Investigating rainfall trend over India using the wavelet technique

Ranjit Kumar Paul and P. S. Birthal

ABSTRACT

In this paper, using the wavelet technique we analysed rainfall behaviour in the country across different agro-climatic zones over a century. Findings indicate that at the national level there is no significant trend in rainfall in the long run, but there are pockets of change in the rainfall pattern. There was a significant increase in the rainfall in the arid zone, whereas in the humid, semi-arid tropics and semi-arid temperate zones the trend was downward but insignificant. The behaviour of rainfall was different during this period. Except in the arid zone, we find a similar trend in other zones – increasing initially, tapering off in the middle and then declining but with some difference in time intervals. In the arid zone, the behaviour of rainfall had been erratic. In the short run, the direction of change in trend remains the same as in the long run but the change is statistically significant.

Key words | agro-climate, India, rainfall trend, wavelets

Ranjit Kumar Paul (corresponding author)
ICAR – Indian Agricultural Statistics Research
    Institute,
New Delhi 110012,
India
E-mail: ranjitstat@gmail.com

P. S. Birthal
ICAR – National Institute of Agricultural Economics
    and Policy Research,
New Delhi 110012,
India

INTRODUCTION

Indian agriculture is rain-dependent, with almost two-thirds of the net cropped area being rain-fed. Therefore, a regular rainfall pattern is crucial for agricultural growth; too much or too little rain can have a devastating effect on agriculture. A regular pattern in rainfall may not be observed all the time. The series may behave erratically, with highs and lows. Recent evidence shows monsoon rainfall in India is becoming more erratic (Manton *et al.* 2001; Kumar *et al.* 2004; Goswami *et al.* 2006; Mall *et al.* 2006; Ghosh *et al.* 2010; Paul *et al.* 2011). Auffhammer *et al.* (2011) have shown that the Indian climate has become drier and warmer, adversely affecting the crop yields.

India is a large country, and regional patterns in rainfall and temperature are more important for agriculture. A few studies have analysed behaviour of rainfall at disaggregated spatial units, at the level of meteorological sub-division, state and district. Guhathakurta & Rajeevan (2006) found a sharp decrease in the number of rainy days and amount of rainfall in many sub-divisions of India. They reported a decline in rainfall in the months of June, July and September and an increase in August in a few sub-divisions. Goswami *et al.*

(2006) studied rainfall behaviour over central India for the period 1951–2000 and observed a significant increase in heavy rainfall events and a decrease in moderate rainfall events. They also reported a rise in the frequency of extreme rainfall events such as droughts and floods. Krishnakumar *et al.* (2009) studied variations in monthly, seasonal and annual rainfall in the state of Kerala for the period 1871–2005, and found a significant decline in the south west monsoon (June–September) rains, especially in the months of June and July; and an upward trend in winter rainfall in the months of January, February and April. Soltani *et al.* (2012) computed rainfall and rainy day trends in Iran.

Some studies have also attempted to predict rainfall behaviour and estimation of trends using recently developed nonparametric techniques like wavelets. Kallache *et al.* (2005) applied discrete wavelet transform (DWT) analysis to river run-off data from gauges in Southern Germany to separate the variability of small and large time-scales, assuming that the trend component is part of the latter. Adamowski *et al.* (2009) proposed a method for trend detection and estimation based on continuous wavelet transform

(CWT), which provides a very flexible and accurate tool for detecting and estimating complicated signals in hydro-climatologic series in Canada. Paul *et al.* (2011) have also used DWT for the estimation of trends in Indian monsoon rainfall and reported that that there is a significant decreasing trend in the same. Nalley *et al.* (2012) studied the DWT technique to analyse and detect trends in monthly, seasonally-based, and annual data from eight flow stations and seven meteorological stations in southern Ontario and Quebec during 1954–2008. Using wavelet and fuzzy logics to describe the frequency decomposition of annual rainfall and summer monsoon rainfall, Singh & Bhatla (2012) suggest that a combination of these two models provides a better forecast of rainfall behaviour. Ramana *et al.* (2013) used wavelet and artificial neural network (ANN) models for predicting monthly rainfall and found wavelet neural network models more effective in predicting rainfall than ANN models alone. Mahajan & Mazumdar (2013) applied non-linear regenerative predicting codes for rainfall forecasts, and predicted more than 1 year's rainfall trend with 90% or more accuracy. Paul *et al.* (2013) applied maximal overlap DWT (MODWT) for forecasting Indian monsoon rainfall. Adarsh & Reddy (2015) carried out trend analysis based on DWT on the post-monsoon rainfall time series of the Kerala subdivision, and the results show that there is a dominance of short-term periodicity of less than a decade in the subdivision.

Most of these studies have analysed rainfall behaviour either at national level or at state level. Estimates at these levels are too aggregative to provide regional variation in climate. India, being a large country, has considerable heterogeneity in agro-climatic conditions even within a state or sub-division, while an analysis of climate variables at a more disaggregated level, say the district, is desirable for a better understanding of their location-specific behaviour. In this paper, we have examined rainfall behaviour at the level of an agro-climatic zone by clustering the districts having similar agro-climatic conditions into homogeneous zones. The study of rainfall trends at the zone level may provide inputs to policymakers in designing regionally differentiated adaptation and mitigation strategies to cope with climate change. Moreover, in almost none of the previous studies, has MODWT been explored for estimation of trends in climate variables. In the present study, MODWT and DWT have been applied.

The paper is organized as follows: the next section deals with the methodology, followed by the results and discussion section, and ending with the conclusions in the final section.

METHODOLOGY

To examine the behaviour of rainfall we use district level data from 1901 to 2002, which are then aggregated into four broad homogeneous agro-climatic zones following Rao *et al.* (2005). Rao *et al.* (2005) have grouped districts into different agro-climatic zones based on the average annual temperature and length of growing period (LGP). The districts with an LGP range from 70 to 180 days, and with a mean monthly temperature (all months) of greater than 18 °C and a daily mean temperature greater than 20 °C during the growing period, are classified as semi-arid tropics. For the districts in the semi-arid temperate zone the LGP ranges from 70 to 180 days, and the daily mean temperature during the growing period ranges from 5–20 °C. One or more months in these districts have a monthly mean temperature below 5 °C. The districts in the arid zone have an LGP of fewer than 75 days, and those in the humid zone have an LGP of more than 270 days. These zones are arid, semi-arid tropics, semi-arid temperate and humid (Figure 1). Semi-arid tropics comprise the largest zone, and the arid zone has the smallest area.

