Wavelets based estimation of trend in sub-divisional rainfall in India

RANJIT KUMAR PAUL, SANDIPAN SARKAR, DIPANKAR MITRA, SANJEEV PANWAR*, A. K. PAUL and L. M. BHAR
ICAR-Indian Agricultural Statistics Research Institute, New Delhi – 110 012, India
*Indian Council of Agricultural Research, New Delhi – 110 012, India
(Received 8 February 2018, Accepted 14 June 2019)
e mail : ranjitstat@gmail.com, ranjit.paul@icar.gov.in

ABSTRACT. Presence of long memory in climatic variables is frequently observed. The trend assessment becomes difficult in the presence of long-memory as the usual methods are not capable to take care of this property during trend estimation. In order to estimate the trend in presence of long memory, the non-parametric wavelet method has become popular in the recent time. The discrete wavelet transformation (DWT) re-expresses a time-series in terms of coefficients that are associated with a particular time and a particular scale. In the present study, DWT has been applied to estimate the monthly rainfall trend for the monsoon months: June-June in ten selected sub-divisions of India using “Haar” wavelet filter. The results from DWT were cross checked with the non-parametric Mann-Kendall (M-K) test. The investigation reveals that the monthly rainfall trend for the monsoon months of different sub-divisions in India are significantly decreasing over the years. However, in some of the sub-divisions, rainfall trend is increasing. DWT reveals significant trend in most of the sub-divisions whereas M-K test reveals that most of the trends are not significant at 5% level.

Key words – Haar filter, Long memory, Mann-Kendall test, Trend, Wavelets.

1. Introduction

A common issue in time series analysis is the decomposition of several time series components viz., low frequency (trend) component and high frequency (noise and periodicity) component. Most of the time series of aggregated variables show a steadily increasing or decreasing pattern, known as trend. Over last few decades Box Jenkin’s Autoregressive integrated moving average (ARIMA) methodology (Box et al., 2007), a parametric approach of time series analysis has been used for forecasting time series data. But there are certain circumstances where it is not possible to postulate appropriate parametric relationship for the underlying phenomena; in this case nonparametric approach is called for.

A plausible statistical model for such series can consist of non-stochastic or trend component and stochastic component: \( Y(t) = T(t) + X(t) \), where \( Y(t) \) is the value of an observable time series at time \( t \), \( T(t) \) is the trend component and \( X(t) \) is called the noise process with mean zero. Trend assessment is the problem of determining whether or not \( T(t) \) is actually present in the time series, \( i.e., \) test the null hypothesis \( H_0: T(t) = 0 \). In many agricultural data, like daily commodity price data,
daily rainfall and temperature data it is seen that the distant observations are dependent that means the dataset have characteristic feature of long memory or long range dependency. A time-series process is called as long memory or Fractional differenced (FD) process if the autocorrelation function decays very slowly towards zero unlike the exponential decay in usual ARIMA model. A major problem in practice is the discrimination of a FD(d) (d denotes the long memory parameter) process with trend from a FD(d) process without trend, since FD(d) processes for 0 < d < ½ usually looks like having some trend in the series. This makes it hard to distinguish the existence of trend in the time series.

Here our interest lies on the testing of low-frequency part or called as trend, in wavelet domain (Antoniadis, 1997). Let us consider the assumptions that the trend T(t) is well approximated at least locally by a low order polynomial (such as linear) and that the stochastic component X(t) is a FD process. Under these assumptions, the discrete wavelet transform (DWT) (Percival and Walden, 2000) based upon the Haar family of wavelet filters can be used to transform the series Y(t). The ability of the DWT to cleanly separate Y(t) into these components allows us to test for its significance.

Almasri et al. (2010) have proposed a test statistic by using wavelet decompositions to test the significance of trend in a time series data. The most difficult problem of testing for linear trend is the presence of dependence among the residuals because of which, tests for trend based on the classical ordinary least squares (OLS) regression are inappropriate. In many situations, the error auto covariance function exhibits a slow decay reflecting the possible presence of long memory process. The wavelet analysis, however, has been extensively used for such purposes, since it suitably matches the structure of these processes. Paul et al. (2011) studied wavelet methodology for estimation of trend in Indian monsoon rainfall time-series data and reported significant trend.

