Modelling long memory in maximum and minimum temperature series in India

RANJIT KUMAR PAUL

ICAR-Indian Agricultural Statistics Research Institute, New Delhi – 110 012, India

e mail : ranjitstat@gmail.com; ranjitstat@iasri.res.in

सार – मौसम आँकड़ों के समय श्रृंखला विश्लेषण उनके परिवर्तिता पैटर्न की जाँच करने तथा, समय श्रृंखला में लघु और दीर्घ काल के परिवर्तनों के पूर्वानुमान लगाने में बहुत महत्वपूर्ण माध्यम हो सकते हैं। इस अध्ययन में फ्रैक्शनल एकीकृत तकनीकों के माध्यम से 1901 से 2007 तक की अवधि के भारत के मासिक न्यूनतम और अधिकतम तापमान के दीर्घ मैमोरी व्यवहार की जाँच की गई है। परिणाम दर्शाते हैं कि समय श्रृंखला को स्वसमाश्रयी फ्रैक्शनली एकीकृत मूविंग एवरेज (ए आर एफ आई एम ए) प्रक्रिया के अनुसार विशेषीकृत किया जा सकता है। दोनों श्रृंखलाएं दीर्घ मैमोरी स्टेशनैरिटि सुनिश्चित करते हुए 0.5 से कम एकीकरण के अनुसार एकीकृत पाई गई। श्रृंखला के विभिन्न पैमानों पर तथा विभिन्न समय कालवधि में दोलन को देखने के लिए हार वेवलेट फिल्टर के साथ आवृति डोमेन में वेवलेट पद्धति का अनुप्रयोग किया गया। पिछले कुछ वर्षों में तापमान की दोनों श्रृंखलाओं में स्थानीय तथा वैश्विक परिवर्तिता को जानने के लिए बहुविभेदन विश्लेषण (एम आर ए) किए गए। न्यूनतम तापमान में परिवर्तिता अधिकतम तापमान से अधिक पाई गई। हालांकि दीर्घअवधि में तापमान श्रृंखलाओं में कोई स्पष्ट महत्वपूर्ण प्रवृत्ति नहीं है तथापि तापमान के पैटर्न में परिवर्तन के संकेत हैं। ए आर एफ आई एम ए मॉडल की पूर्वानुमान क्षमता की जाँच सापेक्ष माध्य निरपेक्ष प्रतिशत त्रुटि के संबंध में की गई।

ABSTRACT. Time series analysis of weather data can be a very valuable tool to investigate its variability pattern and, maybe, even to predict short- and long-term changes in the time series. In this study, the long memory behaviour of monthly minimum and maximum temperature of India for the period 1901 to 2007 by means of fractional integration techniques has been investigated. The results show that the time series can be specified in terms of autoregressive fractionally integrated moving average (ARFIMA) process. Both the series were found to be integrated with orders of integration smaller than 0.5 ensuring the long memory stationarity. Wavelet methodology in frequency domain with Haar wavelet filter was applied in order to see the oscillation at different scale and at different time epochs of the series. Multiresolution analysis (MRA) was carried out to explore the local as well as global variations in both the temperature series over the years. The variability in minimum temperature is found to be more than maximum temperature. Though there is no clear significance trend in the temperature series in the long run, but there are pockets of change in the temperature pattern. The predictive ability of ARFIMA model was investigated in terms of relative mean absolute percentage error.

Key words – Long memory, Stationarity, Validation, Wavelet.

1. Introduction

Large number of research papers have been published on long memory and fractionally integrated processes since the initial publication of the work of Granger (1980); Granger and Joyeux (1980) and Hosking (1981) which parameterized the processes of Hurst (1951) on the time series with hyperbolically decaying autocorrelations. The long memory or long term dependence property describes the high-order correlation structure of a time series. If a series exhibits long memory, there is persistent temporal dependence even between distant observations. Such series are characterized by distinct but non-periodic cyclical patterns. The presence of long memory dynamics causes nonlinear dependence in the first moment of the distribution and hence a potentially predictable component in the series dynamics. Fractionally integrated processes can give rise to long memory (Beran, 1994).

A popular class of models for time series with long memory behaviour is the autoregressive fractionally integrated moving average (ARFIMA) model by Granger and Joyeux (1980). This kind of models extended classical

ARIMA models by assuming the differencing parameter *d* as a fractional value. It is well known that ARFIMA models are linear time series model. Fractional integration is part of the larger classification of time series, commonly referred to as 'long memory' models. The recent empirical evidence suggests that temperature series may be well described in terms of fractionally integrated processes (Gil-Alana, 2004). Fractionally integrated I(d) processes have attracted growing attention among empirical researchers. In fact this is because I(d) processes provide an extension to the classical dichotomy of I(0) and I(1) time series and equip us with more general alternatives in long range dependence (Shimotsu, 2010). Empirical research continues to find evidence that I(d) processes can provide a suitable description of certain long range characteristics.