To identify the trend in rainfall we have used two non-parametric methods, viz. Mann-Kendall (M-K) and MODWT. The advantage of nonparametric methods is that these are based on fewer assumptions underlying the rainfall distribution and are not very sensitive to abrupt breaks, which are a common feature of a rainfall series. We have also used a parametric method, that is, the deterministic linear trend. A brief description of these is given below.

Deterministic linear trend

We use a simple linear regression model where rainfall is regressed on time. Mathematically, the model can be written as:
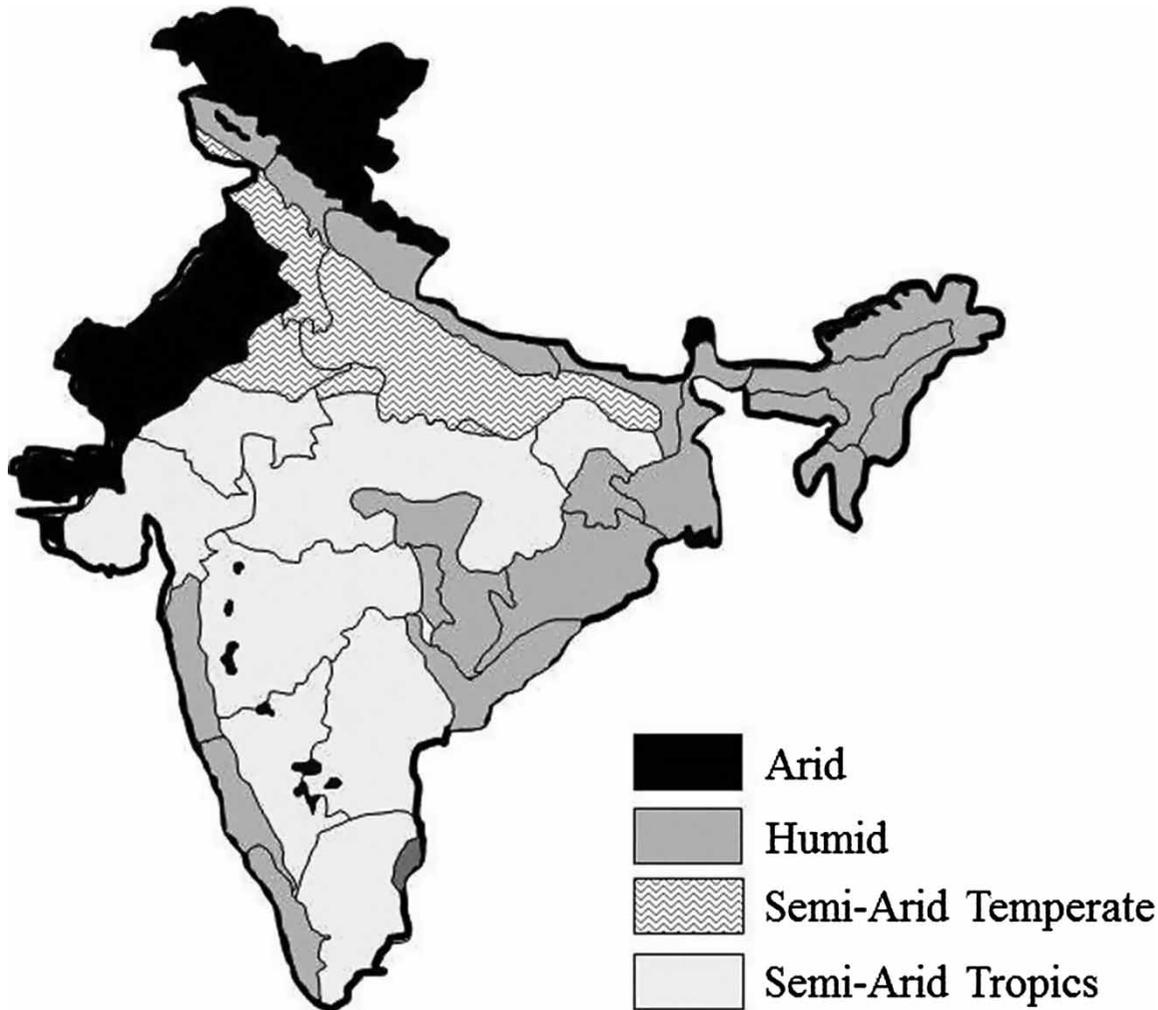
$$Y_t = \beta_0 + \beta_1 t + \varepsilon_t \tag{1}$$

Figure 1 │ Agro-climatic zones of India.

where Y_t is rainfall at time t, β's are the unknown parameters and ε_t is the error term that follows independent, identically distributed (IID) normal distribution, i.e. $\varepsilon_t \sim$ Normal $(0, \sigma^2)$. The advantage of linear regression is that it provides an estimate of slope, confidence interval and goodness of fit. However, it cannot handle missing data and may be greatly affected by outliers.

M-K test

The M-K test is a rank based nonparametric test which is used to detect the existence of a linear trend in a time series. The null hypothesis (H_0) for this test is that there is no trend in the series. Following Jayawardene *et al.* (2005), the first step

in this test is to replace observations (x_i) by their ranks (k_i), and then calculate N_i as the number of k_j preceding it such that $k_j > k_i$. The *t*-statistic for the test is computed as:

$$t_m = \frac{4\sum_i^{n-1} N_i}{n(n-1)} - 1 \tag{2}$$

Then, we define a range, r_m, for a desired probability point, r_g, as:

$$r_m = \pm r_g \sqrt{\frac{4n+10}{9n(n-1)}} \tag{3}$$

If t_m lies outside the range $\pm r_g$, then there exists a significant trend in the series. Further, to find the direction of the

trend, the S-statistic is used.

$$S = \sum_{k}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (4)$$

where $sgn()$ is the signum function. A signum function extracts the sign of a real number. If S is greater or lower than zero, the trend is increasing or decreasing respectively. For large values of n, S is approximately normally distributed with mean zero, i.e. $E(S) = 0$ and variance:

$$V(S) = \frac{n(n-1)(2n+5) - \sum_{i=1}^{m} t_i(i)(i-1)(2i-5)}{18} \qquad (5)$$

where t_i is the number of ties among observations having the same rank.

The Z-statistic can be used to test the significance of S:

$$Z = \begin{cases} \dfrac{S-1}{V(S)}, & \text{if } S > 0 \\ 0, & \text{if } S = 0 \\ \dfrac{S+1}{V(S)}, & \text{if } S < 0 \end{cases} \qquad (6)$$

It follows standard normal distribution. A positive value of Z signifies an upward trend in the series, and a negative value otherwise.