Indian agriculture is mainly rain-dependent; approximately two-third net cropped area is under rain-fed cultivation. Therefore, regular and uniform rainfall pattern is crucial for betterment of agriculture; extreme situations can affect the agriculture devastatingly. But problem is that the regular rainfall pattern may not be observed all the time. Several studies have been done on the pattern of regional rainfall in India (Kumar et al., 2004; Gowsami et al., 2006). Guhathakurta and Rajevean (2006) have found that there is a decline in rainfall in the months of June, July and September and an increase in the month of August in few sub-divisions in India. Gowsami et al. (2006) have studied the rainfall behaviour over central India of the period 1951-2000 and found that a significant increase in heavy rainfall events and a decrease in moderate rainfall events. Some studies have also carried out to predict the rainfall and estimate the trend using some non-parametric methods like wavelet (Kallache et al., 2005). Paul et al. (2013) have investigated the modelling of Indian monsoon rainfall data and concluded that wavelet methodology has greater accuracy than that of ARIMA model. Paul et al. (2015) investigated the trend in mean temperatures in different agro-climatic zones in India using both parametric and nonparametric methods. Paul and Birthal (2016) have applied wavelet approach for describing variability in rainfall in different agro-climatic zones of India. Paul (2017) found the significance presence of long memory in maximum and minimum temperatures in India. In this study we use wavelet based test statistic and non-parametric Mann-Kendall test statistic to test the existence of trend in the time series data. The paper is organized as follows: section 2 describes about long memory process; section 3 deals with the basics of wavelets; section 4 describes methods of detection of trend; section 5 deals with empirical illustration followed by conclusions in section 6.

2. Long memory process

Most of the research works in time-series analysis assume that the observation separated by long time span are independent of each other or nearly so. But in many practical situations it is seen that many empirical economic series show that the distant observations are dependent, though the correlation is small but not negligible. Let X_i(t = 0, 1, 2,...) be a stationary time-series process and the autocorrelation function of the time-series with a time lag of k is ρ_k. For long memory processes, decaying of autocorrelations functions occur at much slower rate (hyperbolic rate) which is consistent with ρ_k ≈ Ck^d, as k increases indefinitely, where C is a constant and d is the long memory parameter. A study of long memory time series in climate can be found in Paul and Anjoy (2018).

3. Wavelet

Wavelets (Daubechies, 1992; Ogden, 1997 and Vidakovic, 1999) are fundamental building block functions, analogous to the trigonometric sine and cosine functions. As with a sine or cosine wave, a wavelet function oscillates about zero. This oscillating property makes the function a wave. However, the oscillations for a wavelet damp down to zero, hence the name wavelet. If ψ(.) be a real valued function defined over the real axis (−∞, ∞) and satisfying two basic properties: the integral of ψ(.) is zero and the square of ψ(.) integrates to unity. The detail discussion of wavelets can be found in (Percival and Walden, 2000).
3.1. Discrete wavelet transform (DWT)

DWT of a time-series observation is used to capture high and low frequency components. The DWT re-expresses a time-series in terms of coefficients that are associated with a particular time and a particular dyadic scale $2^j$ (Nason & Von Sachs, 1999). Let, $X = (x_0, x_1, \ldots, x_{N-1})$ be a column vector containing $N$ observations of a real-valued time series, where we assume that $N$ is an integer multiple of $2^M$, where $M$ is a positive integer. The DWT of level $J$ is an orthonormal transform of $X$ defined by

$$d = (d_1, d_2, \ldots, d_j, \ldots, d_J, s_J) = WX \quad (1)$$

where, $W$ is an orthonormal $N \times N$ real-valued matrix, i.e., $W^{-1} = W'$ so $WW' = WW = I^N$ and called the wavelet matrix. $D_j = \{d_{jk}\}, j = 1, 2, \ldots, J,$ are $N/2^j \times 1$ real-valued vectors of wavelet coefficients at level $j$ associated with scale $\lambda_j$ and location $k$, where $\lambda_j = 2^j$. The real-valued vector $S_j = \{s_{jk}\}$ is made up of $N/2^j$ scaling coefficients. Thus, the first $N - N/2^j$ elements of $D$ are wavelet coefficients and the last $N/2^j$ elements are scaling coefficients, where $J \leq M$. In DWT several filters can be used, “Haar filter” is one of them.

