IJIREEICE

ISSN (Online) 2321-2004
ISSN (Print) 2321- 5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2018

Stationarity Estimation for Time Series Rainfall Data

Vivekananda Mukherjee[1], Arnab Ghosh[2], Bayan Das[3], Binisha Chowdhury[4], Ankan Das[5]

Assistant Professor, Department of Electronics & Communication Engineering,

Techno India College of Technology, New Town, Kolkata[1]

Final Year Student, Department of Electronics & Communication Engineering,

Techno India College of Technology, New Town, Kolkata[2,3,4,5]

Abstract: In this paper, fractality and stationarity of average rainfall value of India have been investigated with a focus on the self similarity pattern and nature of frequency oscillation of the two parameters namely the average rainfall value of North India and South India. The time series of these parameters in between 31st Oct 2006 to 1st Nov 2017, of North India and South India have been chosen for extracting the nature of scaling (fractality) and stationary behavior using statistical methodologies. Hurst Exponent for the time series have been calculated using the methodologies like Higuchi Fractal Dimension (HFD) and General Hurst Estimation (GHE). It has been noticed that both the time series show Short Range Dependent (SRD) anti-persistent behavior. Continuous Wavelet Transform (CWT) method has been incorporated to identify the stationarity/non-stationarity of the data-series where both the time series exhibit the non stationarity.

Keywords: Rain Fall Data; Hurst Exponent; Generalized Hurst Estimation (GHE); Higuchi's Fractal Dimension (HFD); CWT

I. INTRODUCTION

The variational trends in rainfall patterns in India have been extensively studied by researchers. However the focus has generally been on identification of seasonal trends in rainfall data, using homogeneous data series constructed from average rainfall data (1). Area weighed methods have been used to construct rainfall data series using data from 306 uniformly distributed stations distributed over the country (2)(3). Short duration (below 12 hours) temporal changes have also been analyzed in recent years for determination of trends in a more accurate manner (4). Extreme rainfall events have also been studied in detail to enable the accurate identification of variational patterns in the obtained data (5). In this paper a proposal has been taken to expose the nature of the scaling behavior and time dependency of the frequency (stationarity or non-stationarity) of the average rainfall value of North India and South India during 31st October, 2006 to 1st November, 2017 which can be treated as the signatory representative of any rain fall data analysis.

A concurrent study of North Indian(NI) average rain fall value and South Indian (SI) average rain fall value time series may give a feasible nature of the monsoon trend, the agricultural stability of the country, and fluctuation of market price . In this work Hurst exponent has been calculated for revealing the scaling behavior of the time series, North Indian (NI) average rain fall and South Indian(SI) average rain fall. Two different methods like Higuchi Fractal Dimension (HFD) and General Hurst Estimation (GHE) have been used to calculate the Hurst Exponents to understand the nature of the signals with respect to different scales to identify the signals as fractional Brownian motion in order to check whether they are stationary or non-stationary. For getting an unarguable conclusion about the scaling property of the time series, it will be useful to apply more than one method to estimate the Hurst Exponent. Hence two methods (mentioned above) have been chosen to calculate the Hurst Exponent for confirming the authenticity of the conclusions taken out of the results. Stationary or non-stationary behavior of the data series could be completed by analyzing the fluctuating nature of the average rain fall value of NI and SI. A non-stationary signal has changing frequency whereas stationary signal has constant frequency. The signals are checked with respect to time. The analysis for non-stationary behavior is necessary due to: 1) asymptotic analysis which will not be applicable for the regression model with non-stationary variables. Usually "t-ratios" does not follow a t-distribution, and hence valid tests about the regression parameters cannot be undertaken. 2) The properties of the signal are highly affected by the stationary or non-stationary behavior. Different methods can be used to check the stationary/ non-stationary behavior of the signals. Continuous Wavelet Transform (CWT) based methods have been incorporated in this paper to determine the nature of frequency dependency of the stock market. The advantages of using CWT are: a) simultaneous localization in time and frequency domain and is computationally fast. b) Wavelets have the great advantage of being able to separate the fine details in a signal. Very small wavelets can be used to isolate very fine details in a signal, while very large wavelets can identify coarse details. It decomposes a signal into component wavelets.