Spearman's rank statistic can also be used to confirm the results of the M-K test. The ranks (k_i) are arranged in increasing order, and then $d_i = k_i - 1$ is calculated. The test statistic is given by:

$$\rho_s = 1 - 6 \frac{\sum_i^n d_i^2}{n(n-1)} \qquad (7)$$

The significance of ρ_s can be tested by calculating t_S, where t_S is:

$$t_s = r_s \sqrt{\frac{n-2}{1-r_s^2}} \qquad (8)$$

The trend is absent if t_s lies within the desired limits.

DWT

Wavelets are fundamental building block functions, analogous to the trigonometric sine and cosine functions. As with a sine or cosine wave, a wavelet function oscillates around zero. This property makes the function a *wave*. A description of wavelets can be found in Daubechies (1992), Ogden (1997) and Percival & Walden (2000). A completely different use of wavelets in statistical time series analysis is described in Rioul & Vetterli (1991) and Flandrin (1998). Wavelets, being time-scale representation methods, deliver a tool complementary to both classical and localized Fourier analyses. The search for localized time-scale representations of stochastic correlated signals leads to both analysing and synthesizing (i.e. modelling) mainly non-stationary processes with wavelets or wavelet-like bases (Nason & von Sachs 1999).

There are two main waves of wavelets; the first is known as a CWT designed to work with time-series defined over the entire real axis; another is the DWT which deals with series defined essentially over a range of integers. The DWT of a time-series is used to capture high and low frequency components. This enables modelling of time-series data through computation of the inverse DWT. Computation of the DWT is carried out by the 'pyramid algorithm', as described in Percival & Walden (2000). A brief description of this protocol can be written as follows.

Let $\mathbf{X}$ be an $N \times 1$ vector of time series observations with $N = 2^J$; $\mathbf{P}$ be an $N \times N$ real valued matrix defining DWT and satisfying orthonormality property, i.e. $\mathbf{P'P} = \mathbf{I}_N$, where $\mathbf{I}_N$ is the $N \times N$ identity matrix. Then the DWT ($\mathbf{W}$) of the time-series vector $\mathbf{X}$ is computed by $\mathbf{W} = \mathbf{P}\ \mathbf{X}$. Now the elements of the vector $\mathbf{W}$ are decomposed into $J + 1$ sub-vectors. The first J sub-vectors contain all of the DWT coefficients for scale τ_j. Then $\mathbf{W}$ can be formed by concatenating all the $J + 1$ sub-vectors as follows:

$$\mathbf{W} = [\mathbf{W'}_1\, \mathbf{W'}_2 \ldots \mathbf{W'}_J\, \mathbf{V'}_J]' \qquad (9)$$

MODWT

The MODWT is a linear filtering operation that transforms a series into coefficients related to variations over a set of scales. It is similar to the DWT in that both are linear filtering

operations producing a set of time-dependent wavelet and scaling coefficients. Both have basis vectors associated with a location t and a unit-less scale $\tau_j = 2^j - 1$ for each decomposition level $j = 1, \ldots, J_0$. The MODWT retains down-sampled values at each level of decomposition that would otherwise be discarded by the DWT. The MODWT is well defined for all sample sizes N, whereas for a complete decomposition of J levels, the DWT requires N to be an integer multiple of 2^J. In the MODWT, the output signal is never sub-sampled (not decimated). Instead, the filters are upsampled on each level. The redundancy of the MODWT facilitates alignment of the decomposed wavelet and scaling coefficients at each level with the original time-series, thus enabling a ready comparison between the series and its decomposition. Finally, the redundancy of the MODWT wavelet coefficients modestly increases the effective degrees of freedom on each scale, and thus decreases the variance of certain wavelet-based statistical estimates.

MODWT is a highly redundant non-orthogonal transform. For a redundant transform like the MODWT, an N samples input time-series will have an N samples resolution scale for each resolution level. Therefore, the features of wavelet coefficients in a multi-resolution analysis (MRA) will be lined up with the original time-series in a meaningful way.

For a time-series X with arbitrary sample size N, the jth level MODWT wavelet ($\tilde{W}_j$) and scaling ($\tilde{V}_j$) coefficients are defined as:

$$\tilde{W}_{j,t} \equiv \sum_{l=0}^{L_j - 1} \tilde{h}_{j,l} X_{t-l \bmod N} \tag{10}$$

$$\tilde{V}_{j,t} \equiv \sum_{l=0}^{L_j - 1} \tilde{g}_{j,l} X_{t-l \bmod N} \tag{11}$$

where $\tilde{h}_{j,l} \equiv h_{j,l}/2^{j/2}$ are the jth level MODWT wavelet filters, and $\tilde{g}_{j,l} \equiv g_{j,l}/2^{j/2}$ are the jth level MODWT scaling filters. L_j is the width of the jth level equivalent wavelet and scaling filter. A wavelet filter must satisfy the following three basic properties:

$$\sum_{l=0}^{L-1} h_l = 0, \ \sum_{l=0}^{L-1} h_l^2 = 1 \text{ and } \sum_{l=0}^{L-1} h_l h_{l+2n} = \sum_{l=-\infty}^{\infty} h_l h_{l+2n}$$
$$= 0 \tag{12}$$

The first stage scaling coefficients are obtained as: $g_l = (-1)^{l+1} h_{L-1-l}$.

Estimation of trend by wavelets

Sometimes it is important to decompose a time-series into different components of variations such as low frequency (trend), and high-frequency (noise) components, and the MRA is used for decomposing and describing the low frequencies and high-frequency components in the data on a scale by scale basis. Consider the following model for a time-series data $\{X_t\}$:

$$X_t = \mu + T_t + Z_t, \ t = 0, \ldots, N-1, \tag{13}$$

where μ is a constant term, T_t is an unknown deterministic polynomial trend function of order r, Z_t is a residual term which is a long-memory process defined by $(1 - B)^\delta Z_t = \varepsilon_t$, where, δ is the long memory parameter, $\{\varepsilon_t\}$ is a Gaussian white noise process with mean zero and $\sigma_\varepsilon^2 > 0$. Here, B, is the back shift operator such that $BZ_t = Z_{t-1}$.