4. Trend

Trend is defined as a long term change in the underlying mean level per unit time (Jain and Kumar, 2012). There are different models for describing various forms of trend, both linear and nonlinear and both stochastic and deterministic. The model that represents the time series by using a polynomial trend with $p = 1$ and a stationary process with long memory. Based on that we consider the following model:

$$X_t = T_t + Z_t = a + \beta t + Z_t \quad for \quad t = 0, \ldots, N - 1 \quad (3)$$

where, the process $Z_t$ is a residual term which is a long-memory process defined by

$$(1 - B)\delta Z_t = \epsilon_t, \quad (4)$$

where, $0 \leq \delta < 1/2$ is the long-memory parameter, $\{\epsilon_t\}$ is a Gaussian white noise process with mean zero and $\sigma^2 > 0$.

From (1), we can write

$$d = d_w + d_s \quad (5)$$

where, $d_w$ is an $N \times 1$ vector containing the wavelet coefficients and zeros at all other positions and $d_s$ is an $N \times 1$ vector containing the scaling coefficients and zeros at all other positions. Since $X = W'd$, we can write,

$$X = W'd = W'd_s + W'd_w = \hat{T} + \hat{Z} \quad (6)$$

where, $\hat{T}$ is an estimator of the polynomial trend $T$ at level $J$, while $\hat{Z}$ is a tapered ‘version’ of $X$.

An important issue is how to choose the wavelet filter. The Haar wavelet, which is a piecewise constant function, preserves the discontinuities and therefore it is most suitable to identify a structural break in the data.
TABLE 1
Descriptive statistics