IJIREEICE

ISSN (Online) 2321-2004
ISSN (Print) 2321- 5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2018

II. HURST EXPONENT ESTIMATION

The Hurst exponent provides a statistical measure for long range memory and fractality of a time series (6). The interpretation of Hurst value H is simple. When H is equal to 0.5 indicates a random series and when the value of H ranges between 0 and 0.5 indicates up value will follow the lower value and shows anti-persistent series while H is greater than 0.5 and less than 1 means the direction of next value more likely the same as current value and shows persistent time series (7). More specifically Hurst exponent referred to as Index for long range dependence. Different estimators for the estimation of the Hurst Exponent of any signal or data are available. In this paper, two Hurst estimation methods have been used. Higuchi's Fractal Dimension (HFD) Method has been used along with Generalized Hurst Exponent (GHE) estimation method. HFD method is used for statistical measurement of a time series. Its aim is to provide an estimation of how the variability of a series changes with the length of the time-period. GHE provides the best finite sample behavior among all the methods in respect of the bias and lowest variance. GHE is suitable for any data series/signal irrespective of the size of its distribution tail.

A. Higuchi's Fractal Dimension (HFD) Method: Higuchi's method is a technique use to check the irregularity of time series directly. This procedure has been used to calculate the fractal dimension D of a time series. This technique is proposed by Higuchi (8). In order to derive the fractality, HDF is calculated (9) (10) for the discrete data point series:

$$S : S(1), S(2), S(3), \ldots, S(n) \tag{1}$$

Where n is the data points number. From the original data series d new data series $S_\kappa(d)$ with $\kappa = 1, 2, 3, \ldots d$ are constructed. Where κ is the initial time and d is time interval.

$$S_\kappa(d) : S(\kappa), S(\kappa+d), S(\kappa+2d), \ldots.. S\left(\kappa + \left\lfloor \frac{n-\kappa}{d} \right\rfloor d\right) \tag{2}$$

$L_\kappa(d)$ is the length of the time series $S_\kappa(d)$ and define as:

$$L_\kappa(d) = \frac{1}{d}\left\{\left(\sum_{i=1}^{\frac{n-\kappa}{d}} \left|S(\kappa+id) - S(\kappa+(i-1)d)\right|\right)\frac{n-1}{\left\lfloor \frac{n-\kappa}{d} \right\rfloor d}\right\} \tag{3}$$

Where the term $\frac{n-1}{\left\lfloor \frac{n-\kappa}{d} \right\rfloor d}$ represents a normalization factor and $L_\kappa(d)$ is the normalized sum of the differences of values and calculated as follows:

$$L(d) = \frac{1}{d}\sum_{i=1}^{d} L_\kappa(d) \tag{4}$$

Where $L_\kappa(d)$ is the mean value.

B. Generalized Hurst Exponent (GHE) Method: This method was invented by Hurst (11) which defines a function $\mathrm{K}_p(\gamma)$ as (12) (13)

$$\mathrm{K}_p(\gamma) = \frac{\left(\left\langle \mathrm{X}(t+\gamma) - \mathrm{X}(t)\right\rangle^p\right)}{\left\langle \mathrm{X}(t)^p\right\rangle} \tag{5}$$

Where X(t) is the time series and p is the order of the moment of distribution and γis the lag which ranges between 1 and $\gamma_{\max}$. Generalized Hurst Exponent (GHE), is related to $K(p)$ through a power law:

$$\mathrm{K}_p(\gamma) \propto \gamma^{ph(p)} \tag{6}$$

Depending upon whether it is independent of p or not, a time series can be judged as uni-fractal or multi-fractal (13) respectively. The GHE $h(p)$ yields the value of original Hurst Exponent H for $p = 1$, i.e. $h(1) = H$.

IJIREEICE

ISSN (Online) 2321-2004
ISSN (Print) 2321- 5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2018

III. TEST FOR STATIONARITY / NON-STATIONARITY

A. Continuous Wavelet Transform (CWT) Method

Real world data or signals frequently expose the slow changing trend or oscillations punctuated with transient. Though Fourier Transform (FT) is a potential method for data analysis, however it does not characterize sudden changes proficiently. FT indicates data as sum of sine waves which are not contained in time or space. These sine waves fluctuate forever, therefore to precisely analyze signals that have abrupt changes, need to use new class of functions that are well localized with time and frequency. This introduced the topic of wavelets concept. The primary goal of the Continuous Wavelet Transform (CWT) (14) (15) is to get the signal's energy distribution in the time and frequency domain concurrently. The continuous wavelet transform is a generalization of the Short-Time Fourier Transform (STFT) that allows for the analysis of non-stationary signals at multiple scales. Key features of CWT are time frequency analysis and filtering of time localized frequency components. The mathematical equation for CWT is given below:

$$C(a,\tau)=\int \frac{1}{\sqrt{a}}\psi\left(\frac{t-\tau}{a}\right)\ x(t)\ dt \tag{6}$$

Where $C(a,\tau)$ is the function of the parameter a,τ .