Since $\mathbf{W} = [\mathbf{W'}_1 \mathbf{W'}_2 \ldots \mathbf{W'}_J \mathbf{V'}_J]'$, the vector $\mathbf{W}$ can be written as the sum of two vectors: $\mathbf{W} = \mathbf{W_w} + \mathbf{W_s}$, where $\mathbf{W_w}$ is an $N \times 1$ vector containing the wavelet coefficients and zeros at all other locations, and $\mathbf{W_s}$ is an $N \times 1$ vector containing the scaling coefficients and zeros at all other locations. Since $\mathbf{X} = \mathbf{P'W}$, therefore:

$$\mathbf{X} = \mathbf{P'W} = \mathbf{P'W_s} + \mathbf{P'W_w} = \hat{T} + \hat{Z}, \tag{14}$$

Here, $\hat{T}$ is an estimator of the polynomial trend T at level J, while $\hat{Z}$ is the estimate of residual Z. The issue of choosing the level of the estimate depends on the goal of the application. J should be chosen to be small for detecting the local trends and cycles. In other applications, J is set to be large, if the objective is to detect the global trend.

The orthonormality of the matrix $\mathbf{P}$ implies that the DWT is an energy preserving transform so that:

$$\|\mathbf{X}\|^2 = \|\mathbf{W}\|^2 = \sum_{t=1}^{N} X_t^2 \tag{15}$$

Given the structure of the wavelet coefficients, the energy in $\mathbf{X}$ is decomposed, on a scale by scale basis, via:

$$\|\mathbf{X}\|^2 = \|\mathbf{W}\|^2 = \sum_{j=1}^{J} \|\mathbf{W}_j\|^2 + \|\mathbf{V}_J\|^2 \tag{16}$$

so that $\|\mathbf{W}_j\|^2$ represents the contribution to the energy of $\{X_t\}$ due to changes at scale τ_j, whereas $\|\mathbf{V}_J\|^2$ represents the contribution due to variations at scale τ_J. So the estimated variance of the time-series in terms of wavelet and scaling coefficients can be expressed as:

$$\sigma_X^2 = \frac{1}{N} \sum_{t=1}^{N} (X_t - \bar{X})^2 = \frac{1}{N} \|\mathbf{W}\|^2 - \bar{X}^2$$

$$= \frac{1}{N} \sum_{j=1}^{J} \|\mathbf{W}_j\|^2 + \frac{1}{N} \|\mathbf{V}_J\|^2 - \bar{X}^2 = \sum_{j=1}^{J} \hat{v}_X^2(\tau_j) + \hat{\sigma}_{S_J}^2 \tag{17}$$

where $\hat{v}_X^2(\tau_j)$ is the estimated variance of the wavelet coefficients at scale τ_j, and $\hat{\sigma}_{S_J}^2$ is the estimated variance of the trend.

For testing the null hypothesis H_0: *Trend = 0*, against the alternative hypothesis H_1: *Trend ≠ 0*, Almasri *et al.* (2008) proposed a test statistic as:

$$G = \frac{\hat{\sigma}_{S_J}^2}{\sum_{j=1}^{J} \hat{v}_X^2(\tau_j)} \tag{18}$$

The test statistic $(N - N/2^J)/(N/2^J - 1)$ G follows F distributed with $(N/2^J - 1)$ and $(N - N/2^J)$ degrees of freedom (under the normality assumption of the scaling coefficients).

RESULTS AND DISCUSSION

A preliminary investigation into rainfall behaviour across agro-climatic zones shows stable mean and variance, indicating that the series is stationary. The series is also not autoregressive as indicated by the values of autocorrelation and partial autocorrelation function at various lag orders (results are not reported here due to lack of space, but are available on request). The coefficients of the linear regression model and Theil–Sen slope estimates are given in the last two columns of Table 1, respectively. At the all India level, the trend in rainfall is positive but insignificant. It is positive and significant only in the arid zone. In other zones, the trend is negative and not significant. From these we find that over the years rainfall is becoming less, except in the arid zone where rainfall increased by 0.64 mm per year. The M-K test and Spearman's rank test confirm these observations except for the humid zone where one of the tests is significant. Here, the direction of change in rainfall in M-K and Spearman's test is opposite to that provided by linear regression. One possible reason could be the presence of some extreme events, i.e. very high or very low rainfall in some years.

For examining the local and global fluctuations in rainfall series, we computed MODWT coefficients along with the MRA for all the zones. The MODWT can accommodate any sample size N. In order to preclude decomposition at scales longer than the total length of the time-series, we select the highest scale (J_0) such that $J_0 \leq \log_2(N)$. In particular, for alignment of wavelet coefficients with the original series, condition $L_{J_0} < N$ (i.e. the width of the equivalent filter at the J_0th level is less than

Table 1 | Zone-wise rainfall trend during 1901–2002

Zones	M-K test			Spearman's rank test		Linear regression	
	s	t_m	r_m	t_s	Direction	b	Theil–Sen slope
Humid	−527	0.1023	0.1316	−1.6641	Decreasing	−0.6272	−0.553
Semi-arid temperate	109	−0.0211	0.1316	0.1841	Increasing	−0.0351	−0.161
Semi-arid tropic	−125	0.0242	0.1316	−0.5410	Decreasing	−0.2104	−0.180
Arid	805	−0.1562[a]	0.1316	2.2670 [a]	Increasing	0.6418 [a]	0.597[a]
All India	437	−0.0813	0.1316	0.8971	Increasing	0.4690	0.113

[a]Denotes significant trend at 5% level of significance.