Understanding the nature and scale of possible climate changes in India is of importance to the policy makers and farmers as it gives them a chance to be prepared for better mitigation and adaptation measures. For that purpose time series analysis of weather data can be a very valuable tool to investigate its variability pattern and, maybe, even to predict short- and long-term changes in the time series. Various researchers have carried out studies on temperatures. Woodcock (1984) described some experimental MOS forecasts of daily maximum and minimum temperature for seven Australian cities. Raj (1998) evolved a scheme for predicting minimum temperature at Pune by analogue and regression methods. Mohan *et al.* (1989) developed a method for forecasting maximum temperature over Ozar situated in Maharastra using maximum and dew point temperature of the previous day. In the present investigation monthly minimum temperatures in India, for the period 1901–2007 were examined by means of fractional integration techniques. Dhimri *et al.* (2005) have carried out forecast of minimum temperature at Manali, India. Paul *et al.* (2014) have investigated the trend in mean surface temperature in different agro-climatic zones in India. Paul *et al.* (2015) have also investigated the structural break in mean surface temperature in different agro-climatic zones in India and reported that there is significant structural break during 1970's. However, in none of the above studies, long memory nature in maximum and minimum surface temperature in India has been investigated. In the present investigation, an attempt has been made to apply long memory model for forecasting maximum and minimum surface temperature in India with more accuracy. To this end, nonparametric wavelet technique has also been applied to study the pattern of maximum and minimum surface temperature in India over the last century or so both globally as well as locally. Some applications of this technique in modelling climate variables may be found in Paul *et al.* (2013),

Paul *et al.* (2011); Paul and Birthal (2015). The paper is organized as follows: section 2 describes the data set used in the present investigation; section 3 describes the long memory definition, ARFIMA model, testing stationarity and testing of presence of long memory followed by section 4 which deals with results and discussion.

2. Data

For the present investigation, all India monthly maximum and minimum temperature data during the period January, 1901 to December, 2007 is used. The data is collected form Indian Institute of Tropical Meteorology, Government of India. The data for the period January 1901 to December, 2006 have been used for model building and the remaining data have been used for model validation purpose.

3. Methodology

3.1. *Long memory process*

Long memory in time-series can be defined as autocorrelation at long lags (Robinson, 1995). According to Jin and Frechette (2004), memory means that observations are not independent (each observation is affected by the events that preceded it). The autocorrelation function (acf) of a time-series y_t is defined as :

$$\rho_K = \text{cov}(y_t, y_{t-1}) / \text{var}(y_t) \qquad (1)$$

for integer lag *k*. A covariance stationary time-series process is expected to have autocorrelations such that $\lim_{k \to \infty} \rho_k = 0$. Most of the well-known class of stationary and invertible time-series processes have autocorrelations that decay at the relatively faster exponential rate, so that $\rho_k \approx |m|^k$, where |m|<1 and this property is true, for example, for the well-known stationary and invertible ARMA(*p*, *q*) process. For long memory processes, the autocorrelations decay at an hyperbolic rate which is consistent with $\rho_k \approx Ck^{2d-1}$, as *k* increases without limit, where C is a constant and *d* is the long memory parameter.

3.2. *ARFIMA Model*

Fractional integration is the primary conceptual framework for describing long memory in financial time-series. Fractional integration is a generalization of integer integration, under which time-series are usually presumed to be integrated of order zero or one. For example, an

autoregressive moving-average process integrated of order d [denoted by ARFIMA(p, d, q)] can be represented as :

$$(1-L)^d \, \varphi(L) y_t = \theta(L) u_t \qquad (2)$$

where, u_t is an independently and identically distributed (i.i.d.) random variable with zero mean and constant variance, L denotes the lag operator; and $\varphi(L)$ and $\theta(L)$ denote finite polynomials in the lag operator with roots outside the unit circle. For $d = 0$, the process is stationary, and the effect of a shock to u (t) on y ($t + j$) decays geometrically as j increases. For $d = 1$, the process is said to have a unit root, and the effect of a shock to $u(t)$ on y ($t + j$) persists into the infinite future. In contrast, fractional integration defines the function $(1 - L)^{-d}$ for non-integer values of the fractional differencing parameter d. It turns out that for $-0.5 < d < 0.5$ the process $y(t)$ is stationary and invertible. A detail description of ARFIMA model can be found in Robinson (2003).

3.3. *Estimation procedures*

We deal with some well known estimation methods of the long memory parameter d. The first one is the semi parametric method based on an approximated regression equation obtained from the logarithm of the spectral density function of a model. This method is proposed by Geweke and Porter-Hudak (1983). The second is the Gaussian semi parametric method developed by Robinson (1995).

3.4. *Testing of Long Memory*

$H = 1 - d$, a Hurst exponent produced by the rescaled range analysis, or R/S, analysis and applied to economic price analysis by Booth *et al*. (1982) and Helms *et al*. (1984). For a given time-series, the Hurst exponent measures the long-term non-periodic dependence, and indicates the average duration the dependence may last.

