INTERNATIONAL JOURNAL OF HIGHER EDUCATION AND RESEARCH
IJHER, Vol., 9(1), 2019, 319-332.
www.ijher.com
IJHER
(ISSN 2277 260X)
CODEN:IJHER
ORIGINAL ARTICLE

Spectral Analysis of Rainfall using Haar Wavelet Transforms

Anil Kumar* and Uttam Kumar Rawat

Innovative Physics Research Laboratory, Department of Physics

Hindu College, Moradabad-244001, Uttar Pradesh (India)

*Email: akumarmbd@gmail.com

ABSTRACT

The study and understanding of variation in rainfall in space, time and amounts, and their attendant effects on the ecosystem arevery important for the existence of all forms of life.We have studied the rainfall variability on basis of its average daily record in Moradabad, the Metropolitan city of Western Uttar Pradesh, Republic of India, during the period from 01/07/2008 to 30/06/2018(10 years).The wavelets are very useful in signaldecompositionand reconstruction algorithms where it is desirable to recover the original information with minimal loss. There is slight significant decreasing trend for long term northeast monsoon rainfall over Moradabad region.Approximations are low frequency and details are high frequency content of the signal and hence play important role in signal or data analysis. Approximations provide information about the trend of signal. The trend is its slowest part of the signal and represents the behavior of signal corresponding to greatest scale value. A peak in the details shows rapid change or fluctuation in the quantity of rainfall in that time period. Skewness is a measure of symmetryand more precisely, the lack of symmetry while Kurtosis is a measure of whether the data are peakedness or flatness relative to a normal distribution.

Key words: Rainfall, Wavelet, Approximation, Detail, Trend.

INTRODUCTION

Water is one of the most important natural resources on earth that is necessary for survival of all forms of life. Rainfall is one of the main sources of water in any region. The availability of water for various purposes is very much depending upon the amount of precipitation in that particular region.Rainfall affects the environment and society in various ways ranging from water availability for livelihood, agriculture for functioning of various industries, hydroelectric power generation etc. and thus affects the every aspect of life inthat region. In order to understand the hydrological balance and the complex interactions among the components within the hydrologic cycle, we need to focus ourselves on all the information regarding precipitation. It is the meteorological phenomenon that has the greatest impact on human activities. Moreover, in order to determine the amount of water available to meet various demands such as agricultural, industry, domestic water supply and for hydroelectric power supply, the rainfall received in that particular area plays an important role. Understanding rainfall variability is essential to optimally manage the scarce water resources that are under continuous stress due to the increasing water demands, increase in population and the economic development(Longobardiet al., 2010). Though, climate change is global in nature, potential changes may be consists of some dramatic regional differences. Long term rainfall patterns may get influenced by the global climatic changes and this may result with the danger of increasing the occurrences of droughts and floods. Due to population growth and economic developments, the need for water is increasing daily.Consequently, the water management using all the available resources is becoming more challenging and crucial day by day. In order to develop an effective water management strategy, it is necessary to understand and assess its impact on the ecosystem. The study and understanding of variation in rainfall in space, time and amounts, and their attendant effects on the ecosystem is very important.

Moradabad city is very near to the forest and hill range of Uttarakhand (around 60 Km). This forest and hill range direct affects the weather and climate of Moradabad region. It is an agriculture-based district where about 80% of its people are directly or indirectly engaged in a wide range of agricultural activities. Rainfall is one of the important natural factors for the agricultural production. The variability of rainfall with respect to time and space are important for the agriculture as well as the economy of the region. The rainfall is changing on both the global and the regional scale due to global warming, which is well established. Study on rainfall variability and their statistical analysis are therefore important for long-term water resources planning, agricultural development and disaster management mainly in the context of global climatic change. Although a number of studies have been carried out on rainfall patterns like Kumar (2003) on Rainfall characteristics of Shimla district (HP), Santos et al. (2003) on analysis of rainfall using wavelet transforms, Asati (2012) on analysis of rainfall data for drought investigation at Brahmapuri (MS), Hasan et al. (2014) on statistical behaviour of daily maximum and monthly average rainfall along with rainy days variation in Sylhet, Bangladesh, but nobody work has been found on rainfall trends and their variability particularly for Moradabad. The discrete nature of rainfall in time and space has always posed unique problems for the climatologist compared to more continuous climatic elements such as temperature and pressure. Rainfall totals have often been used in studies that examine large-scale fluctuations in rainfall.

