IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-12, Dec 2018 E-ISSN: 2347-2693

Undermining the Fractal and Stationary Nature of Earthquake

Bikash Sadhukhan[1*], Somenath Mukherjee[1]

[1] Department of Computer Science and Engineering, Techno International New Town, Kolkata, West Bengal, India

[*]Corresponding Author: bikash.sadhukhan@tict.edu.in, Tel.: +91-9433635474

Available online at: www.ijcseonline.org

Accepted: 04/Dec/2018, Published: 31/Dec/2018

Abstract— In this paper, an investigation has been made to detect the self-similarity and stationarity nature of magnitude of occurred Earthquake by exploring the fractal pattern and the variation nature of frequency of the essential parameter, Magnitude of occurred earthquake across the different place of the world. The time series of magnitude (19.04.1997 to 07.11.2017), of occurred earthquakes, collected from U.S.G.S. have been analysed for exposing the nature of scaling (fractality) and stationary behaviour using different statistical methodologies. Four conventional methods namely Visibility Graph Analysis (VGA), Wavelet Variance Analysis (WVA) Higuchi's Fractal Dimension (HFD) and Detrended Fluctuation Analysis (DFA) are being used for computing the value of Hurst parameter. In addition, Artificial Neural Network, a pre-trained fully connected 3-layer has finally been used to compute the Hurst parameter. It has been perceived that the specified dataset reveals the anti-persistency and Short-Range Dependency (SRD) behaviour. Binary based ADF, KPSS test and Time Frequency Representation based Smoothed Pseudo Wigner-Ville Distribution (SPWVD) test is incorporated in order to explore the nature of stationarity/non-stationarity of the specified profile, which here displays the non-stationary character of the magnitude of earthquake.

Keywords**—** Earthquake; Hurst Parameter (H); Fractality; Stationarity; Artificial Neural Network (ANN);

I. INTRODUCTION

The shear stress and the shear strain are described in terms of forces produced in Earth's shell. The shear stress is defined as the force per unit area which is functional tangent to a body whether the shear strain is the alteration of a body formed due to shear stress. New advancement in space-based geodesy, GPS and Satellite-Interferometry produces a pure outline of strain build up and crustal measure [1]. When the stress at point on the Earth's shell surpasses a critical value, an unexpected failure arises. The plane where let-down arises is known as the fault plane and the focus is define by a point where let-down starts. Naturally, an abrupt dislocation of the earth's shell near to the fault plane resulting a disaster that release elastic waves. This natural phenomenon is known as earthquake. In most of the earthquake, the dislocation happens at a prevailing geological fault, specifically, a plane which is already weak.

Earthquakes are typical instances of composite scenario which reveal scale-invariance and fractality characteristics together. Temporal, spatial and size distribution of earthquake has been taken in consideration during the research experiments of earthquakes, the property of fractality and scale variation nature are being exposed. As the Earthquake predicting is a demanding exploration topic so single prediction technique is not enough being the best. After long observation it is shows that the time series earthquake data, shows a composite pattern which is jointly an association of statistical parameters. The common goal of this kind of analysis is to differentiate the original frame of the earthquake movement in a seismic province, estimate dynamical progression by qualitatively and, quantitatively.

The rest of the paper is organized as follows, Section I contains the introduction and the physics behind the earthquake, Section II contain the related work of analysing the probability of earthquake using different predictive models, Section III contain the methodology used to expose the scaling behaviour and the time-based frequency representation of earthquake, Section IV discussed the Result by exploring the self-similarity and stationarity nature of magnitude of occurred earthquake and finally in the Section V conclude the paper and give the future scope.

II. RELATED WORK

Many researchers across the world are continuously being engaged to analyse the probability of occurrences of earthquake using the different methods and to design perfect predictive models which are discussed as follows:

A reasonable level of determination in the occurrences and strength, relapse of earthquakes in Indonesia may rise in the coming years [2]. The scaling frequency spectra are being flexible by spectacles of normal and constrained selforganized criticality, spectra of white and color rackets and rare change of Markov and non-Markov effects of long-range memory during earthquake [3]**.** The prospects of earthquake predictability research in order to realize practical operational forecasting in the near future has been analyzed [4]. The statistics and soft computing techniques to analyze the earthquake have been focused [5]. Multifractal method has been used to understand the time-based dynamics of the current earthquake in the Corinth- Rift which indicates heterogeneity in clustering and correlations in every time scale that recommend non-Poissonian performance [6]. Context of automatic predicting of the earthquake, appropriate for instantaneous earthquake monitor and an examination on how various characteristics of the data series are coupled with different prediction algorithms for the best conceivable precision has been proposed [7]. A multi-fractal distribution can also be fitted to the observations using Bayesian techniques in earth quake prediction [8]. Several variables processes having common contact and common attributes the study for earthquake occurrence has been devised [9]. Wavelet Transform Modulus Maxima Method which designated qualitatively and quantitatively the complex temporal patterns of seismicity has been described [10]. Prediction of earthquake can be achieved more accurately by using numerous neural network algorithms as proposed in [11]. Numerous Data mining algorithms are used to develop the predictive model for natural disasters that forecasts the future events using past data. Some of the methods used for prediction are classification, regression, supervised and unsupervised learning algorithms, support vector machine and artificial neural networks [12]

III. METHODOLOGY

The endeavour has been taken to expose the scaling behaviour and the time-based frequency representation of earthquake in different location of the world. Here magnitude of the earthquake dataset (19.04.1997 to 07.11.2017) has been chosen from U.S Geological Survey (https://earthquake.usgs.gov/earthquakes) which can be preserved as the signatory characteristic of any Earthquake events. Magnitudes are generally determined from dimensions of an earthquake's seismic waves as recorded on a seismogram. Based on the measurement and calculation of the type and component of the seismic waves, magnitude scales will vary. Different magnitude scales are essential because of differences in earthquakes, and in the purposes for which magnitudes are used. A concurrent study of Magnitude scale of Earthquake may give apparent scenario of the Earthquake events, the strength of the earthquake events and the pattern of the Earthquake.

The scaling behaviour of the said time series has been exposed by estimating the Hurst exponent of earthquake magnitude time series. The properties of self-similarity are shown by the estimation of the Hurst Exponent. Several methods like Visibility Graph Analysis (VGA), Wavelet Variance Analysis (WVA), Higuchi's Fractal Dimension (HFD) and Detrended Fluctuation Analysis (DFA) are used to estimate the Hurst-Exponent to realize the statistical aspects of the signals in diverse scales and to classify whether the signalling system have the characteristics of motion fractional Gaussian noise or fractional Brownian. Since such statistical signal processing techniques have a very high computation time, a pre-trained 3-layer fully connected Artificial neural network (ANN) has been used as the ultimate estimator of the Hurst parameter. The use of neural networks can be justified by their ability to learn complex input-output mappings while being extremely robust to noise. Apart from that, it has been experimentally verified that the pre-trained neural network computes the Hurst parameter at least fourteen times faster than the statistical signal processing techniques. The results obtained using the four other statistical signal processing techniques have been used for cross verification purposes. Hurst Exponent can be calculated using different methods but most of them have certain limitations. These methods, as stated, are being selected to compute the Hurst parameter and thus approve the satisfactory accuracy of the decisions derived from the results.

The study of Earthquake would be incomplete without revealing the frequency fluctuation nature of earthquake magnitude time series data. A stationary signal has persistent frequency components while the non-stationary has ever changing frequency with respect to time.

The test of stationarity/non-stationarity is required for the signal because it determines several activities and properties of the signal. Different approaches are prevailing to test the stationarity/ non-stationarity of the signals. In this paper the conventional techniques and Time-Frequency Representation (TFR) based methods are selected for identifying the stationarity and non-stationarity character of that specified profile. Binary based ADF and KPSS systems have also been used to detect that an observable time series is stationary around a deterministic trend against substitute of a unit root(i.e. trend stationary) which is known as null hypothesis. Basically, TFR method is used to traces the period of presence of frequencies and identify the features of the frequency content within the signal. Time Frequency Representation Based Smoothed Pseudo Wigner Ville Distribution (SPWVD) method is being incorporated here for revealing the scenario of the stationarity/non-stationarity of the magnitude of Earthquake. SPWVD methods are being used due to its advantageous characteristics like 1) SPWVD method can merely be employed for the reason of being close to Fast Fourier transform. 2) Furthermore, when it is run on a finite time series gesture, the method produces a finite integration If the self-similar time series signal belongs to non-stationary, then it exhibits Fractional Brownian Motion (fBm) and if it explores stationarity then it is called Fractional Gaussian Noise (fGn).