The time period spanned by the time series of length T is divided into m contiguous sub-periods of length n such that $m*n = T$. In each sub-period X_{ij}, the elements have two subscripts. The first subscript ($i = 1,\dots,n$) denotes the number of elements in each sub-period and the second one ($j = 1,\dots,m$) denotes the sub-period index. For each sub-period j the R/S statistic is calculated as follows:

$$(R/S)_j = \left(S_j\right)^{-1} \max_{1 \le k \le n} \sum_{i=1}^{k} \left(x_{ij} - \bar{x}_j\right) - \min_{1 \le k \le n} \sum_{i=1}^{k} \left(x_{ij} - \bar{x}_j\right)$$

$$(3)$$

where, S_j is the standard deviation for each sub-period. In (3), the k deviations from the sub-period mean

have zero mean; therefore the last value of the cumulative deviations for each sub-period will always be zero. Because of this, the maximum value of the cumulative deviations will always be greater or equal to zero, while the minimum value will always be less or equal to zero. Rescaling the range is crucial since it allows diverse phenomena and time periods to be compared, which means that R/S analysis can describe time series with no characteristic scale. The $(R/S)_n$ is computed by the average of the $(R/S)_j$ values for all the m contiguous sub-periods with length n as

$$(R/S)_n = m^{-1} (R/S)_j \qquad (4)$$

Eq. (4) computes the R/S value which corresponds to a certain time interval of length n. This is repeated by increasing n to the next integer value, until $n = T/2$, since at least two sub-periods are needed, to avoid bias. As n increases**,** the following holds:

$$\log\left[(R/S)_n\right] = \log \alpha + H \log n \qquad (5)$$

When $0.5 < H < 1$, the long memory structure exists. If $H \ge 1$, the process has infinite variance and is nonstationary. If $0 < H < 0.5$, anti-persistence structure exists. If $H = 0.5$, the process is white noise.

3.5. *Geweke and Porter-Hudak (GPH) estimate*

The GPH estimation procedure is a two-step procedure, which begins with the estimation of d. This method is based on least squares regression in the spectral domain, exploits the sample form of the pole of the spectral density at the origin: $f_x(\lambda) \sim \lambda^{-2d}, \quad \lambda \to 0$. To illustrate this method, we can write the spectral density function of a stationary model X_t, $t = 1, \dots, T$ as:

$$f_x(\lambda) = \left[4\sin^2\left(\frac{\lambda}{2}\right) \right]^{-d} f_\varepsilon(\lambda)$$

where, $f_\varepsilon(\lambda)$ is the spectral density of ε_t, assumed to be a finite and continuous function on the interval $\left[-\pi, \pi\right]$. Taking the logarithm of the spectral density function $f_x(\lambda)$ the log-spectral density can be expressed as :

$$\log[f_x(\lambda)] = \log[f_\varepsilon(0)] - d \log\left[4\sin^2\left(\frac{\lambda}{2}\right) \right] + \log\frac{f_\varepsilon(\lambda)}{f_\varepsilon(0)}$$

Let, $I_x(\lambda_j)$ be the periodogram evaluated at the Fourier frequencies $\lambda_j = 2\pi_j / T$; $j = 1, 2, \dots, m$; T is the
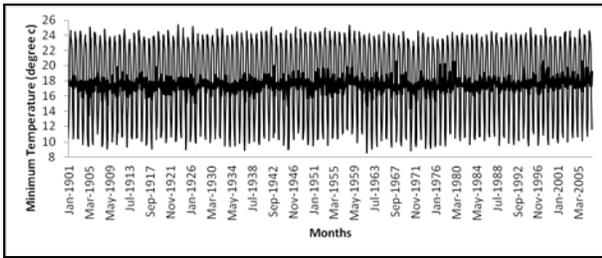
Fig. 1. All India minimum temperature in degree centigrade (dark line is seasonally adjusted series and light line is average monthly minimum temperature)
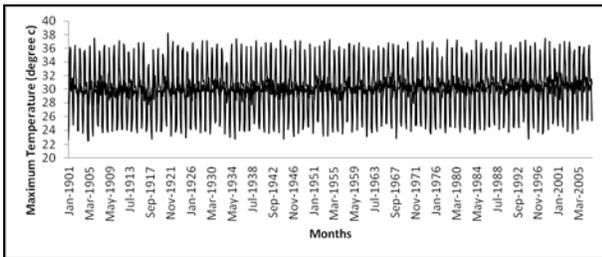
Fig. 2. All India maximum temperature in degree centigrade (dark line is seasonally adjusted series and light line is average monthly minimum temperature)

number of observations and *m* is the number of considered Fourier frequencies, that is the number of periodogram ordinates which will be used in the regression:

$$\log\left[I_x\left(\lambda_j\right)\right] = \log\left[f_\varepsilon(0)\right] - d\log\left[4\sin^2\left(\frac{\lambda}{2}\right)\right]$$
$$+ \log\frac{f_\varepsilon(\lambda)}{f_\varepsilon(0)} + \log\frac{I_x\left(\lambda_j\right)}{f_x\left(\lambda_j\right)}$$

where, $\log\left[f_\varepsilon(0)\right]$ is a constant, $\log[4\sin^2(\lambda/2)]$ is the exogenous variable and $\log[I_x(\lambda_j)/f_x(\lambda_j)]$ is a disturbance error. The GPH estimate requires two major assumptions related to asymptotic behaviour of the equation:

H_1: for low frequencies, we suppose that $\log\left[f_\varepsilon(\lambda)/f_\varepsilon(0)\right]$ is negligible.