Wavelet is anew mathematical toolused to extract information from many different kinds of data related to signal processing, atmospheric and environmental physics etc. (Kumar et al., 2015; Kumar, 2017). In wavelet, we use a single function and its dilation and translation to generate a set of orthonormal basis functions to represent a signal. Number of such functions is infinite and we choose one that suits to our data. The range of interval over which scaling function and wavelet function are defined is known as support of wavelet. Beyond this interval (support) the functions should be identically zero. There is an interesting relation between length of support and number of coefficients in the refinement relation. For orthogonal wavelet system, the length of support is always less than no. of coefficients in the refinement relation. It is also very helpful to require that the mother function have a certain number of zero moments, according to:

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \quad (1.1)$$

The mother function can be used to generate a whole family of wavelets by translating and scaling the mother wavelet as following:-

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) = T_b D_a \psi \quad (1.2)$$

Here $b$is the translation parameter and a is the dilation or scaling parameter. Provided that ψ (t) is real-valued, this collection of wavelets can be used as an orthonormal basis. A critical sampling of the continuous wavelet transform is as following:-

$$W_{a,b} = \int f(\mathrm{t}) \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) dt$$

$$= \int f(t)\psi_{a,b}(t)\, dt \quad (1.3)$$

By taking $a = 2^j \& b/a = k$ where $j$and $k$are integers representing the set of discrete translations and discrete dilations (Mallat, 1998; Kumar et al., 2018), we obtain the discrete wavelet transformas:-

$$W_{j,k} = \int f(t) 2^{j/2} \psi\left(2^j t - k\right) dt \qquad (1.4)$$

In discrete form, Haar wavelets are related to a mathematical operation called the Haar transform. One distinctive feature that the Haar transform enjoys is that it lends itself easily to simple hand calculations and more involved computer computations.Haar thought that there exists orthonormal system $h_0(t), h_1(t) \dots \dots \dots \dots h_n(t) \dots \dots$of functions defined on $[0,1]$ such that for any function $f(t)$continuous on $[0,1]$, the series;

$$\langle f, h_0 \rangle h_0(t) + \langle f, h_1 \rangle h_1(t) + \dots\dots\dots\dots \langle f, h_n \rangle h_n(t) + \cdots \dots \dots \dots$$

converge to $f(t)$uniformly on $[0,1]$. Here $\langle f, h_0 \rangle = \int_0^1 f(t)\overline{h(t)} dt$, where $\bar{h}$ is the complex conjugate of h. This problem has infinite number of solutions. Haar himself discovered the simplest solution and at the same time open one of the routes leading to wavelets. Haar begins with the function that is equal to 1 on [0,1] and 0 outsite the interval $[0,1]$.

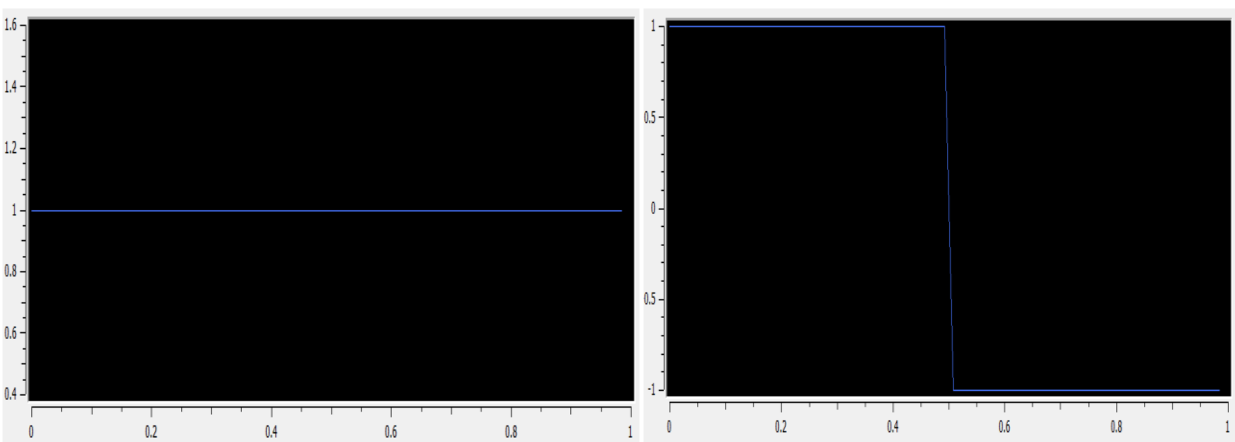


Figure 1: The Haarscaling function ϕ and its wavelet ψ

Haar wavelet is constructed from the MRA generated by scaling function $\phi(t)$=$\chi_{[0,1]}(t)$ (Figure 1). Since,

$$\phi(\mathrm{t}) = \phi(2t) + \phi(2t + 1) \qquad (1.5)$$

and$\psi(t) = \phi(2t) - \phi(2t + 1)$(1.6)

$= \chi_{\left[0,\frac{1}{2}\right]} - \chi_{\left[\frac{1}{2},1\right]}$

The present work analyzes average daily rainfall data by estimating different statistical and wavelet analytical parameters. Daily rainfall and its monthly, seasonal and yearly variation as well as rainy day pattern are estimated and discussed focusing climate change. Understanding the rainfall pattern is tough for the solution of several regional environmental issues of water resources management, with implications for agriculture, climate change, and natural calamity such as floods and droughts. This study aims to establish the rainfall trends of the Moradabad region. This work studies rainfall variability in Moradabad, the Metropolitan city of Western Uttar Pradesh, Republic of India, during the period from 01/07/2008 to 30/06/2018 (10 years). Most of the rainfall is found out to be taking place in Northeast monsoon season. The analysis revealed great degree of variability in precipitation with time. There is slight significant decreasing trend for long term northeast monsoon rainfall over Moradabad region. Information about the trends of rainfall are important as it is closely related to the practical water relates issues in the region especially flood related problems. Thus it becomes increasingly important to study the trend in precipitation and their physical explanation.