the sample size) should be satisfied to prevent multiple wrappings of the time-series at level J_0. In our case, we found J_0 as 6. The choice of wavelet filter has a great significance on feature extraction of a time series. The decomposition would vary significantly, particularly for high frequency components with the choice of a different filter. A variety of 'mother wavelet functions' is available that can be used for DWT analysis (e.g. the Haar wavelet, Dsaubechies wavelet). The mother wavelet function should reflect the type of features present in the time
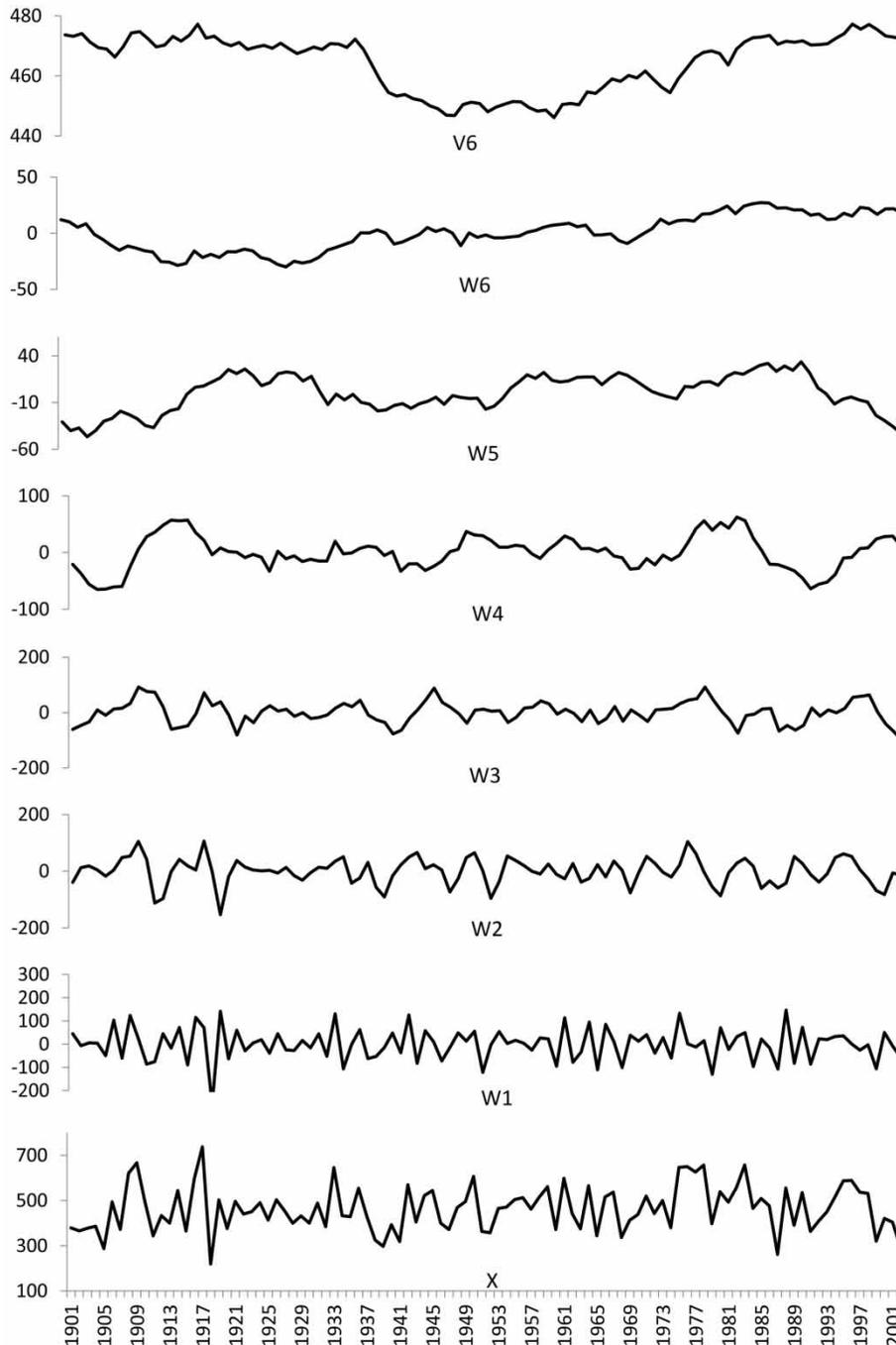
Figure 2 | (a) MODWT coefficients of rainfall in arid zone. (b) Vertical clustering of wavelet coefficients. *(Continued.)*
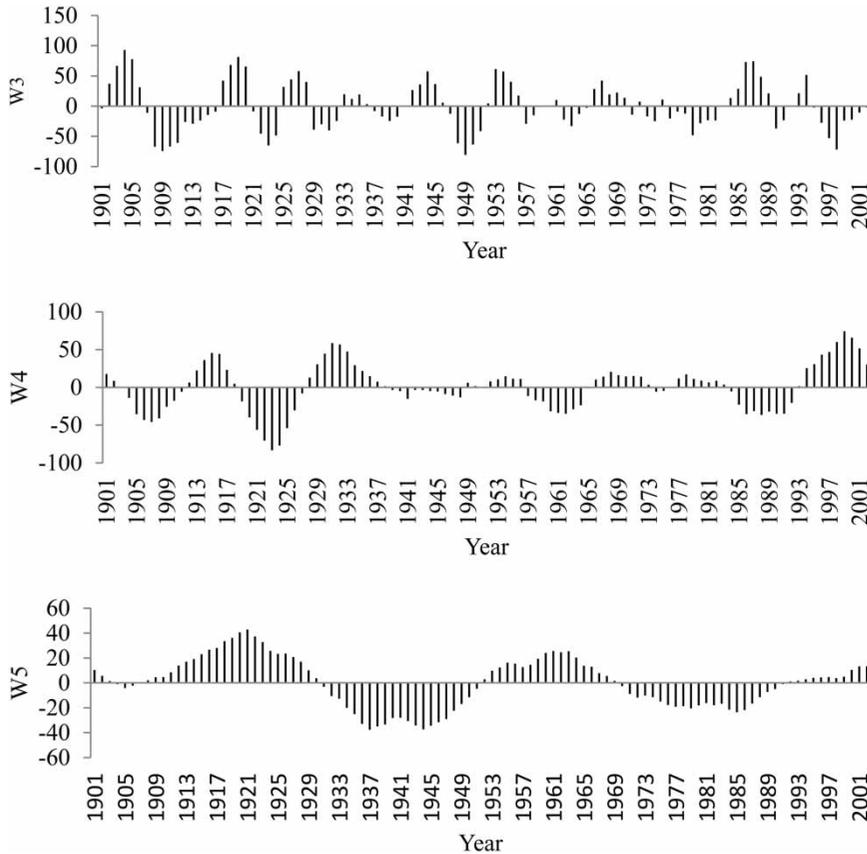
Figure 2 | Continued.

series. For time series with 'steps', one would choose a boxcar-like mother wavelet function such as the Haar wavelet, while for more smoothly varying time series one would choose a smoother mother wavelet function such as the Dsaubechies wavelet. In the present paper, both the filters have been used. Daubechies wavelets optimally capture the polynomial trends, while the Haar wavelet is discontinuous and resembles a step function. It is to be noted that for trend detection only, a low frequency component is used.