H_2: the random variables $\log\left[I_x\left(\lambda_j\right)/f_x\left(\lambda_j\right)\right]$; $j = 1, 2, \ldots, m$ are asymptotically iid.

Under the hypotheses H_1 and H_2, we can write the linear regression

$$\log\left[I_x\left(\lambda_j\right)\right] = \alpha - d\log\left[4\sin^2\left(\frac{\lambda_j}{2}\right)\right] + e_j$$

where, $e_j \sim iid\left(-c, \pi^2/6\right)$.

Let $Y_j = -\log\left[4\sin^2(\lambda_j/2)\right]$ the GPH estimator is the OLS estimate of the regression log $I_x\left(\lambda_j\right)$ on the constant α and y_j. The estimate of *d* is

$$\hat{d}_{GPH} = \frac{\sum_{j=1}^{m}\left(y_j - \overline{y}\right)\log\left[I_x\left(\lambda_j\right)\right]}{\sum_{j=1}^{m}\left(y_j - \overline{y}\right)^2}, \text{ where } \overline{y} = \sum_{j=1}^{m}y_j/m$$

The parameter *m* is selected so that $m = T^u$, with $u = 0.5; 0.6; 0.7$. Robinson (1995), Hurvich *et al.* (1998) and Tanaka (1999) have analyzed the GPH estimate in detail. Under the assumption of normality for X_t, it has been proved that the estimate is consistent and asymptotically normal. An alternative semiparametric estimator has been proposed by Robinson (1995).

3.6. *Wavelets*

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 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 $\psi(.)$ is a real-valued function defined over the real axis $(-\infty, \infty)$ and satisfies two basic properties: (*i*) Integral of $\psi(.)$ is zero, *i.e.*, $\int_{-\infty}^{\infty}\psi(u)\,du = 0$ (*ii*) Square of $\psi(.)$ integrates to unity, *i.e.*, $\int_{-\infty}^{\infty}\psi^2(u)\,du = 1$, then the function $\psi(.)$ is called a wave. A good description of wavelets can be found in Daubechies (1992); Ogden (1997) and Percival and Walden (2000).

3.7. *Maximal Overlap Discrete Wavelet Transforms (MODWT)*

The Maximal Overlap Discrete Wavelet Transforms (MODWT) is a linear filtering operation that transforms a series into coefficients related to variations over a set of scales. It is similar to 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 $j = 1, .., J_0$. Both

TABLE 1

Descriptive Statistics of maximum and minimum temperature

Descriptive statistics	Maximum temperature	Minimum temperature
Mean	30.24	18.33
Median	30.7	19.7
Maximum	38.2	25.4
Minimum	22.5	8.6
Std. Deviation	3.81	5.15
CV	12.60	28.10
Skewness	-0.12	-0.39
Kurtosis	2.13	1.6

TABLE 2

Testing stationarity of seasonally adjusted temperature series

Test	Test Statistic		1% critical value	5% critical value
	Max temp	Min temp		
Augmented Dickey Fuller (ADF)	-40.821	-12.174	-3.438	-2.864
Philips and Peron (PP)	-12.190	-27.966	-3.438	-2.864

are suitable for the analysis of variance (ANOVA) and for multiresolution analysis (MRA). However, MODWT differs from DWT in the sense that it is a highly redundant, nonorthogonal transform (Percival and Walden, 2000). It retains downsampled values at each level of the decomposition that would otherwise be discarded by DWT. The MODWT is well defined for all sample sizes N, whereas for a complete decomposition of J levels, DWT requires N to be a multiple of 2^J.

3.8. MODWT coefficients

For a redundant transform, like MODWT, an N sample input time-series will have an N sample resolution scale for each resolution level. Therefore, features of wavelet coefficients in a multiresolution analysis (MRA) will be lined up with 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:

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

$$\tilde{V}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{g}_{j,l} X_{t-l \bmod N} \tag{6}$$

where, $\tilde{h}_{j,l} \equiv h_{j,l}/2^{j/2}$ are j^{th} level MODWT wavelet filters, and $\tilde{g}_{j,l} \equiv g_{j,l}/2^{j/2}$ are j^{th} level MODWT scaling filters, L_j is width of jth level equivalent wavelet and scaling filters. For a time-series X with N samples, MODWT yields an additive decomposition or MRA given by:

$$X = \sum_{j=1}^{J_0} \tilde{D}_j + \tilde{S}_{J_0}, \tag{7}$$

where,

$$\tilde{D}_{j,t} = \sum_{l=0}^{N-1} \tilde{u}_{j,l}\ \tilde{W}_{j,t+l \bmod N} \text{ and}$$

$$\tilde{S}_{j,t} = \sum_{l=0}^{N-1} \tilde{v}_{j,l}\ \tilde{V}_{j,t+l \bmod N} \tag{8}$$