DATA METHODOLOGY

The scaling functions or associated wavelets do not play role directly in the computation of signal expansion coefficients. This is because of the relationship between expansion coefficients at a lower scale and expansion coefficients at a higher scale. We use Haar scaling and wavelet functions to establish the relation between these coefficients. We have the refinement relation responsible for signal decomposition:-

$\phi(t) = \sum_{n=0}^{N-1} \alpha(n)\sqrt{2}\phi(2t-n) = \sum_{n=0}^{N-1} \alpha(n)\phi_{1,n}(t)$(2.1)

where N is the number of coefficients in the refinement relation. For Haar $N = 2$ and the relation is,

$$\phi(t) = \alpha(0)\sqrt{2}\phi(2t) + \alpha(1)\sqrt{2}\phi(2t-1)$$

$= \alpha(0)\phi_{1,0}(t) + \alpha(1)\phi_{1,1}(t)$(2.2)

Here $\alpha(0)$ and $\alpha(1)$ are normalized coefficients. For Haar, $\alpha(0) = \frac{1}{\sqrt{2}}$and $\alpha(1) = \frac{1}{\sqrt{2}}$. In general we can write,

$$\phi(2^j t - k) = \sum_{n=0}^{N-1} \alpha(n)\sqrt{2}\phi(2(2^j t - k) - n)$$
$$= \sum_{n=0}^{N-1} \alpha(n)\sqrt{2}\phi(2^{j+1} t - 2k - n)$$

By putting m=$2k + n$, we have,

$\phi(2^j t - k) = \sum_{m=2k}^{2k+N-1} \alpha(m - 2k)\sqrt{2}\phi(2^{j+1} t - m)$ (2.3)

Similarly,

$\psi(2^j t - k) = \sum_{m=2k}^{2k+N-1} \beta(m - 2k)\sqrt{2}\phi(2^{j+1} t - m)$(2.4)

The subspace V_j is described as

$V_j = \ span\overline{\{\phi(2^{j/2} t - k)\}}$, Then $f(t) \in V_{j-1}$

$\Rightarrow f(t) = \sum_k s_{j-1}(k)\, \phi\left(2^{(j+1)/2}(t - k)\right)$(2.5)

That is, $f(t)$ is expressed as a linear combination of bases inV_{j+1}. At this scale, at this scale wavelets are not coming to picture. Since $V_j = V_{j+1} \oplus W_{j+1}$, to represent same signal, in the next lower scale, we require the help of wavelets(Grinsted et al., 2004; Antoine, 2004). So we have,

$f(t) = \sum_k s_j(k)\, 2^{j/2}\phi(2^j t - k) + \sum_k d_j(k)\, 2^{j/2}\psi(2^j t - k)$(2.6)

The $2^{j/2}$ terms maintain the unity norm of the basis functions at various scales. The $s_j(k)$ is the projection of $f(t)$ on to the corresponding normalized base $2^{j/2}\phi(2^j t - k)$, that is,

$s_j(k) = \langle f(t), \phi_{j,k}(t)\rangle = \int f(t)\, 2^{j/2}\phi(2^j t - k) dt$(2.7)

Substituting equation (7.3) in above equation and interchanging summation and integral, we have,

$s_j(k) = \sum_{m=2k}^{2k+N-1} \alpha(m - 2k) \int f(t)\, 2^{(j+1)/2}\phi(2^{(j+1)} t - m) dt$(2.8)

The integral in above equation is basically the projection of $f(t)$ on to the normalized base $2^{(j+1)/2}\phi(2^{(j+1)} t - m)$and according to equation (2.7), it is $s_{j-1}(m)$. Therefore,

$s_j(k) = \sum_{m=2k}^{2k+N-1} \alpha(m - 2k) s_{j-1}(m)$ (2.9)

The corresponding relationship for the wavelet coefficient is,

$d_j(k) = \sum_{m=2k}^{2k+N-1} \beta(m - 2k) s_{j-1}(m)$ (2.10)

Thus equation (2.9) is equivalent to convolution followed by down sampling of even indexed terms in the output series (Figure 2):

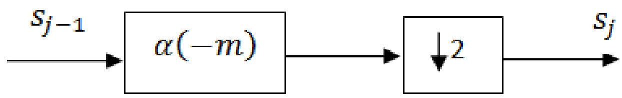


Figure 2: Relation between s_{j-1} and s_j

Similarly, Equation (2.10) can be shown as (Fig. 3):-

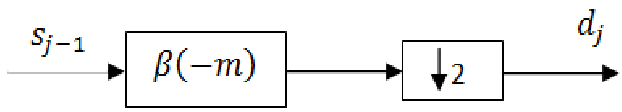


Figure 3: Relation between s_{j-1} and d_j

We can also write,

$s_j(k) = \sum_{m=2k}^{2k+N-1} \alpha[-(2k-m)] s_{j-1}(m)$ (2.11)