Haar and Daubechies (D4) wavelet filters are used for analysing data on a scale by scale basis to reveal its localized nature as exhibited by MODWT coefficients at level 6. In Figure 2(a) we report MODWT coefficients only for the arid zone, for which we obtained a significant trend in rainfall. MODWT coefficients for other zones are not reported due to space constraints, but are available on request. In Figure 2(a), X denotes the original time-series; W_1 to W_6 denote wavelet details components, and V_6

denotes the smoothed component of MODWT. A perusal of Figure 2(a) indicates that localized variation in the data is detected at a lower scale, whereas global variation is detected at a higher scale. At the lower scales the series are much more irregular compared to the series in the upper scale. In the lower scale we can detect trends for small time intervals, whereas in the larger scale we can visualize trends for the whole time period. The V_6 series is considered to be the proxy for trend.

Coefficients plotted at the top of Figure 2(a) are low-frequency components, and at the bottom are high-frequency components. The wavelet coefficients do not remain constant over time, and reflect changes in the data series at various time points. The locations of abrupt jumps can be spotted by looking at vertical clustering of the relatively larger coefficients. To obtain a visual idea regarding this, the MODWT coefficients at 3–5 levels, i.e. W_3, W_4 and W_5, have been plotted in Figure 2(b). A perusal of Figure 2(b) represents that whenever there are abrupt changes in

rainfall, it is depicted by clusters of long spikes on either side of the X-axis.

Rainfall trends estimated through wavelets are plotted in Figure 3(a)–(e). In the humid zone, overall, there is a slightly decreasing trend in rainfall (Figure 3(a)). However, the behaviour is different over the years; an increasing trend between 1901 and 1933; stagnation between 1934 and 1943 and a decreasing trend thereafter. Figure 3(b) shows rainfall behaviour in the semi-arid temperate zone. We observe almost a similar pattern of rainfall as in the humid zone, but with differences in time intervals; the trend is increasing from 1901 to 1946, then fairly constant until 1957 and decreasing thereafter. Similar trends are observed in the semi-arid tropics but at different time intervals (Figure 3(c)).

The behaviour of rainfall in the arid zone is in contrast with that in others. In this zone, the overall trend in rainfall is increasing, but the intervals of change in direction are greater, indicating more erratic rainfall behaviour. The intensity of rainfall decreases during 1901–1916, remains almost constant until 1931, increases until 1976 and then starts decreasing. At the all India level, rainfall has increased from 1950 onwards (Figure 3(e)).

The spatial average of the annual rainfall calculated on the basis of all grids over India show no discernible increasing or decreasing trend (Figure 4(a)). However, over the years the moving temporal variance calculated from the spatial means shows a statistically significant decreasing trend (Figure 4(b)). These findings are similar to those reported in Ghosh et al. (2012). A global increase in extreme rainfall events is a consequence of the emission of anthropogenic greenhouse gases causing atmospheric and oceanic warming (Ghosh et al. 2012). However, the spatial heterogeneity in rainfall as observed here could also be due to local influences such as topography, deforestation, non-uniform population growth and urbanization.

The trend analysis studies by Guhathakurta & Rajeevan (2006), Goswami et al. (2006) and Krishnakumar et al. (2009) indicate that trends significantly depend upon the
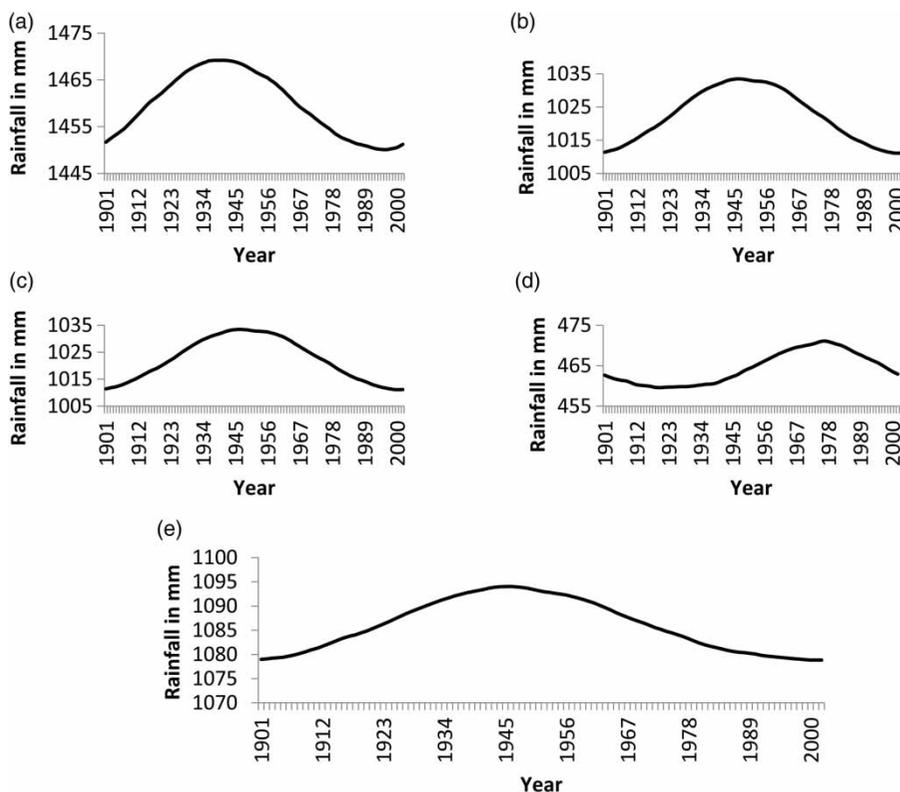

Figure 3 | Estimate of rainfall trend by MODWT using Haar wavelet for humid (a), semi-arid temperate (b), semi-arid tropic (c), arid (d) zone and all India (e).
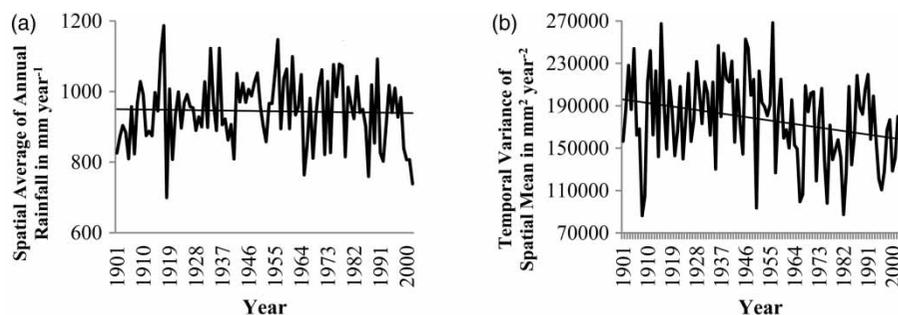
Figure 4 | Spatiotemporal trends and variability in all India rainfall with 102-year overlapping time windows. (a) Spatial average of the annual rainfall; (b) temporal variance of the spatial mean.

period and the locations. To see this, we analysed rainfall behaviour for the period 1939–2002, applying the same methods as for the entire period but instead of MODWT we applied DWT (the DWT requires that the number of observations (N) should be sufficient such that $N = 2^{\wedge J}$, J takes a positive integer). Also, we used two wavelet filters, viz. Haar and Daubechies (D4).