$\tilde{u}_{j,l}$ and $\tilde{v}_{j,l}$ being the filters obtained by periodizing $\tilde{h}_{j,l}$ and $\tilde{g}_{j,l}$. According to eq. (7), at a scale j, a set of coefficients $[D_j]$ each with the same number of samples (N) as in the original signal (X) is obtained. These are called wavelet "details" and capture local fluctuations over whole period of a time-series at each scale. Set of values S_{J_0} provide a "smooth" or overall "trend" of the original signal and adding D_j to S_{J_0}, for $j = 1, 2,..., J_0$, gives an increasingly more accurate approximation for it. This additive form of reconstruction allows prediction of each wavelet subseries (D_j, S_{J_0}) separately and adding individual predictions an aggregate forecast is generated.

3.9. Choosing number of levels

A time-series can be completely or partially decomposed into a number of levels. For complete decomposition of a series of length $N = 2^J$ using DWT, maximum number of levels in the decomposition is J. In practice, a partial decomposition of level $J_0 \le J$ suffices for many applications. A J_0 level DWT decomposition requires that N be an integral multiple of 2^{J_0}. The MODWT can accommodate any sample size N and in theory, any J_0. In practice, largest level is commonly
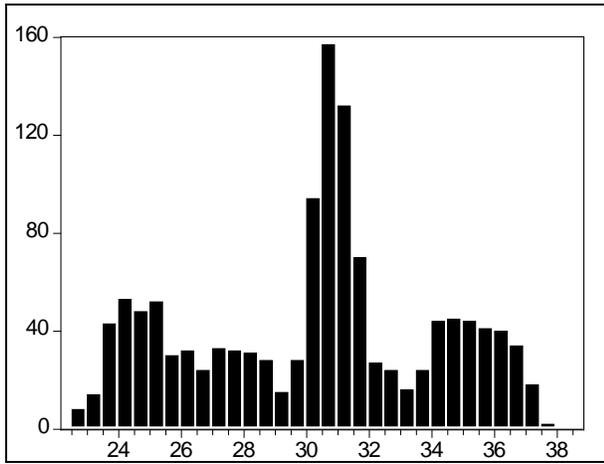
Fig. 3. Histogram of maximum temperature
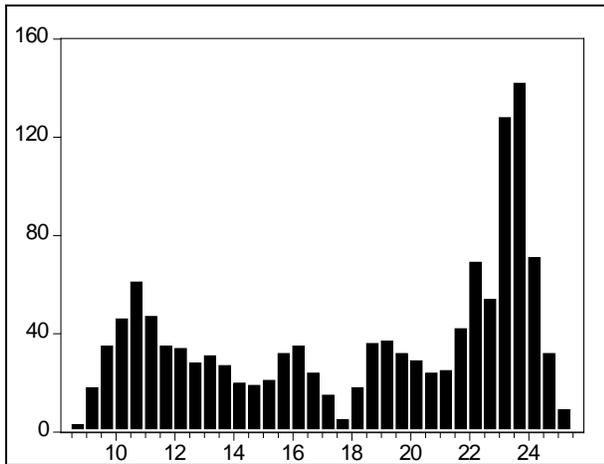
Fig. 4. Histogram of minimum temperature
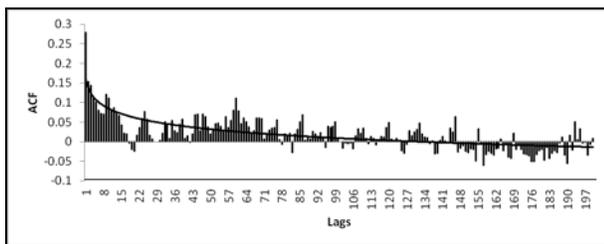
Fig. 5. Autocorrelation function (ACF) of seasonally adjusted minimum temperature series up to 200 lags
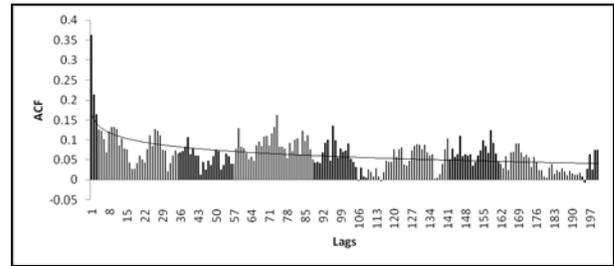
Fig. 6. Autocorrelation function (ACF) of seasonally adjusted maximum temperature series up to 200 lags
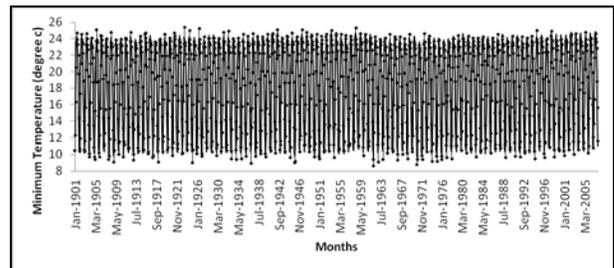
Fig. 7. Observed (marker) *vs* predicted (line) of minimum temperature
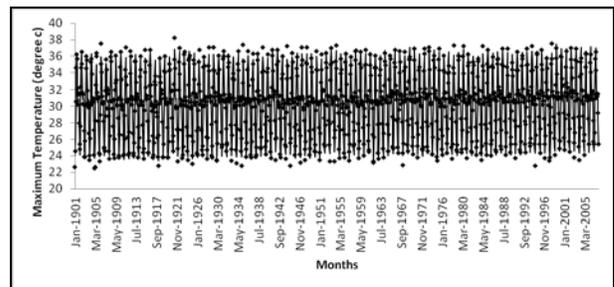
Fig. 8. Observed (marker) *vs* predicted (line) of maximum temperature