$= \sum_{m=2k}^{2k+N-1} \alpha'(2k-m) s_{j-1}(m)$(2.12)

Where $\alpha' = \alpha(-n)$ represent a time reversed series of $\alpha(n)$. The parameter k is fixed by the left hand side of the equation. Comparing equation (2.11) and (2.12), we find that equation (2.12) is a sort of convolution between the series $\{\alpha(-m)\}$and $\{s_{j-1}(m)\}$. However, there is a difficulty in interpreting the formula as a convolution. Translation quantity in equation (2.11) is $2k$. That is, to find $s_j(k)$, first we translate $\{\alpha(-m)\}$ by $2k$, fold it backwards, multiplying the corresponding term in the series $\{s_{j-1}(m)\}$ and add. Since translation is by $2k$, we see the process as an ordinary convolution followed by decimation of alternative terms. This will become obvious if we take a small example and substitute in equation (2.9). Let us assume that our filter is given by $\alpha(0)$and $\alpha(1)$, and the signal coefficients when expressed using the bases of V_{j-1} is $s_{j-1}(0)$, $s_{j-1}(1)$, $s_{j-1}(2)$,$s_{j-1}(3)$, $s_{j-1}(4)$, $s_{j-1}(5)$. For all other index assume that $h(k)$ and $s_{j-1}(k)$ are zero. By equation (2.9),

$s_j(k) = \sum_{m=2k}^{2k+N-1} \alpha(m-2k) s_{j-1}(m)$ (2.13)

Where $N = 2$, because we have only $\alpha(0)$ and $\alpha(1)$. Therefore, for $k = 0$, we have,

$$s_j(0) = \sum_{m=0}^{1} \alpha(m) s_{j-1}(m)$$

$= \alpha(0)s_{j-1}(0) + \alpha(1)s_{j-1}(1)$(2.14)

For $k = 1$, we have,

$s_j(1) = \sum_{m=2}^{3} \alpha(m-2)s_{j-1}(m)$(2.15)

$$= \alpha(0)s_{j-1}(2) + \alpha(1)s_{j-1}(3)$$

For $k = 2$, we have,

$$s_j(2) = \sum_{m=2}^{3} \alpha(m-4)s_{j-1}(m)$$

$= \alpha(0)s_{j-1}(4) + \alpha(1)s_{j-1}(5)$(2.16)

To give interpretation in terms of convolution, let us convolve series $\{\alpha(-m)\}$ with the series $\{s_{j-1}(m)\}$. In the present case, $\{\alpha(-m)\}$ series becomes $\alpha(1)$and $\alpha(0)$, the time reversed form of series $\{\alpha(0), \alpha(1)\}$. When we convolve these two series, the output sequence $a_j(k)$ can be represented by;

$$a_j(0) = \alpha(1)s_{j-1}(0)$$

$$a_j(1) = \alpha(0)s_{j-1}(0) + \alpha(1)s_{j-1}(1)$$

$$a_j(2) = \alpha(0)s_{j-1}(1) + \alpha(1)s_{j-1}(2)$$

$a_j(3) = \alpha(0)s_{j-1}(2) + \alpha(1)s_{j-1}(3)$and so on.

It can be easily seen that $s_j(0)$, $s_j(1)$, $s_j(2)$ etc. given by equations (2.13) –(2.15) are $a_j(0)$, $a_j(1)$ and $a_j(2)$ etc.We can say that the decomposition of a signal in V_{j-1} into sum of two signals, one in space V_j and the other in W_j is visualized as shown below (Figure 4):

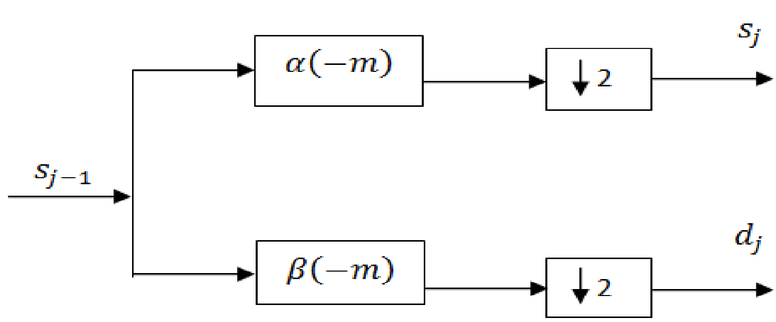


Figure4: Relation among s_{j-1}, s_j and d_j

A fundamental task in many statistical analyses is to characterize the location and variability of a data set. A further characterization of the data includes skewness, kurtosis and variance(Rockingeret al. 2002). Skewness is a measure of symmetry, or more precisely, the lack of symmetry. A distribution of data set is symmetric if it looks the same to the left and right of the center point. Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution. That is, data sets with high kurtosis tend to have a distinct peak near the mean, decline rather rapidly, and have heavy tails. Data sets with low kurtosis tend to have a flat top near the mean rather than a sharp peak. The standard deviation (SD) is a measure that is used to quantify the amount of variation or dispersion of a set of data values.