| Zone | Month | Mean | SD  | Maximum | Minimum | CV  | Skewness | Kurtosis |
|------|-------|------|-----|---------|---------|-----|----------|----------|
| ASMEG | Jun   | 4405 | 1060.08 | 7760 | 2333 | 24.065 | 0.380 | 3.217     |
|       | Jul   | 4158 | 1029.17 | 7812 | 2398 | 24.751 | 0.844 | 3.625     |
|       | Aug   | 3508 | 823.535 | 6672 | 1952 | 23.476 | 0.664 | 4.104     |
|       | Sep   | 2867 | 785.427 | 5325 | 1308 | 27.395 | 0.641 | 3.291     |
| GNWBL | Jun   | 2502 | 980.389 | 5835 | 699  | 39.184 | 0.731 | 3.477     |
|       | Jul   | 3316 | 900.475 | 6736 | 1590 | 27.156 | 0.647 | 3.837     |
|       | Aug   | 3236 | 865.187 | 5632 | 1721 | 26.736 | 0.713 | 2.894     |
|       | Sep   | 2522 | 939.822 | 5967 | 924  | 37.258 | 1.308 | 4.813     |
| WUPPL | Jun   | 890  | 646.674  | 3007 | 39  | 72.660 | 1.116 | 3.889     |
|       | Jul   | 2589 | 876.867  | 4320 | 201  | 33.869 | -0.263 | 2.585     |
|       | Aug   | 2568 | 937.033  | 4948 | 650  | 36.489 | 0.149 | 2.663     |
|       | Sep   | 1497 | 905.677  | 3732 | 41   | 60.499 | 0.451 | 2.355     |
| PUNJ   | Jun   | 547  | 486.415  | 3360 | 13   | 88.875 | 2.597 | 13.572    |
|       | Jul   | 1774 | 800.455  | 5390 | 221  | 45.728 | 1.078 | 6.016     |
|       | Aug   | 1669 | 788.405  | 3847 | 237  | 47.238 | 0.637 | 3.134     |
|       | Sep   | 953  | 968.369  | 4412 | 0    | 101.55 | 1.749 | 5.671     |
| GUJRT  | Jun   | 1279 | 941.213  | 5682 | 8   | 73.590 | 1.322 | 6.302     |
|       | Jul   | 3433 | 1567.91  | 8625 | 126  | 45.672 | 0.59  | 3.27      |
|       | Aug   | 2441 | 1311.06  | 6385 | 101  | 53.710 | 0.533 | 2.738     |
|       | Sep   | 1435 | 1184.49  | 5800 | 25   | 82.560 | 0.979 | 3.532     |
| EMPRA  | Jun   | 1559 | 857.622  | 3863 | 262  | 55.025 | 0.667 | 2.790     |
|       | Jul   | 3842 | 1082.27  | 6716 | 869  | 28.169 | 0.113 | 2.615     |
|       | Aug   | 3715 | 1041.44  | 7424 | 1779 | 28.033 | 0.607 | 3.776     |
|       | Sep   | 1977 | 997.997  | 5627 | 177  | 50.545 | 0.856 | 4.122     |
| MADMH  | Jun   | 1261 | 515.891  | 2642 | 233  | 40.905 | 0.457 | 2.995     |
|       | Jul   | 1815 | 614.799  | 3201 | 409  | 33.873 | -0.025 | 2.507     |
|       | Aug   | 1315 | 515.319  | 3198 | 261  | 39.191 | 0.296 | 3.289     |
|       | Sep   | 1403 | 621.782  | 2977 | 212  | 44.312 | 0.208 | 2.396     |
| COAPR  | Jun   | 893  | 433.041  | 2528 | 252  | 48.509 | 1.222 | 4.714     |
|       | Jul   | 1334 | 459.005  | 3155 | 477  | 34.408 | 0.764 | 4.169     |
|       | Aug   | 1343 | 468.634  | 2920 | 285  | 34.895 | 0.486 | 3.317     |
|       | Sep   | 1556 | 623.493  | 3910 | 587  | 40.070 | 0.934 | 4.119     |
| TLNAD  | Jun   | 464  | 222.054  | 1838 | 141  | 47.836 | 2.289 | 13.644    |
|       | Jul   | 622  | 317.790  | 1873 | 162  | 51.116 | 1.500 | 5.357     |
|       | Aug   | 903  | 402.817  | 2491 | 282  | 44.599 | 0.885 | 3.977     |
|       | Sep   | 1100 | 412.480  | 2173 | 267  | 37.498 | 0.311 | 2.572     |
| ERJST  | Jun   | 743  | 507.121  | 2275 | 30   | 68.299 | 0.847 | 3.063     |
|       | Jul   | 2257 | 837.666  | 4287 | 213  | 37.114 | 0.154 | 2.814     |
|       | Aug   | 2223 | 978.826  | 4652 | 58   | 44.032 | 0.233 | 2.532     |
|       | Sep   | 1026 | 747.785  | 3442 | 25   | 72.855 | 0.939 | 3.723     |

*SD: Standard deviation and CV: Coefficient of variation

4.3. A wavelet - based test for testing the trend

A test statistics to test the null hypothesis \( H_0: T = 0 \) against the alternative \( H_1: T \neq 0 \) (Almasri et al., 2010) is defined as follows:

\[
\hat{\sigma}^2 = \frac{1}{N} \sum_{t=0}^{N-1} (X_t - \bar{X})^2 = \frac{1}{N} \| D \|^2 - \bar{X}^2
\]

\[
G = \frac{\hat{\sigma}^2_j}{\sum_{j=1}^{J} \hat{\sigma}^2_j \kappa_j (\lambda_j)} = \frac{1}{N} \sum_{j=1}^{J} \| D_j \|^2 + \frac{1}{N} \| S_j \|^2 - \bar{X}^2
\]
\[ \hat{\sigma}^2 = \sum_{j=1}^{l} \hat{\sigma}^2_j (\lambda_j) + \hat{\sigma}^2 \]

\( \hat{\sigma}^2_j (\lambda_j) \) is the estimated variance of the wavelet coefficients at scale \( \lambda_j \) and \( \hat{\sigma}^2 \) is the estimated variance of the trend. The test statistic \( (N - N/2^j)/(N/2^j - 1)G \) will follow an F-distribution with \((N/2^j - 1)\) and \((N - N/2^j)\) degrees of freedom (under the normality assumption of the scaling coefficients). The null hypothesis is rejected if the calculated value is greater than the tabulated F-value.