determines the number of octave bands and thus number of scales of resolution in the decomposition.

4.   Results and discussion

A perusal of the Figs. 1 and 2 indicate that both the series are stationary. In order to test for stationarity, two tests namely Augmented Dickey-Fuller unit root test (Said and Dickey, 1984) and Philips-Peron unit root test (Philips and Peron, 1988) are conducted. The descriptive statistics for monthly maximum and minimum temperature have been computed and are reported in Table 1. A perusal of Table 1 reveals, the variability in minimum temperature is more than maximum temperature. The same can also be observed from the histogram plotted in Figs. 3 and 4. The results of the stationarity tests are given in Table 2. A perusal of Table 2 reveals that both the test statistics reject the null hypothesis of presence of unit root indicating that series are stationary.

selected such that $J_0 \leq log_2\left(N\right)$ in order to preclude decomposition at scales longer than total length of the time-series. In particular, for alignment of wavelet coefficients with the original series, condition $L_{J_0} < N$ (*i.e.*, width of equivalent filter at J_0^{th} level is less than sample size) should be satisfied to prevent multiple wrappings of the time-series at level J_0. Selection of J_0

TABLE 3

Parameter estimate of ARFIMA (0,d,0)

Parameters	Maximum temperature			Minimum temperature		
	Estimate	Standard error	*P* Value	Estimate	Standard error	*P* Value
const	30.025	0.137	<0.001	17.543	0.093	<0.001
d	0.262	0.022	<0.001	0.217	0.021	<0.001
Log likelihood	406.956			452.398		
Akaike Information Criteria (AIC)	-807.912			-898.797		
Bayesian Information Criteria (BIC)	3038.09			2947.200		

TABLE 4

Validation of models

Temperature	MAE	RMSPE	RMAPE (%)
Maximum	0.46	0.54	1.51
Minimum	0.44	0.56	2.72

4.1. *Structure of autocorrelations*

For a linear time series model, typically an autoregressive integrated moving average [ARIMA(p,d,q)] process, the patterns of autocorrelations and partial autocorrelations could indicate the plausible structure of the model. At the same time, this kind of information is also important for modelling nonlinear dynamics. The long lasting autocorrelations of the data suggest that the processes are nonlinear with time-varying variances. The basic property of a long memory process is that the dependence between the two distant observations is still visible. For the series of daily wholesale price, autocorrelations were estimated up to 100 lags, *i.e*., $j = 1,...,100$. The autocorrelation functions of these series are plotted in Figs. 5 and 6. A perusal of figures indicate that, these do not decay exponentially over time span, rather, there is hyperbolic decay of the autocorrelations functions towards zero and they show no clear periodic patterns. There is no evidence that the magnitude of autocorrelations become small as the time lag *j*, becomes larger. No seasonal and other periodic cycles were observed.

Accordingly, ARFIMA model was fitted to the above seasonally adjusted dataset. The best ARFIMA model has been selected on the basis of minimum Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) values. It is found the both the maximum as well as minimum temperature series follow

ARFIMA (0, d, 0) process. The value of long memory parameter for maximum and minimum temperature are found to be 0.262 and 0.217 respectively. Both values are also significant at 1% level as reported in Table 3.

4.2. *Diagnostic checking and validation*

The model verification is concerned with checking the residuals of the model to see if they contained any systematic pattern which still could be removed to improve the chosen ARFIMA. For this purpose, autocorrelations of the residuals were computed and it was found that none of these autocorrelations was significantly different from zero at any reasonable level. This proved that the selected ARFIMA model was an appropriate model for forecasting the data under study.