RESULTS AND DISCUSSION

The rainfall data from 01/07/2008 to 30/06/2018 is taken as original signal *s* which represents the average daily behavior of the rainfall during this periodin Moradabad (Figure 5).

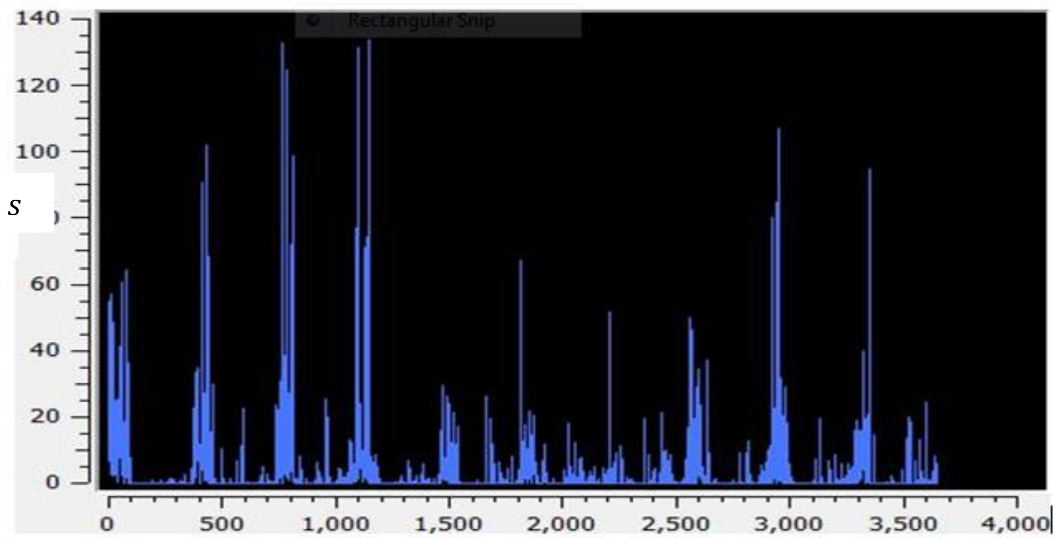
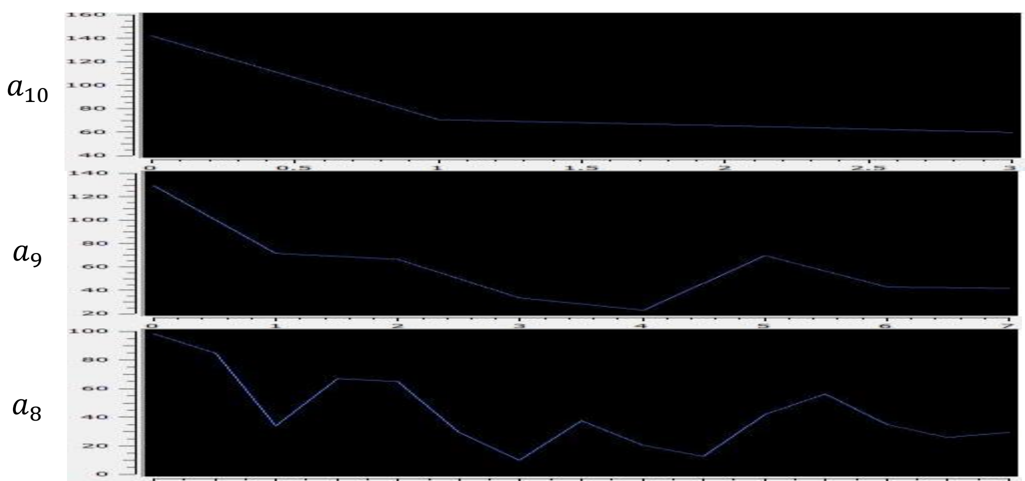


Figure5:Plot of rainfall data from 01/07/2008 to 30/06/2018 in Moradabad

We have decomposed this signal up to 10 levels using Haar wavelet.