The results are reported in Table 2. With application of the wavelet method, the direction of trend in rainfall remains as in the case of the entire period, but the trend coefficients are now significant. These results indicate that there has been a significant decline in rainfall in the past six decades or so (1939–2002) in all the zones, except in the arid zone and at the all India level, where it has increased. The wavelet analysis could determine a significant trend globally as well as locally that could not be captured through other methods.

The rainfall trends for the period 1939–2002 estimated through DWT are shown in Figure 5(a)–(e). In the humid

zone, there is a decreasing trend until 1954, followed by an increasing trend between 1955 and 1969. A mildly decreasing trend is observed thereafter until 1986, and an increasing trend later on. Until the late 1960s we observed similar trends for semi-arid tropics and temperate zones. Afterwards, there is a declining trend, more so in the semi-arid tropics. In the arid zone, the behaviour of rainfall is different; it declines until 1954 and starts increasing thereafter. At the all India level, the rainfall behaves differently. It has a decreasing trend during 1939 to 1954; thereafter it increases up to 1970 and then remains almost stagnant until 2002.

CONCLUSIONS

Agriculture in India is highly sensitive to rainfall, therefore any change in rainfall will influence crop yields. An understanding of the spatial and temporal distribution and changing patterns in rainfall is important for planning and

Table 2 | Zone-wise rainfall trend during 1939–2002

Zones	M-K test			Linear regression		Test statistic based on wavelets	
	S	Tm	Rm	B	Theil–Sen slope	D4	Haar
Humid	−418	0.207[a]	0.167	−1.873[a]	−1.855[a]	17.859[a]	7.122[a]
Semi-arid temperate	−306	0.151	0.167	−1.861[a]	−1.732[a]	15.876[a]	24.191[a]
Semi-arid tropic	−378	0.187[a]	0.167	−1.999[a]	−1.965[a]	13.573[a]	9.241[a]
Arid	208	−0.103	0.167	0.636	0.722	9.608[a]	12.812[a*]
All India	−927	0.308[a]	0.167	−5.268[a]	−4.842[a]	19.645[a]	25.470[a]

[a]Denotes statistically significant trend at 5% level.

Note: Statistics for testing trend using D4 and Haar wavelets were estimated using $(N–N/2^j)/(N/2^j–1)$G described above under the section 'Estimation of trend by wavelets'.
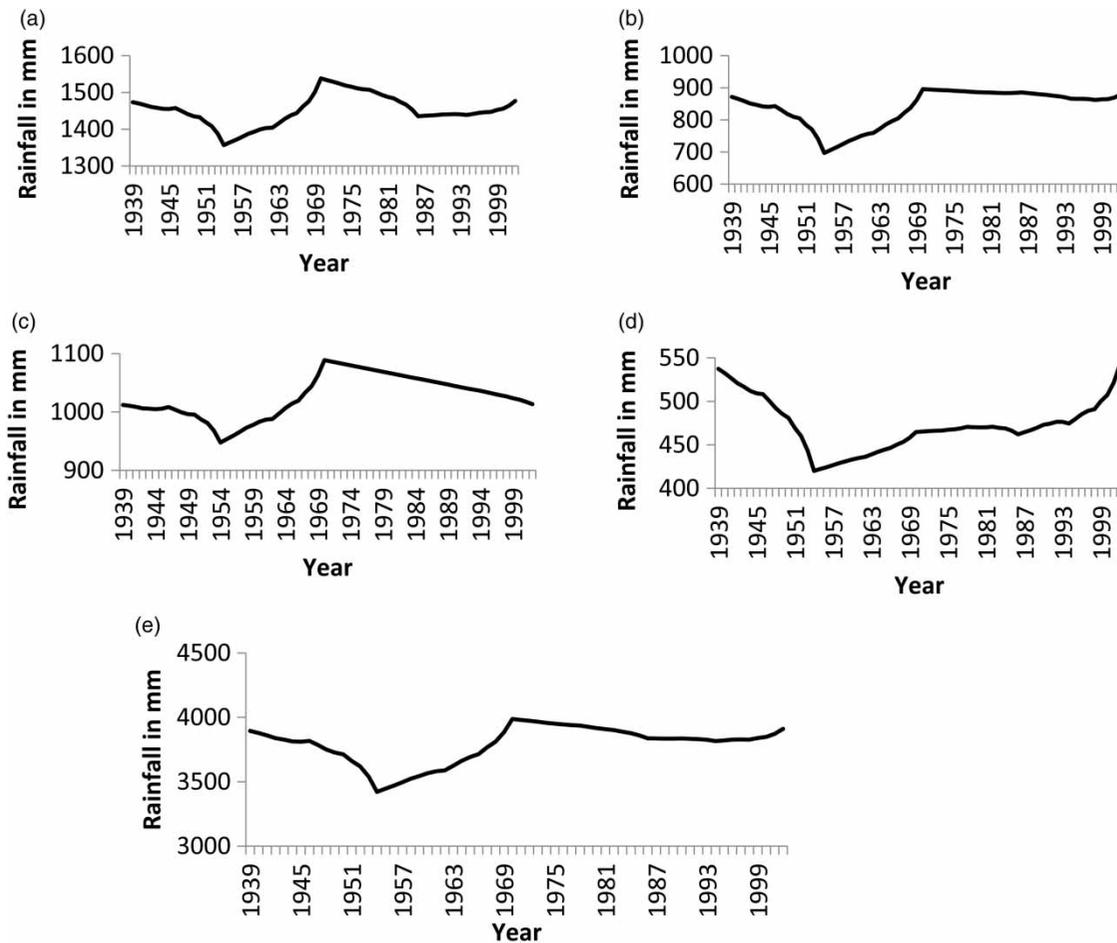
Figure 5 | Estimate of trend in rainfall by DWT using D4 wavelet for humid (a), semi-arid temperate (b), semi-arid tropic (c), arid (d) zone and all India (e).

management of water resources. In this paper we examined behaviour of rainfall in different agro-climatic zones over a period of more than 100 years using wavelet techniques. The trends were tested for two periods, a long-term period (1901–2002) and for a shorter period (1939–2002). Findings indicate that at the national level there is no significant trend in rainfall in the long run, but there are pockets of change in the rainfall pattern. There was a significant increase in the rainfall in the arid zone, whereas in the humid, semi-arid tropics and semi-arid temperate zones the trend was downward but insignificant. In the short run, the direction of change in trend remains the same as in the long run, but the change is statistically significant. Since rainfall has been decreasing in three of the four agro-climatic zones, a possible sign of temporal change in Indian rainfall over the past century may be apparent.