One-step ahead forecasts of temperature series using naïve approach for the period January, 2007 to December, 2007 in respect of above fitted model are computed. For measuring the accuracy in fitted time series model, Root mean square prediction error (RMSPE), Mean absolute error (MAE) and Relative mean absolute prediction error (RMAPE) are computed by using the formulae given below and are reported in Table 4.

$$MAE = 1/12 \sum_{i=1}^{12} \left| y_{t+i} - \hat{y}_{t+i} \right|$$
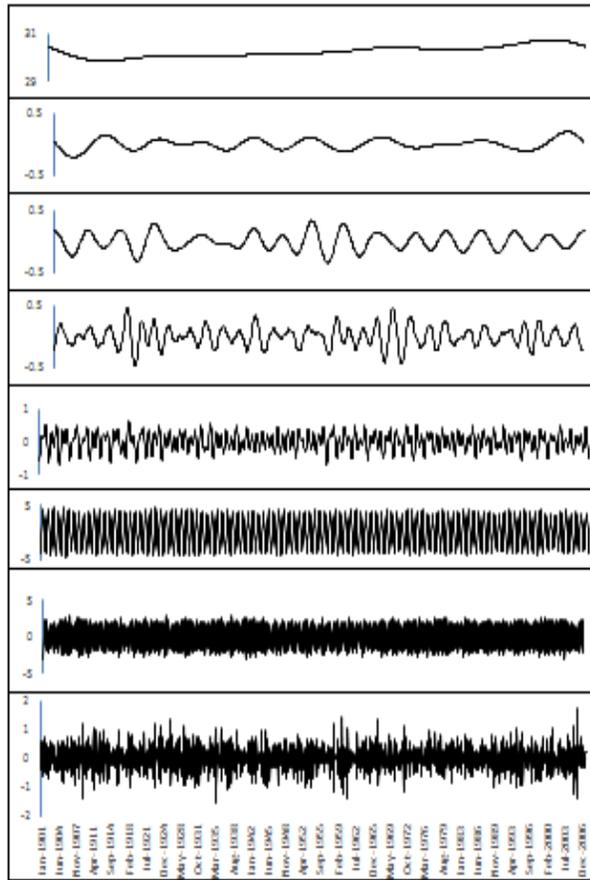
Fig. 9. MRA of Maximum temperature at level 7 D1, D2, D3, D4, D5, D6, D7 and S7 (From bottom to top)
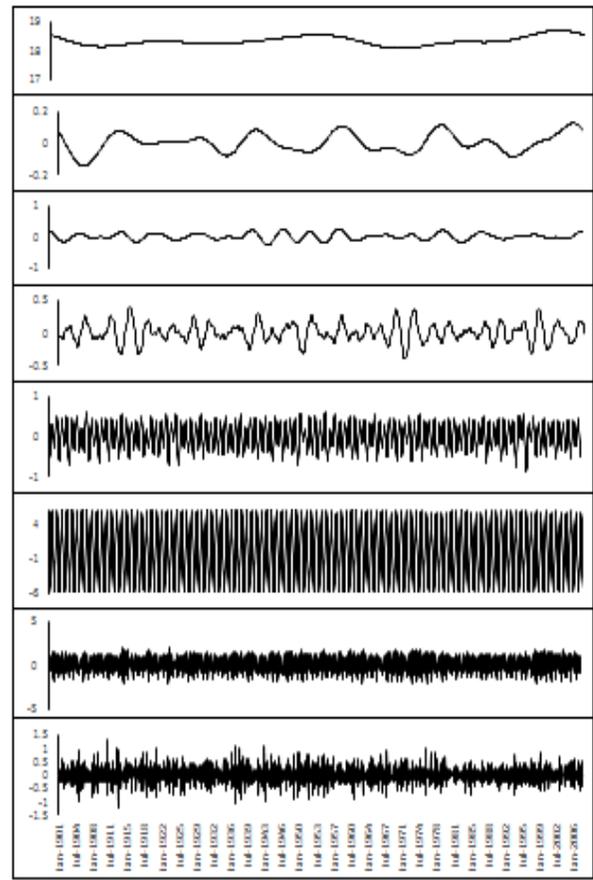
Fig. 10. MRA of Minimum temperature at level 7 D1, D2, D3, D4, D5, D6, D7 and S7 (From bottom to top)

$$RMSPE = \left[1/12 \sum_{i=1}^{12} \left\{ (y_{t+i} - \hat{y}_{t+i})^2 \right\} \right]^{1/2}$$

$$RMAPE = 1/12 \sum_{i=1}^{12} \left\{ \left| y_{t+i} - \hat{y}_{t+i} \right| / y_{t+i} \right\} \times 100$$

A perusal of above table indicates that for both the temperature series data, RMAPE is less than 5% indicating the performance of the model is satisfactory. The fitted *vs* observed maximum as well as minimum temperature series are plotted in Figs. 7 and 8 respectively. Both the figures justify the accuracy of the model fitting.

4.3. *Modelling of temperature series by Wavelet approach*

For computation of MODWT of temperature series by Wavelet approach, methodology discussed in methodology section is followed. Here, we take J_0 as 7.