1. Experimental Dataset:

Primarily the magnitude of Earthquake (Occurred in different location of the world) datasets has been taken from U.S Geological Survey (https://earthquake.usgs.gov/ earthquakes). Magnitude scales will vary based on the various kinds of the seismic waves which were measured and calculated. Several magnitude scales are required because of differences in earthquakes, and in the purposes for which magnitudes are used. The various magnitude scales correspond to several ways of deriving magnitude. All magnitude scales preserve the logarithmic scale as devised by Charles Richter, are used to mid-range approximately which correlates with the original "Richter" scale. The earthquake of magnitude 5.7-5.9 is moderate which causes slight or no damage, the magnitude of 6-6.9 are strong may cause a lot of damage in very populated areas, magnitude of 7-7.9 are major which causes serious damage and magnitude 8 or more are great which can totally destroy communities near the epicentre.

So, in this work magnitude of Earthquake event whose values are greater than 5.7 and duration of 19.04.1997 to 07.11.2017 have been considered to investigate the scaling pattern as well as nature of frequency fluctuation of the earthquake.

The summary statistics of the experimental data sets are tabulated in Table1 which basically provide the statistical outline of the signals under examination in this work and Fig.1 depicts the plotting of dataset respectively.

Table-1: Summary statistics for magnitude of Earthquake dataset.

SCORES	MAGNITUDE
Mean	6.2374
Median	6.1000
Mode	5.8000
Standard Deviation	0.4745
Variance	0.2251
Maximum	9.1000
Minimum	5.7000
Skewness	1.6410
Kurtosis	6.2506

Table 1 indicates that the mean of the magnitudes is 6.2374 which is strong earthquake. The mean may not be a reasonable picture of the data, because the average is easily inclined by outliers. The median is another method to quantify the centre of a numerical data set. The median is 6.1 which also indicates a strong earthquake. The mode is 5.8 which indicates that most often the earthquakes are nearly of magnitude 5.8 which is a moderate earthquake. The standard deviation of the dataset is low which specifies that the data points incline to be adjacent to the mean of the set and not banquet out over a broader range of standards. The variance signifies that how far are the magnitudes in the dataset from the mean. Skewness is typically defined as a degree of symmetry or asymmetry in a data set. Symmetric data set will have a skewness of 0. Kurtosis is a measure of the collective sizes of the two tails. It measures the amount of probability in the tails. The value is often related to the kurtosis of the normal distribution, which is equal to 3. Here the kurtosis is greater than 3, which means the dataset includes heavier tails than a normal distribution (more in the tails).

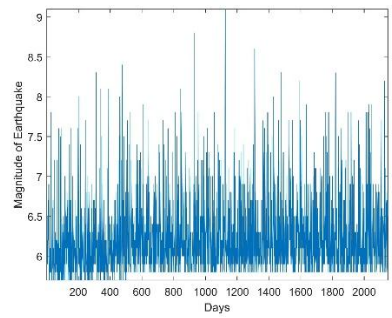

Fig.1: Plot of Magnitude of Earthquake dataset.

2. Exponential Smoothing of Raw Data.

Exponential smoothing is a technique for smoothing time series data using a mathematical function whose value becomes zero outside the chosen interval. The exponential function assigns exponentially decreasing weights over time.

It is an exponential function which is generally applied to smooth the time series data in signal processing, functioning as low-pass filters to remove noisy signals. It learns easily and applies easy procedures for making some determinations based on prior assumptions by the user. It is mostly used for investigation of time-series data.

The simple formula of basic form for exponential smoothing is given by:

$$s_t = \alpha y_t + (1-\alpha)\, s_{\{t-1\}}, t>0 \qquad (1)$$

Where, α is the smoothing constant and $0< \alpha <1$

t is the time

s_t is the statistic(smoothed) of the current observation y_t.

Here the value of α can be a value ranging from 0 to 1, larger values of α reduces the smoothing level and when its value is 1, then the output series is the current observation. So, when α's value is nearly 1, it has lesser smoothing effect while when the values of α are close to 0, it has greater smoothing effect. The best value of α can be found out by method of least squares, that is, the value of α at which $(s_t - y_{t+1})^2$ is minimized.

