D_r	0.592	0.712	0.533	0.544	0.517
		uRMSE	0.827	0.596	0.990	1.038	1.100
	3	r	0.638	0.884	0.596	0.566	0.698
		D_r	0.538	0.778	0.592	0.515	0.617
		uRMSE	0.789	0.473	0.939	1.059	0.822
Basin-II	1	r	0.620	0.796	0.426	0.420	0.262
		D_r	0.588	0.704	0.512	0.516	0.344
		uRMSE	0.929	0.719	1.014	1.020	1.309
	3	r	0.653	0.933	0.754	0.600	0.389
		D_r	0.616	0.829	0.587	0.513	0.546
		uRMSE	0.913	0.425	0.727	0.927	1.019

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

6 Summary and Concluding Remarks

Nonstationary nature of hydrologic variables, owing to various reasons including climatic change, poses a mathematical challenge to its predictability. In this chapter, the potential of wavelet transform is investigated in this regard. Initially, a brief introduction to various wavelets is presented followed by mathematical background of three mostly used wavelet transform. Next, mathematical details of MRSWT are provided which is used in an illustrative problem on drought prediction using the concept of temporal translation of one type of drought to another type.

In the illustrative example, one small and another medium size watersheds were considered from central part of India. For modeling the propagation of one drought type to another, the drought indices series are first transformed to their MRSWT components and their interrelationship is modeled using either ANN- or MLR-based models. Two different types of validation schemes (I and II) are used. Validation scheme I assumes that the model development period and testing period do not change, hence, the underlying relationship between drought components is considered to be the same during development and testing periods. On the other hand, in validation scheme II, the relationship is assumed to evolve with time and such evolution is modeled by shifting the window for development and testing period by one month on each application of the model. Thus, the validation scheme II is more suitable in cases where the relationship between the decomposed components of drought indices time series is expected to be dynamic. On the other hand, validation scheme I assumes the relationship to be static.

The prediction of drought indices at component level is better as compared to when they are analyzed without decomposition. Hence, MRSWT is effective tools for decomposing the hydrological time series and the models developed utilizing the decomposed components usually have better prediction performance. For most of the cases, MLR-based models are found to perform comparable to their ANN counterparts. This observation suggests that decomposed wavelet coefficient has linear relationship, another benefit of MRSWT decomposition in our case (though it is not guaranteed). 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. Further, 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 (validation scheme II) 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 methodological framework based on MRSWT is general in nature can be applied to other similar problems of hydrologic prediction. However, as in case of many statistical models, the methodology heavily depends on the length of the available historical data to capture the temporal evolution properly.

References

1. Milly PCD, Betancourt J, Falkenmark M, Hirsch RM, Kundzewicz ZW, Lettenmaier DP, Stouffer RJ (2008) Stationarity is dead: whither water management? Science 319(5863):573–574
2. Pichuka S, Maity R (2018) Development of a time-varying downscaling model considering non-stationarity using a Bayesian approach. Int J Climatol 38(7):3157–3176
3. Kirchner JW, Feng X, Neal C (2000) Fractal stream chemistry and its implications for contaminant transport in catchments. Nature 403(6769):524
4. Şen Z (1980) Adaptive fourier analysis of periodic-stochastic hydrologic sequences. J Hydrol 46(3–4):239–249
5. Pan F, Smerdon J, 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
6. Westra S, Sharma A (2006) Dominant modes of interannual variability in australian rainfall analyzed using wavelets. J Geophys Res Atmos 111(D5)
7. Karthikeyan L, Kumar DN (2013) Predictability of nonstationary time series using wavelet and emd based arma models. J Hydrol 502:103–119

8. Smith LC, Turcotte DL, Isacks BL (1998) Stream flow characterization and feature detection using a discrete wavelet transform. Hydrol Process 12(2):233–249
9. Özger M, Mishra AK, Singh VP (2012) Long lead time drought forecasting using a wavelet and fuzzy logic combination model: a case study in texas. J Hydrometeorol 13(1):284–297
10. Maity R, Suman M, Verma NK (2016) Drought prediction using a wavelet based approach to model the temporal consequences of different types of droughts. J Hydrol 539:417–428
11. 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(3–4):149–178
12. Daubechies I (1988) Orthonormal bases of compactly supported wavelets. Commun Pure Appl Math 41(7):909–996
13. Maheswaran R, Khosa R (2012) Comparative study of different wavelets for hydrologic forecasting. Comput Geosci 46:284–295
14. Cao L, Hong Y, Fang H, He G (1995) Predicting chaotic time series with wavelet networks. Physica D 85(1–2):225–238
15. Shannon CE (1949) Communication in the presence of noise. Proc IRE 37(1):10–21
16. Bradley AP (2003) Shift-invariance in the discrete wavelet transform. In: *Proceedings of VIIth* digital image computing: techniques and applications. Sydney 2003
17. American Meteorological Society (1997) Policy statement: meteorological drought. Bull Amer Meteor Soc 78:847–849
18. Suman M, Maity R (2016) Foreseeing the agricultural and hydrological drought knowing the ongoing meteorological scenarios through wavelet analysis. In: Proceedings of the international conference on hydraulics, water resources, coastal & environmental engineering (HYDRO 2016 International), December 8–10, 2016 at CWPRS, Pune.
19. Maybank J et al (1995) Drought as a natural disaster. Atmos Ocean 33(2):195–222
20. Peters E, Torfs PJJF, van Lanen HAJ, Bier G (2003) Propagation of drought through groundwater—a new approach using linear reservoir theory. Hydrol Process 17(15):3023–3040. https://doi.org/10.1002/hyp.1274
21. Keyantash J, Dracup JA (2002) The quantification of drought: an evaluation of drought indices. Bull Am Meteor Soc 83(8):1167–1180
22. McKee TB, Doesken NJ, Kleist J others (1993) The relationship of drought frequency and duration to time scales. In: Proceedings of the 8th conference on applied climatology 17:179–183
23. Soltani S (2002) On the use of the wavelet decomposition for time series prediction. Neurocomputing 48(1–4):267–277
24. Krause P, Boyle D, Bäse F (2005) Comparison of different efficiency criteria for hydrological model assessment. Adv Geosci 5:89–97
25. Willmott CJ, Robeson SM, Matsuura K (2012) A refined index of model performance. Int J Climatol 32(13):2088–2094
26. Entekhabi D, Reichle RH, Koster RD, Crow WT (2010) Performance metrics for soil moisture retrievals and application requirements. J Hydrometeorol 11(3):832–840
27. ajeevan M., Bhate J (2009) A high resolution daily gridded rainfall dataset (1971–2005) for mesoscale meteorological studies. Curr Sci, 558–562
28. Fan Y, van den Dool H (2004) Climate prediction center global monthly soil moisture data set at 0.5 resolution for 1948 to present. J Geophys Res Atmos 109(D10)
29. India WRIS version 4 (2014) Non classified hydro observation data. http://india-wris.nrsc.gov.in/HydroObservationStationApp.html