Haar wavelet is used for analysing the data on a scale by scale basis to reveal its localized nature as exhibited by MRA coefficients at level 7 in Figs. 9 and 10. A perusal indicates that localized variation in the data is detected at lower scale, whereas global variation is detected at higher scale. The wavelet coefficients are related to differences (of various order) of (weighted) average values of portions of X_t concentrated in time. Coefficients at the top (below) provide "high frequency" ("low frequency") information. Wavelet coefficients do not remain constant over time and reflects changes in the data at various time-epochs. Locations of abrupt jumps can be spotted by looking for vertical (between levels) clustering of relatively large coefficients.

5.   Conclusions

Long memory time series have been analysed by using ARFIMA models. Model parameter *d* reflects the

long memory in the maximum and minimum temperature series. It is found that in the both the series long memory parameter is significant. The study has revealed that the ARFIMA model could be used successfully for modelling the temperature series. The predictive ability of ARFIMA model was investigated in terms of relative mean absolute percentage error. The model has demonstrated a good performance in terms of explained variability and predicting power. Multiresolution analysis (MRA) was carried out to explore the local as well as global variations in both the temperature series over the years. The variability in minimum temperature is found to be more than maximum temperature. The study reveals that there are pockets of change in the temperature pattern (both in maximum as well as in minimum temperature) which may be clearly visible by vertical clustering of coefficients in MRA.

Acknowledgements

References

Beran, J., 1994, "Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models", *Journal of the Royal Statistical Society*, **B 57**, 4, 659-672

Booth, G. G., F. R. Kaen and P. E. Koveos, 1982, "R/S analyses of foreign exchange rates under two international monetary regimes", *Journal of Monetary Economics*, **10**, 407-415.

Daubechies, I., 1992, "Ten lectures on wavelets", CBS-NSF regional conferences in applied mathematics 61", Society for industrial and applied mathematics (SIAM), Philadelphia, PA, MR 1162107.

Dhimri, A. P., Mohanty, U. C. and Rathore, L. S., 2005, "Minimum temperature forecast at Manali", India. *Current Science*, **88**, 4, 927-934

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

Gil-Alana, L. A., 2004, "Long memory behaviour in the daily maximum and minimum temperatures in Melbourne, Australia", *Meteorological Applications*, **11**, 319-328.

Granger, C. W. J., 1980, "Long memory relationships and the aggregations of dynamical models", *Journal of Econometrics*, **14**, 227-238

Granger, C. W. J. and Joyeux, R., 1980, "An introduction to long-memory time-series models and fractional differencing", *Journal of Time-series Analysis*, **4**, 221-238.

Helms, B. P., F. R. Kaen and R. E. Rosenman, 1984, "Memory in commodity futures contracts", *The Journal of Futures Markets*, **10**, 559-567.

Hosking, J. R. M., 1981, "Fractional differencing", *Biometrika*, **68**, 165-176.

Hurst, H. E., 1951, "Long-term storage capacity of reservoirs", *Transactions of the American Society of Civil Engineers*, **116**, 770-99.

Hurvich, C. M., Deo, R. and Brodsky, J., 1998, "The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time-series", *Journal of Time series Analysis*, **19**, 19-46.

Jin, H. J. and Frechette, D., 2004, Fractional integration in agricultural futures price volatilities", *American Journal of Agricultural Economics*, **86**, 432-443.

Mohan, V., Jargle, N. K. and Kulkarni, P. D., 1989, "Numerical prediction of daily maximum temperature over Ozar", *Mausam*, **40**, 227-28.

Ogden, T., 1997, "Essential Wavelets for statistical applications and data analysis", Birkhauser, Boston.

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 Science*, **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.

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*, 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. and Birthal, P. S., 2015, "Investigating rainfall trend over India using wavelet technique", *Journal of Water and Climate Change*, **7**, 2, 365-378.

Percival, D. B. and Walden, A. T., 2000, "Wavelet methods for time-series analysis", Cambridge University Press.

Phillips, P. C. B. and P. Perron, 1988, "Testing for unit roots in time series regression", *Biometrika*, **75**, 335-346.

Raj, Y. E. A., 1998, "Prediction of winter minimum temperature at Pune by analogue and regression method", *Mausam*, **40**, 175-80.

Robinson, P. M., 1995, "Log-periodogram regression of time-series with long-range dependence", *The Annals of Statistics*, **23**, 1048-1072.

Robinson, P. M., 2003, "Time series with long memory", Oxford university press, Oxford.

Said, S. E. and D. Dickey, 1984, "Testing for unit roots in autoregressive moving-average models with unknown order", *Biometrika*, **71**, 599-607.

Shimotsu, K., 2010, "Exact local whittle estimation of fractional integration with unknown mean and time trend", *Econometric Theory*, **26**, 2, 501-540.

Tanaka, K., 1999, "The nonstationary fractional unit root", *Econometric Theory*, **15**, 549- 582.

Woodcock, F., 1984, "Australian experimental model output statistics forecasts of daily maximum and minimum temperature", *Monthly Weather Review*, **112**, 2112-2121.