### 5. Empirical Illustrations

#### 5.1. Dataset

In this study monthly rainfall data corresponding to different zones of India is collected from Indian Institute of Tropical Meteorology (www.tropmet.res.in), Pune, India for detection as well as estimation of trend in monsoon rainfall data of different sub-divisions of India. The data set comprises of monthly rainfall over 128 years (from 1887 to 2014) measured in mm. There are total 30 sub-divisions in India, out of which 10 zones are considered in this present study. These 10 zones are selected in such a manner that it can represent the whole India. Here we are considering only monsoon months (namely June, July, August and September). The names of the selected sub-divisions along with their short names are listed below:

(i) Assam & Meghalaya sub-divisions (ASMEG)
(ii) Gangetic West Bengal sub-divisions (GNWBL)
(iii) West Uttar Pradesh Plains sub-divisions (WUPPL)
(iv) Punjab sub-divisions (PUNJB)
(v) East Rajasthan sub-divisions (ERJST)
(vi) East Madhya Pradesh sub-divisions (EMPRA)
(vii) Gujarat Sub-divisions (GUJRT)
(viii) Madhya Maharashtra sub-divisions (MADMH)
(ix) Coastal Andhra Pradesh sub-divisions (COAPR)
(x) Tamilnadu and Pondichery sub-divisions (TLNAD)

#### 5.2. Descriptive statistics

The descriptive statistics of the rainfall data is reported in Table 1. Here, we consider mean, maximum value, minimum value, standard deviation (SD), coefficient of variation (CV), skewness and kurtosis to represent the data set. The values of these statistics are calculated for each of the monsoon months corresponding to each of the sub-divisions under consideration. A perusal of Table 1 indicates that the average monthly monsoon rainfall is higher for Assam & Meghalaya (ASMEG) and Gangetic West Bengal (GNWBL) sub-divisions than the other sub-divisions. The table also indicates that the maximum rainfall is higher for Assam & Meghalaya (ASMEG), Gangetic West Bengal (GNWBL), Gujarat (GUJRT) and Madhya Maharashtra (MADMH) sub-divisions for each of the monsoon months. In terms of CV the variability in the monthly monsoon rainfall data is higher for West Uttar Pradesh Plains (WUPPL), Punjab (PUNJB) and Gujarat (GUJRT) sub-divisions. Almost all the series under consideration are positively skewed (except July month of WUPPL and MADMH) and all the series are leptokurtic.

#### 5.3. Test for long memory

The presence of long memory in a time series data can be confirmed either by investigating the autocorrelation function (ACF) plot of the data set or by using some parametric and nonparametric tests like GPH (Geweke and Porter-Hudak, 1983) and
Figs. 1(a-j). DWT plot of different zones using Haar wavelet filter
PAUL et al.: WAVELETS BASED ESTIMATION OF TREND IN SUB-DIVISIONAL RAINFALL

| 12000 | 6000 | 0 | -6000 | -12000 |
|-------|------|---|-------|--------|
| 1.1   |      |   |       |        |

| 6000  | 0    | -6000 | -12000 |
|-------|------|-------|--------|
| 1.3   |      |       |        |

| 4000  | 0    | -4000 | -8000  |
|-------|------|-------|--------|
| 2.1   |      |       |        |

| 8000  | 0    | -4000 | -8000  |
|-------|------|-------|--------|
| 2.3   |      |       |        |

| 2000  | 0    | -2000 | -4000  |
|-------|------|-------|--------|
| 3.1   |      |       |        |

| 4000  | 0    | -4000 | -8000  |
|-------|------|-------|--------|
| 3.3   |      |       |        |

| 4000  | 0    | -4000 | -8000  |
|-------|------|-------|--------|
| 4.1   |      |       |        |

| 4000  | 0    | -4000 | -8000  |
|-------|------|-------|--------|
| 4.3   |      |       |        |

| 2000  | 0    | -2000 |       |
|-------|------|-------|--------|
| 5.1   |      |       |        |