a_7

a_6

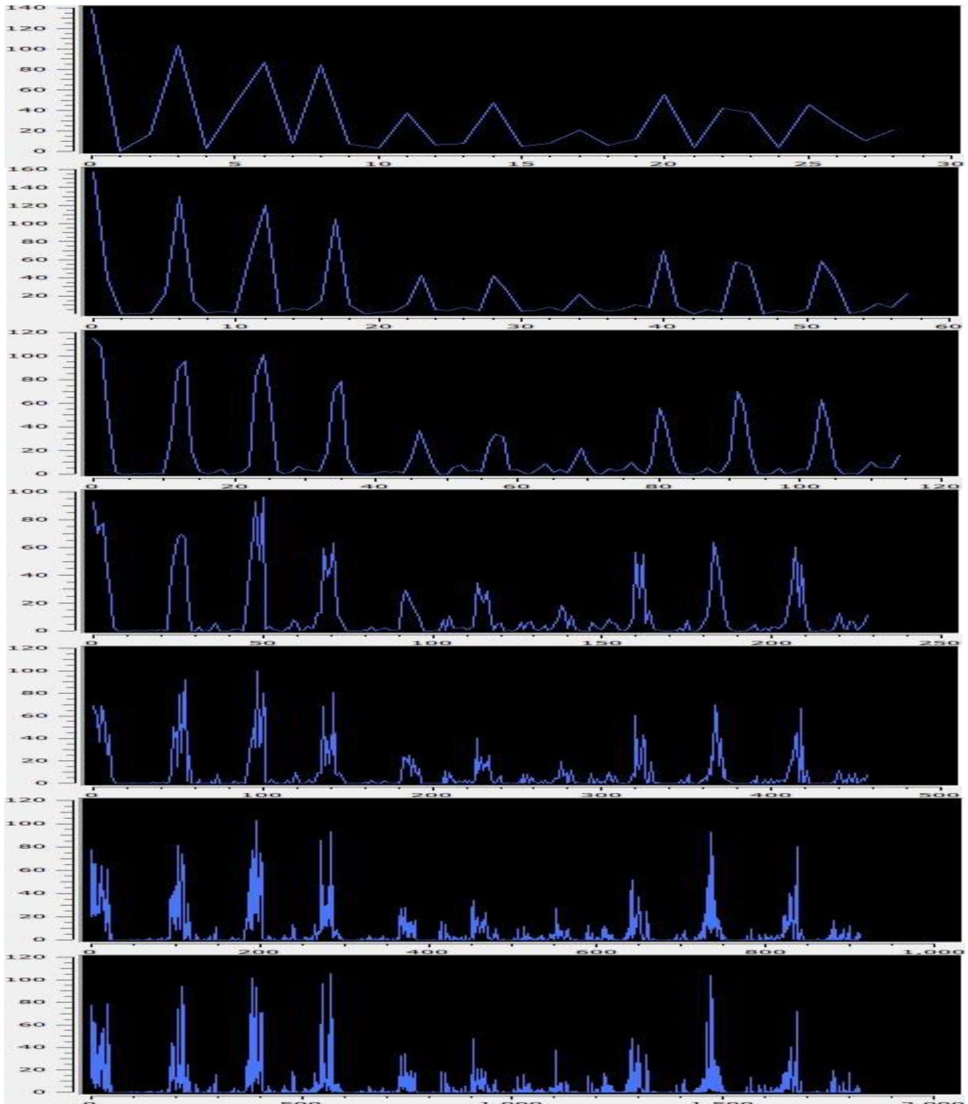
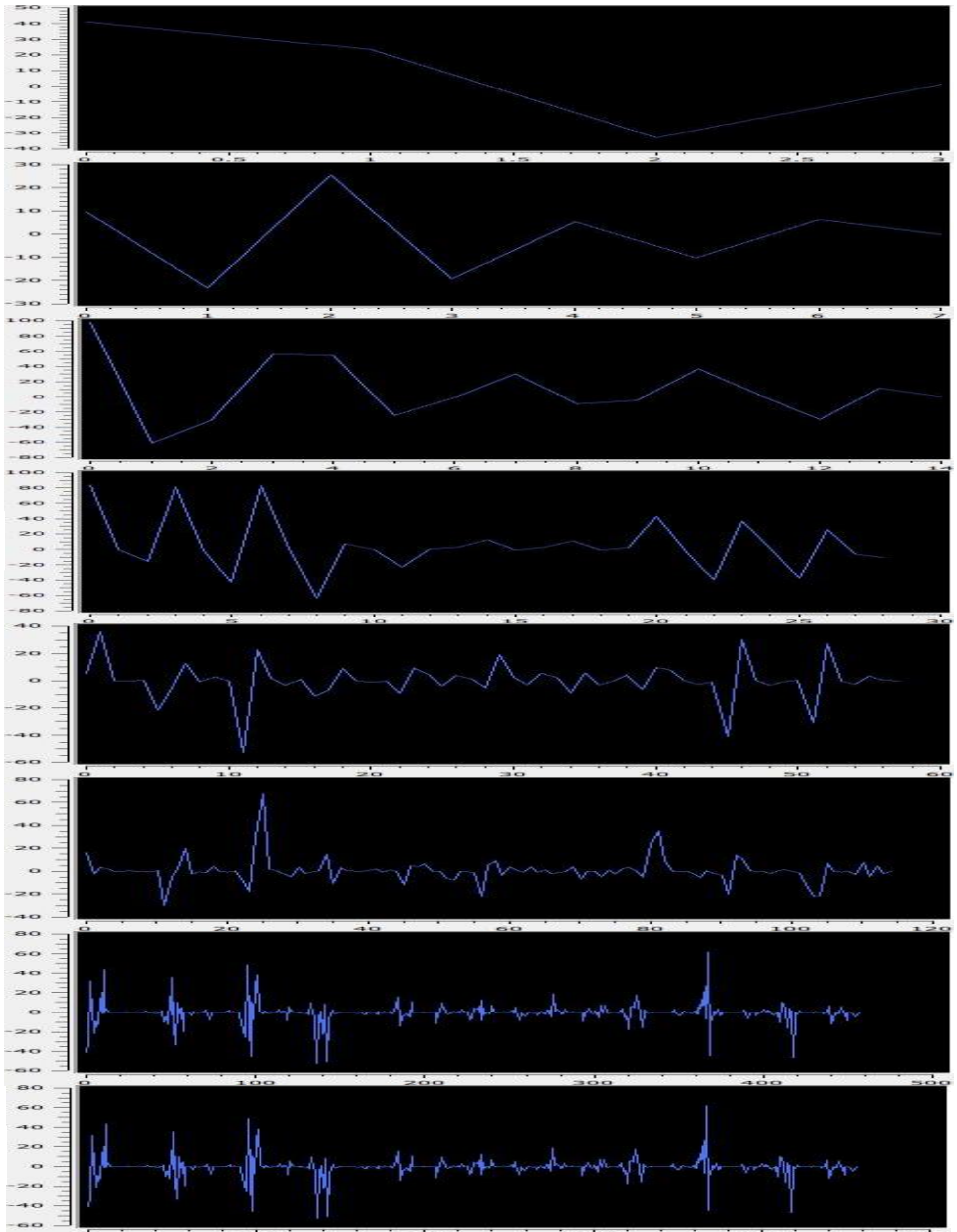