REFERENCES

Adamowski, K., Prokoph, A. & Adamowski, J. 2009 Development of a new method of wavelet aided trend detection and estimation. *Hydrol. Process.* **23**, 2686–2696.

Adarsh, S. & Reddy, M. J. 2015 Trend analysis of rainfall in four meteorological subdivisions of southern India using nonparametric methods and discrete wavelet transforms. *Int. J. Climatol.* **35** (6), 1107–1124.

Almasri, A., Locking, H. & Shukur, G. 2008 Testing for climate warming in Sweden during 1850–1999, using wavelets analysis. *J. Applied Statistics* **35**, 431–443.

Auffhammer, M., Ramanathan, V. & Vincent, J. R. 2011 Climate change, the monsoon, and rice yield in India. *Clim. Chang.* **111**, 411–424.

Daubechies, I. 1992 *Ten Lectures on Wavelets*. SIAM, Philadelphia.

Flandrin, P. 1998 *Time-Frequency/Time-Scale Analysis. Wavelet Analysis and Its Applications*. Vol. 10. Academic Press, San Diego, CA.

Ghosh, H., Paul, R. K. & Gupta, P. 2010 Wavelet frequency domain approach for statistical modeling of rainfall time-series data. *J. Stat. Theory Pract.* **4**, 813–825.

Ghosh, S., Das, D., Kao, S. & Ganguly, A. R. 2012 Lack of uniform trends but increasing spatial variability in observed Indian rainfall extremes. *Nat. Clim. Chang.* **2**, 86–91.

Goswami, B. N., Venugopal, V., Sengupta, D., Madhusoodanan, M. S. & Xavier, P. K. 2006 Increasing trends of extreme rain events over India in a warming environment. *Science* **314**, 1442–1445.

Guhathakurta, P. & Rajeevan, M. 2006 *Trends in Rainfall Pattern Over India*. National Climate Centre. Report 2. India Meteorological Department, Pune, India.

Jayawardene, H. K. W. I., Sonnadara, D. U. J. & Jayewardene, D. R. 2005 Trends of rainfall in Sri Lanka over the last century. *Sri Lankan J. Phys.* **6**, 7–17.

Kallache, M., Rust, H. W. & Kropp, J. 2005 Trend assessment: applications for hydrology and climate research. *Nonlinear Process. Geophys.* **12**, 201–210.

Krishnakumar, K. N., Prasada Rao, G. S. L. H. V. & Gopakumar, C. S. 2009 Rainfall trends in twentieth century over Kerala, India. *Atmos. Environ.* **43**, 1940–1944.

Kumar, K. K., Kumar, K. R., Ashrit, R. G., Deshpande, N. R. & Hansen, J. W. 2004 Climate impacts on Indian agriculture. *Int. J. Climatol.* **24**, 1375–1393.

Mahajan, S. & Mazumdar, H. S. 2013 A regenerative prediction algorithm for Indian rainfall prediction. *Int. J. Eng. Sci.* **5**, 1832–1837.

Mall, R. K., Singh, R., Gupta, A., Srinivasan, G. & Rathore, L. S. 2006 Impact of climate change on Indian agriculture: a review. *Clim. Chang.* **78**, 445–478.

Manton, M. J., Della-Marta, P. M., Haylock, M. R., Hennessy, K. J., Nicholls, N., Chambers, L. E., Collins, D. A., Daw, G., Finet, A., Gunawani, D., Inape, K., Isobe, H., Kestin, T. S., Lefale, P., Leyu, C. H., Lwin, T., Maitrepierre, L., Ouprasitwong, N., Page, C. M., Pahalad, J., Plummer, N., Salinger, M. J., Suppiah, R., Tran, V. L., Trewin, B., Tibig, I. & Yee, D. 2001 Trends in extreme daily rainfall and temperature in southeast Asia and the south-pacific: 1961–1998. *Int. J. Climatol.* **21**, 269–284.

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

Nason, G. P. & von Sachs, R. 1999 Wavelets in time series analysis. *Phil. Trans. R. Soc. Lond. A* **357**, 2511–2526.

Ogden, T. 1997 *Essential Wavelets for Statistical Applications and Data Analysis*. Birkhauser, Boston.

Paul, R. K., Ghosh, H. & Gupta, P. 2011 Wavelet methodology for estimation of trend in Indian monsoon rainfall time series data. *Ind. J. Agric. Sci.* **81**, 290–292.

Paul, R. K., Gupta, P. & Ghosh, H. 2013 Wavelet frequency domain approach for modelling and forecasting of Indian monsoon rainfall time-series data. *J. Ind. Soc. Agric. Stat.* **67**, 319–327.

Percival, D. B. & Walden, A. T. 2000 *Wavelet Methods for Time-Series Analysis*. Cambridge University Press, Cambridge, UK.

Ramana, R. V., Krishna, B., Kumar, S. R. & Pandey, N. G. 2013 Monthly rainfall prediction using wavelet neural network analysis. *Water Resour. Manage.* **27**, 3697–3711.

Rao, K. P. C., Bantilan, M. C. S., Singh, K., Subrahmanyam, S., Deshingkar, P., Rao, P. P. & Shiferaw, B. 2005 *Overcoming Poverty in Rural India: Focus on Rainfed Semi-Arid Tropics*. International Crops Research Institute for the Semi-Arid Tropics, Andhra Pradesh, India, p. 96.

Rioul, O. & Vetterli, M. 1991 Wavelets and signal processing. *IEEE Sig. Proc. Mag.* **8**, 14–38.

Singh, R. C. & Bhatla, R. 2012 Wavelets in meteorology. *AIP Conf. Proc.* **1463**, 176–193.

Soltani, S., Saboohi, R. & Yaghmaei, L. 2012 Rainfall and rainy days trend in Iran. *Clim. Chang.* **110**, 187–213.