| 4000  | 0    | -4000 | -8000  |
|-------|------|-------|--------|
| 5.3   |      |       |        |

| 4000  | 0    | -4000 | -8000  |
|-------|------|-------|--------|
| 6.1   |      |       |        |

| 6000  | 0    | -6000 | -12000 |
|-------|------|-------|--------|
| 6.3   |      |       |        |

| 2000  | 0    | -2000 | -4000  |
|-------|------|-------|--------|
| 7.1   |      |       |        |

| 4000  | 0    | -4000 | -8000  |
|-------|------|-------|--------|
| 7.3   |      |       |        |
Sperio test (Reisen, 1994). In this study we have applied Sperio test to test the presence of long memory for each of the monsoon month rainfall data of the selected zones and the test results are provided in Table 2. It is found that the test is significant at 5% level for some of the series which establishes the presence of long range dependency in the respective monthly rainfall data like July (ASMEG), August (WUPPL) etc.

5.4. Variation in rainfall using DWT

DWT plots of rainfall in different zones for the months of June, July, August and September using Haar wavelet filter at level 5 are reported in Figs. 1(a-j). The number of the plots in Figs. 1(a-j) are given for the different zones accordingly to the order as per section 5.1. Wavelet coefficients are plotted as bars, up or down. The sizes of the bars are relative to magnitudes of coefficients. The number of wavelet coefficients at the lowest resolution level (level = 1) is exactly half the number of original data points and the number of coefficients decreases by half at each level (Nason and Sachs, 1999). The plot of actual data is shown at the bottom of each plot of Figs. 1(a-j) which is very rough and it is full of noise component. DWT attempts to extract the actual signal from the noisy time series data at each level. The graph of $W_2$ is much smoother than the $W_1$. Similarly smoothness increases as we are going to top of the graphs with upper coefficients. In the graph $V_2$ scaling coefficient shows the smooth plot.

5.5. Detection of trend using wavelet

Trends of different sub-divisional zones are estimated using “Haar” DWT and plotted in the Fig. 2. Each plot in Fig. 2 is represented by two subscripts. The first subscript refers to zone as described in section 5.1 and the second subscript refers to the monsoon months, i.e., 1 for June; 2 for July; 3 for August and 4 for September. Therefore, the first plot, i.e., 1.1 represents the trend in rainfall in ASMEG zone for the month June. The rainfall trend in the month of June for ASMEG zone clearly indicates that trend is negative. From the year 1887 to 1915 trend remains constant then it falls during 1916. Again it remains stagnant during the year 1918 to 1949. After that a slight increase in the rainfall has been detected till the year 1951. A rapid decreasing trend has been observed from 1997 to 2014. For other three months of this zone, the rainfall pattern remains same with a difference in magnitude. The trend in other zones can be interpreted in similar pattern. The overall pattern of trend in rainfall remains almost same in all the zones except the variation in different time epochs.

5.6. Test for detection of trend

In this section we test the existence of trend using Mann-Kendall test. The values of test statistics along with corresponding p-values are given in Table 3. This table also includes the values of G statistic along with their
**TABLE 3**