Figure6:Approximation Coefficients up to level 10



d_{10}

d_9

d_8

d_7

d_6

d_5

d_4

d_3

d_2

d_1

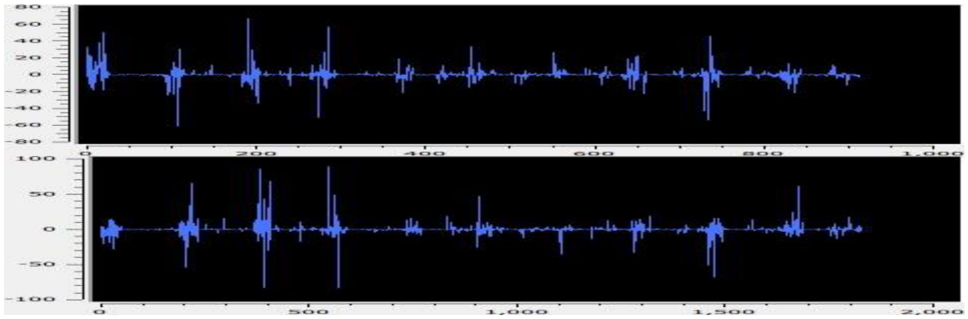


Figure7:Detail Coefficients up to level 10

In Figure 6 and 7, the approximations and the details belonging to Haar wavelet decomposition up to level 10 are shown. Approximations a_1, a_2, a_3,……….a_{10} represent 2^1, 2^2, 2^3……….2^{10} days average behavior of the signal. Likewise details represent 2^1, 2^2, 2^3……….2^{10} days variations in the signal. In wavelet analysis, Approximations are low frequency and details are high frequency content of the signal.Due to low frequency content, the approximations are very important part of the signal and provide information about the trend of signal. The trend represents the behavior of signal corresponding to greatest scale value. As the scale value is increased, the resolution of the signal decreases, that provides better estimation of unknown trend of the signal. From the trend of the signal, it is obvious that the average rainfall is continuously decreasing in the Moradabad region in that time period. It is also clear that the rate of decrement of average rainfall in current years is not so high that was in the previous years. A peak in the details shows rapid change or fluctuation in the quantity of rainfall in that time mode.