The parameter a is the dilation of wavelet (scale) and τ defines a translation of the wavelet and indicates the time localization, $\psi(t)$ is the wavelet. The coefficient $\frac{1}{\sqrt{a}}$ is an energy normalized factor (the energy of the wavelet must be the same for different a value of the scale).

IV. RESULTS & DISCUSSION

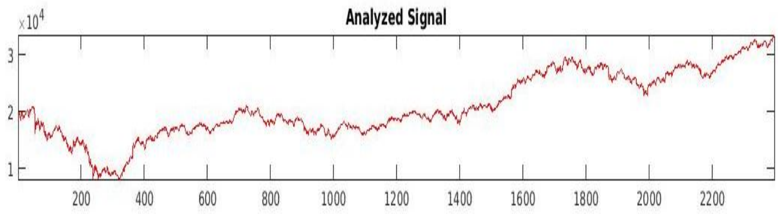
The values of Hurst exponents for the two time series like NI and SI has been calculated using the two methods, GHE and HFD which are being tabulated below in TABLE I.

Table I: Hurst Parameter Values For Ni & Si Data

Methods	Hurst exponent (H)	
	NI	SI
HFD	0.2787	0.2435
GHE	0.2452	0.1552

Value of the Hurst exponents for both the series is less than 0.5. The Hurst exponent for SI is lower than that of the NI. These results state the anti-persistent behavior of both of them i.e. their future values have the tendency to revert to their long-term mean with the SI profile has more tendency to come back to its mean compared to the NI profile. As there are the tendencies for both the profiles to back again to their respective mean, it can be said that there must be some moving forces which bring back the series towards their means when the profiles deviate from the mean. This signifies that some negative feedback system must be carrying out which constantly try to stabilize the systems. However these low values of H signify that both the signals have short-range dependent (SRD) memory. The self-similar nature in short scale for both the times series is apparent from this SRD phenomenon of them.

The CWT based time-frequency spectrum for the two time series are shown in Fig.1 and Fig.2 respectively.



IJIREEICE

ISSN (Online) 2321-2004
ISSN (Print) 2321- 5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified

Vol. 6, Issue 7, July 2018

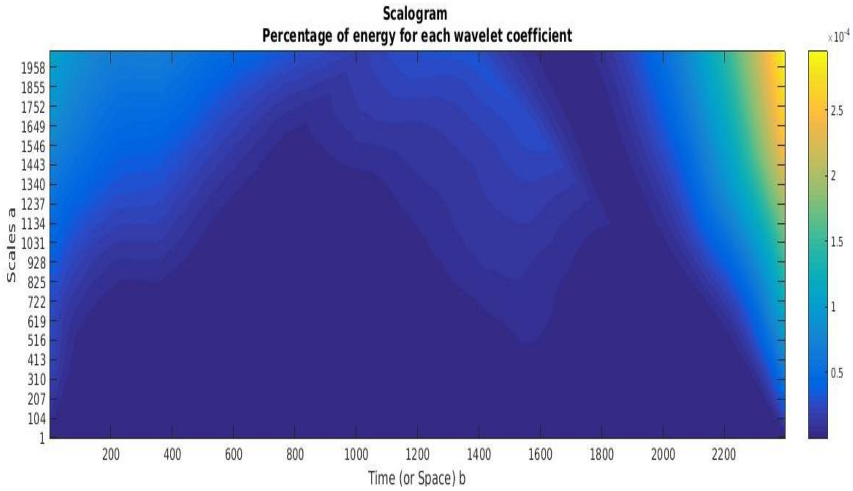


Fig.1 CWT analysis for South Indian rain fall data

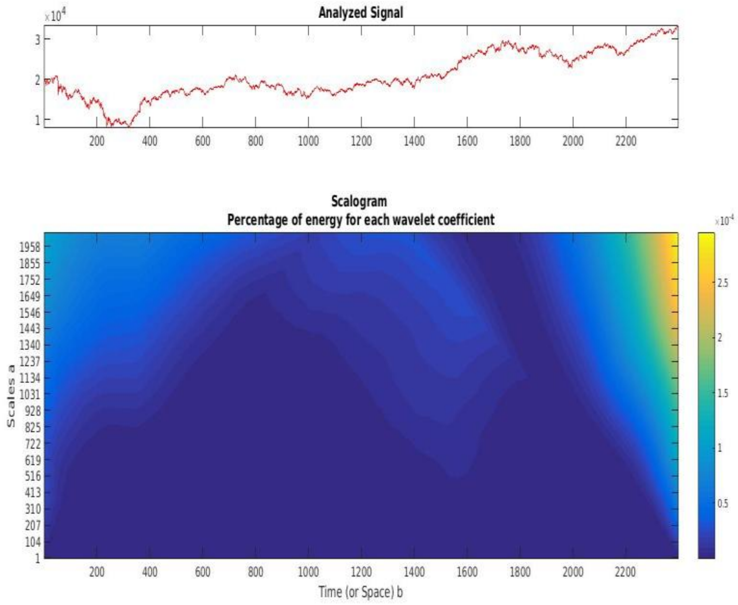


Fig.2 CWT analysis for North Indian rainfall data

From Fig.1 and Fig.2 the scalogram obtained from the given data, it can thus be inferred that the small CWT coefficients dominate the majority of the time-scale plane and also that density of the various CWT coefficients vary considerably all through, giving an indication of non-stationarity in the input time series. Relatively small portions of the scalogram are occupied by the intermediate sized CWT coefficients towards the beginning and large sized CWT coefficients towards the end. Both the time series of stock market data are non stationary in nature.

V. CONCLUSION

The value of Hurst Exponent of any method is greater than 1/2 or less than 1/2 but not equal to 1 are normally being believed to confirm nonlinear dynamics. It can be concluded to consider that both time series are nonlinear in their dynamics as the Hurst value less than 1/2. In addition the anti-persistent behavior of average value of rain fall of both NI and SI give the outline of the existence of some negative feedback system which needs to be exposed more in the successive works. The low value of H denotes more steadiness of the SI than that of the NI. Again it is found that there is non-Stationary in both the time series. So, it can be concluded that the NI and SI are not a random trend rather it is much more complex and non-linear, stable process. As the rain fall value of NI and SI are used as the vital figures of worth to evaluate the economical stability of country, the consequent works will be to explore the nature of the non-linear dynamics and originate the model depending on the present work findings.

IJIREEICE

ISSN (Online) 2321-2004
ISSN (Print) 2321- 5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2018

REFERENCES

[1]. Trends in the Rainfall Pattern over India, Guhathakurata, P., Rajeevan, M., National Climate Centre, Research Report No.: 2/2006, 2006.
[2]. All-India monthly and seasonal rainfall series 1887-1993, Parthasarathy, B., Munot A, Kothawale D R, Theoretical and Applied Climatology, 49, pp 217-224, 1994.
[3]. Inter-annual and long term variability of Indian summer monsoon rainfall, Parthasarathy, B., Proc. Indian Acad. Sci. (Earth Planet Sci.), 93, pp 371-385, 1984.
[4]. Characteristic Features of Hourly Rainfall in India, Deshpande, N. R., Kulkarni, A., Krishna Kumar, K., International Journal of Climatology, 32, 1730-1744, 2012.
[5]. Analysis of variability and trends of extreme rainfall events over India using 104years of daily gridded rainfall data, Rajeevan, M, Bhate, J, Jaiswal, AK, Geophical Research Letters, 35, L18707, pp 1-6, 2008.
[6]. The Hurst Exponent: Predictability of Time Series. Mansukhani, S. s.l. : Predictability of Time Series, pp. 29-46, 2012.
[7]. The characteristics of wireless network traffic : fractality and stationarity. Mukherjee, S, et al. [ed.] ICRCICN 17. Kolkata : IEEE, 2017.
[8]. Approach to An Irregular Time Series on the Basis of the Fractal Theory. Higuchi, T. 1988, Physica D, Vol. 31, pp. 277–283.
[9]. Use of the Higuchi's fractal dimension for the analysis of MEG recordings from Alzheimer's disease patients. Gómez, C., et al. 3, Medical Engineering Physics, Vol. 31, pp. 306-313, 2008.
[10]. Wairimu, M J. Features Affecting Hurst Exponent estimation on time series. Jomo Kenyatta University Of Agriculture and Technology. Juja : University of Kenya, 2013.
[11]. Long-Term Storage: An Experimental Study. Hurst, H E, Black, R P and Simaika, Y M. London: Constable : s.n., p. 145, 1965.
[12]. Fractals and scaling in finance. Mandelbrot, Benoit. New York : Springer, 1997.
[13]. Dynamical Generalized Hurst exponent as a Tool to Monitor Unstable Periods in Financial Time Series. Morales, Raffaello, Matteo, T. Di and Gramatica, Ruggero. Physica A. Vol. 391, pp. 3180–3189, 2012.
[14]. Rioul, Olivier and Vetterli, Martin. Wavelet and Signal Processing. s.l. : IEEE SP MAGAZINE, pp. 14-38, 1991.
[15]. Multiscale Activity in Physiological Signal. Shoeb, Ali and Clifford, Gari. s.l. : Biomedical Signal and Image Processing Spring 2005, 2006.