Test for detection of trend

| Zone  | Months | Method       |       |       |       |       |
|-------|--------|--------------|-------|-------|-------|-------|
|       |        |              | Haar Filter | Mann-Kendall |
|       |        | G  | F     | S    | tau | p-value |
| ASMEG | Jun    | 0.171 | 7.068* | -1121 | -0.138* | 0.021 |
|       | Jul    | 0.046 | 1.901 | -229  | -0.028 | 0.630 |
|       | Aug    | 0.043 | 1.777 | -1316 | -0.162* | 0.006 |
|       | Sep    | 0.034 | 1.405 | -520  | -0.064 | 0.285 |
| GNWBL | Jun    | 0.494 | 20.419* | -505  | -0.062 | 0.299 |
|       | Jul    | 0.232 | 9.589* | 8     | 0.001  | 0.989 |
|       | Aug    | 0.474 | 19.592* | 569  | 0.070 | 0.242 |
|       | Sep    | 0.615 | 25.420* | 1133 | 0.139* | 0.019 |
| WUPPL | Jun    | 0.127 | 5.249* | 338  | 0.041  | 0.487 |
|       | Jul    | 0.080 | 3.307* | -235 | -0.028 | 0.629 |
|       | Aug    | 0.345 | 14.260* | -407 | -0.050 | 0.403 |
|       | Sep    | 0.018 | 0.744 | -167  | -0.020 | 0.732 |
| PUNJB | Jun    | 0.037 | 1.529 | -111  | -0.013 | 0.820 |
|       | Jul    | 0.510 | 21.080* | 270  | 0.033  | 0.579 |
|       | Aug    | 0.277 | 11.449* | -28  | -0.003 | 0.950 |
|       | Sep    | 0.102 | 4.216* | 612  | 0.075  | 0.208 |
| GJRT  | Jun    | 0.037 | 1.529 | -111  | -0.013 | 0.820 |
|       | Jul    | 0.272 | 11.243* | -190 | -0.023 | 0.690 |
|       | Aug    | 0.101 | 4.133* | 585  | 0.072  | 0.220 |
|       | Sep    | 0.134 | 5.539* | 372  | 0.045  | 0.440 |
| EMPRA | Jun    | 0.199 | 8.225* | -305  | -0.030 | 0.530 |
|       | Jul    | 1.15  | 47.533* | -879 | -0.108 | 0.070 |
|       | Aug    | 0.272 | 11.243* | -612 | -0.075 | 0.208 |
|       | Sep    | 0.025 | 1.033 | -517  | -0.060 | 0.287 |
| MADMH | Jun    | 0.475 | 19.633* | 722  | 0.088  | 0.137 |
|       | Jul    | 0.17  | 7.027* | -568  | -0.069 | 0.242 |
|       | Aug    | 0.69  | 28.520* | 1145 | 0.141* | 0.010 |
|       | Sep    | 0.112 | 4.629* | 48   | 0.005  | 0.922 |
| COAPR | Jun    | 0.195 | 8.060* | -32   | -0.003 | 0.949 |
|       | Jul    | 0.259 | 10.705* | 446  | 0.054  | 0.359 |
|       | Aug    | 0.441 | 18.228* | 858  | 0.106  | 0.077 |
|       | Sep    | 0.021 | 0.868 | -151  | -0.018 | 0.757 |
| TLNAD | Jun    | 0.054 | 2.232 | -468  | -0.057 | 0.336 |
|       | Jul    | 0.26  | 10.747* | 403  | 0.049  | 0.407 |
|       | Aug    | 0.103 | 4.257* | 29   | 0.003  | 0.954 |
|       | Sep    | 0.152 | 6.283* | -141 | -0.017 | 0.773 |
| ERJST | Jun    | 0.056 | 2.315 | 322   | 0.039  | 0.508 |
|       | Jul    | 0.566 | 23.395* | -34  | -0.004 | 0.940 |
|       | Aug    | 0.063 | 2.604 | 94   | 0.011  | 0.840 |
|       | Sep    | 0.075 | 3.100* | 178  | 0.021  | 0.710 |

*denotes significance at 5% level

corresponding F-values for testing the significance of existing trend as described in section 4.3. It has been seen that the Mann-Kendal test could not capture the significant trend in many of the series whereas the Wavelet based test successfully captured the trend in almost all the months of different zones. This is because, wavelet based test has more power than that of the Mann-Kendal test in capturing the trend particularly when there is long memory present in the series.
6. Conclusions

In the present study, two nonparametric methods namely wavelet analysis and Mann-Kendall test have been used for estimation of trend in monthly rainfall in ten selected sub-divisions of India. Presence of long memory in rainfall data was tested using suitable statistical test and it was found that long memory was significantly present in rainfall for the month of June in COAPR; July of ASMEG, PUNJB and EMPRA; August of GUJRT and EMPRA; September of EMPRA and COAPR. In presence of long memory, Mann-Kendall test may not be appropriate to detect the trend. Therefore, wavelet based trend test was advocated and it was found that in most of the sub-divisions, the rainfall trend was significant at 5% level. Mostly, significant negative trend in rainfall has been found in the monsoon months which is not a positive signal for the water sector in the country. The variation in monthly rainfall for June-September for the studied zones were also investigated by wavelet using Haar filter. The local as well as global variation in rainfall is observed by DWT plot of the respective months in the respective locations. The variability in rainfall is evident in the recent decades in all the locations.