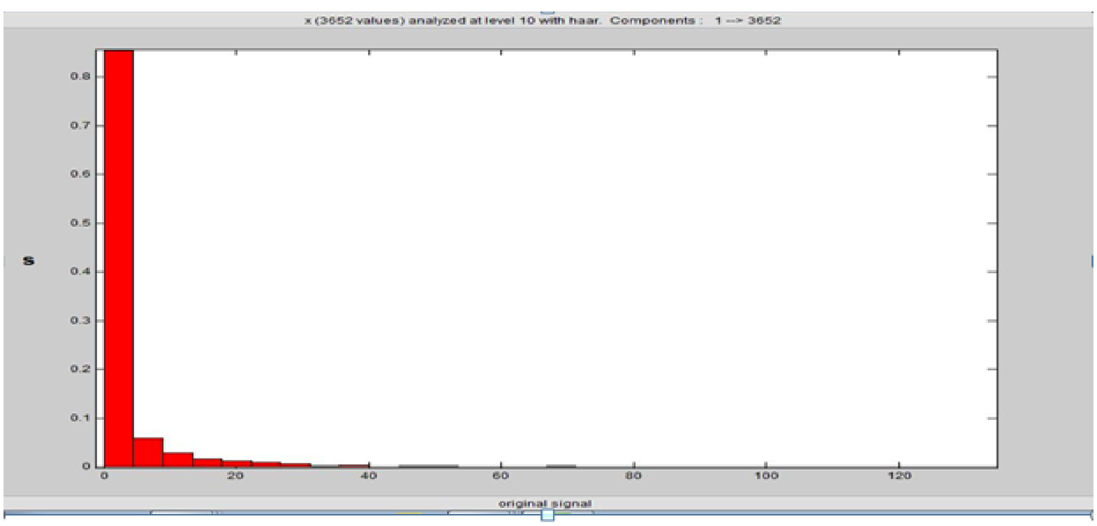


Figure8:Histogram of rainfall

The maximum value of rainfall in that time period is 133.5 mm in a day and average value is 2.726791 mm. The maximum variation on daily average in rainfall is 89.0356 mm. on 29 June, 2011. Histogram provides a visual interpretation of numerical data by indicating the number of data points that lie within a range of values. Histogram of the rainfall data represents a strong skewnesstowards right as shown in Fig. 8. The value of skewnessparameter is 7.375535, which measures the asymmetry of the probability distribution of a real-valued random variableabout its mean. The value of Kurtosis parameter is 76.25836for rainfall data in that time mode, which is measure of the peakedness of the probability distribution. The standard deviation of given data is 8.930644. High value of standard deviation indicates that the data points are spread out widely in that time period.

CONCLUSION

The trend of the rainfall becomes more sharp and clear from the plot of approximation coefficients of up to level 10. From the trend of the signal, it is obvious that the average rainfall is continuously decreasing in the Moradabad region in that time period.The rapid change or fluctuation in the quantity of rainfall in that time mode is described by the peaks in the details. Positive value of skewness indicates that the rainfall data is skewed to right. High positive value of kurtosis indicates the strong intermittency in the rainfall variability. High value of standard deviation indicates that the data points are spread out over a wider range of values. Analyzing the time series of rainfall variability in given time period, it is clear that the energy of high pass filtered time series has a small contribution to the time series as compared with the low-pass filtered time series. Taking into account the results, we can say that the generalized wavelet analytical approach provides a simple and accurate framework for modeling the behavior of time series of rainfall data.

REFERENCES

Longobardi, A. and Villani, P. (2010). Trend analysis of annual and seasonal rainfall time series in the Mediterranean area, *International Journal of Climatology*, 30, 1538-46, (www.interscience.wiley.com), DOI: 10.1002/joc.2001.

Kumar, V. (2003). Rainfall characteristics of Shimla district (HP), *Journal of Indian Water Resources Society*, 23(1), 1-10.

Santos, C.A.G.,Galvao, C.O. and Trigo, R.M. (2003). Rainfall data analysis using wavelet transforms, *International Symposium at Montpelier on Hydrology of the Mediterranean and Semiarid regions,* USA, 195-201.

Asati, S.R. (2012). Analysis of rainfall data for drought investigation at Brahmapuri (MS), *International Journal of Life science, Biotechnology and Pharma Research*, 1, 81-86.

Hasan, G. M. J., Chowdhury, M. A. and Ahmed, I. S. (2014). Analysis of the statistical behaviour of daily maximum and monthly average rainfall along with rainy days variation in Sylhet, Bangladesh *Journal of Engineering Science and Technology*, 9(5), 559 – 578.

Kumar, A., Kumar, S. and Pathak, J.K. (2015). Spectral analysis of river ramganga hydraulics using discrete wavelet transform, *International Journal of Advance Research and Innovation (Special issue for Proceedings of International Conference of Advance Research and Innovation)*, 370-373.

Kumar, A. (2017). Wavelet analysis of air pollution due to ground level nitrogen dioxide, *Asian Journal of Science and Technology,* 8(8), 5218-5222.

Kumar, A.,Rawat, U.K. and Singh, P. (2018). Spectral analysis of ground level nitrogen dioxide and sulphur dioxide, *Progress in Nonlinear Dynamics and Chaos*, 6(1), 11-17.

Mallat, S. G. (1998). A wavelet tour of Signal Processing, Academic Press, New York.

Grinsted, A., Moore, J. C. and Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series, *Nonlinear Progress in Geophysics*, 11, 561-566.

Antoine, J. P.(2004).Wavelet analysis: A new tool in Physics,*Wavelets in Physics*, Van Der Berg, J. C., ed., 9-21.

Rockinger, M. and Jondeau, E.(2002). Entropy densities with an application to autoregressive conditional skewness and kurtosis, *Journal of Econometrics*, 106(1),119-142.