Acknowledgements

Authors are thankful to the anonymous reviewers for providing useful comments which helped improve the paper. The contents and views expressed in this research paper are the views of the authors and do not necessarily reflect the views of the organizations they belong to.

References

Almasri, A., 2010, “Tests for Trend: A Simulation Study”, Communications in Statistics-Simulation and Computation, 39, 3, 598-611.

Antoniadis, A., 1997, “Wavelets in statistics: A review”, Journal of the Italian Statistical Society, 6, 291-304.

Box, G. E. P., Jenkins, G. M. and Reinsel, G. C., 2007, “Time-series analysis: Forecasting and control”, 3rd edition, Pearson Education, India.

Daubechies, I., 1992, “Ten Lectures on Wavelets”, SIAM, Philadelphia.

Geweke, J. and Porter-Hudak, S., 1983, “The estimation and application of long memory time series models”, Journal of Time Series Analysis, 4, 4, 221-238.

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

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

Jain, S. K. and Kumar, V., 2012, “Trend analysis of rainfall and temperature data for India”, Current Science, 102, 37-49.

Jayawardene, H. K. W. I., Sonnadara, D. U. J. and Jayawardene, D. R., 2005, “Trends of rainfall in Sri-Lanka over the last century, Sri Lankan Journal of Physics, 6, 7-17.

Kallache, M., Rust, H. W. and Kropp, J., 2005, “Trend assessment: Applications for hydrology and climate research”, Nonlinear Processes in Geophysics, 12, 201-210.

Kumar, K. K., Kumar, K. R., Ashrit, R. G., Deshpande, N. R. and Hansen, J. W., 2004, “Climate impacts on Indian agriculture”, International Journal of Climatology, 24, 1375-1393.

Nason, G. P. and Von Sachs, R., 1999, “Wavelets in time series analysis”, Philosophical Transaction of the Royal Society London A, 357, 2511-2526.

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

Paul, R. K. and Anjoy, P., 2018, “Modeling fractionally integrated maximum temperature series in India in presence of structural break”, Theoretical and Applied Climatology, 134, 1-2, 241-249.

Paul, R. K. and Birthal, P. S., 2016, “Investigating rainfall trend over India using the wavelet technique”, Journal of Water and Climate Change, 7, 2, 353-364.

Paul, R. K., 2017, “Modelling long memory in maximum and minimum temperature series in India”, Mausam, 68, 2, 317-326.

Paul, R. K., Birthal, P. S. and Khokhar, A., 2014, “Structural Breaks in Mean Temperature over Agro-climatic Zones in India”, The Scientific World Journal, 434325, doi:dx.doi.org/10.1155/2014/434325.

Paul, R. K., Birthal, P. S., Paul, A. K. and Gurung, B., 2015, “Temperature trend in different agro-climatic zones in India”, Mausam, 66, 4, 841-846.

Paul, R. K., Prajneshu and Ghosh, H., 2011, “Wavelet methodology for estimation of trend in Indian monsoon rainfall time-series data”, Indian Journal of Agricultural Sciences, 81, 3, 96-98.

Paul, R. K., Prajneshu and Ghosh, H., 2013, “Wavelet Frequency Domain Approach for Modelling and Forecasting of Indian Monsoon Rainfall Time-Series Data”, Journal of the Indian Society of Agricultural Statistics, 67, 3, 319-327.

Percival, D. B. and Walden, A. T., 2000, “Wavelet methods for time series analysis”, Cambridge University Press, U. K.

Reisen, V. A., 1994, “Estimation of the fractional difference parameter in the ARFIMA (p,d,q) model using the smoothed periodogram”, Journal Time Series Analysis, 15, 1, 335-350.

Vidakovic, B., 1999, “Statistical Modeling by Wavelets”, John Wiley, New York.

Yajima, Y., 1988, “On estimation of regression model with long memory stationary errors”, Annals of Statistics, 16, 791-807.