The following plot shows the exponential smoothing for the time series data, that is, of the magnitude of earthquakes:

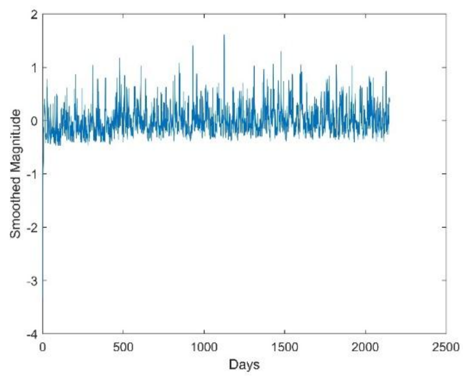

Figure 2: Plot of Smoothed Magnitude of Earthquake dataset.

3. Hurst Exponent Estimation

The Hurst exponent has wide applications in numerous fields of applied mathematics, including chaos theory and fractals, spectral analysis and long memory processes. In hydrology, approximation of the Hurst parameter was formerly developed. Nevertheless, the current procedures for approximating the Hurst parameter originated from fractal concepts. It defines the autocorrelations concepts used in time series signal and the proportion at which it is being reduced as the lag between pairs of values rises. It reflects the comparative inclination of a signal either to revert powerfully to the mean or to gathering in a track. Hurst analysis exposes the scaling nature of a time series Y(i), the time series having length (i) for a sampling interval (t). Through Hurst parameter (H) the fractal behaviour of this signal is numerically signified. The scale of this specified parameter is fluctuating between 0 and 1. Computing the standards of Hurst Parameter, the features within the signal can be predicted. At, $H = 0$ signifies the system having white noise whose autocorrelation function reduces fast with interval, value of H in the series of 0 – ½ specifies a signal with long term basis, converting among the high and low values in nearby sets, a low value possibly trailing a high value and the next value in turn will most probably tend to be high. The range $0 < H < 0.5$ specifies that a time series is anti-persistent in nature i.e. increase in the value is followed by a decrease or vice-versa, which means the series has negative autocorrelation. The values change with respect to a mean, which indicates that new values possess an inclination to convert to a mean on long-term basis and when the range $0.5 < H < 1$ it shows that a time series is persistent in nature, that is, it has an autocorrelation which is positive. In a persistent time, series, with an increase in values the short-term increases, and it decreases with the decrease in the values of the persistent time series data.

Among different estimators five Hurst estimation methods like Visibility Graph Analysis (VGA), Wavelet Variance Analysis (WVA), Higuchi's Fractal Dimension (HFD), Detrended Fluctuation Analysis (DFA) and Artificial Neural Network (ANN) has been used to accurately determine the persistency or anti persistency nature of the time series data.

The VGA is initiated to be vigorous sufficient to invent long range dependency as VGA is independent on the fractal scale and in the visibility graph the power law degree distribution's exponent is linearly dependent on the Hurst Exponent. Wavelet *analysis has a benefit* to decompose macroeconomic signal, and overall data into their time scale mechanisms. Higuchi's method has been preferred as it is efficient, highly accurate and not very sensitive to artefacts. DFA is used here due its high accuracy for computing the Hurst parameter value. The use of neural network can be justified by its relatively small computation time compared to other signal processing techniques and it is fast and more accurate.

METHODS USED:

3.1 Visibility Graph Analysis (VGA) Method

A visibility graph [13] [14] [15] is attained by mapping time series data into a network corresponding to the subsequent perceptibility standard. Taking two random data points (t_a, y_a) and (t_b, y_b) in the time signal have visibility, and therefore develop two linked nodes in the allied graph, if any other data (t_c, y_c) such that $t_a < t_c < t_b$ satisfies:

$$y_c < y_a + (y_b - y_a)\frac{t_c - t_a}{t_b - t_a} \quad (2)$$

Now points y_a and y_b is linked as two nodes to create a graph. In this manner the visibility graph is generated by tracing and connecting the perceptibility of similar data points of the supposed signal. The behaviour or characteristics of the signal is determined by the consistency allocation of the visibility graph. It means a consistent graph designates an interrupted sequence. however, an arbitrary graph suggests for an arbitrary signal. Correspondingly, a scale-independent graph is attained for a scaling sequences with the distribution of power law dependent, $P(x) = y^{-\alpha}$. The α exponent is well-defined as:

$$\alpha = 1 + N\left[\sum_{i=1}^{N} \log \frac{y_i}{y_{min}}\right]^{-1}. \quad (3)$$

Here, N is the total numbers of values, y_i $(i = 1, \dots\dots\dots, N)$ is the signal and y_{min} is the smallest value of y having power law behaviour. α is linearly dependent on the Hurst Exponent (H)

$$H = 1 - \frac{\alpha}{2} \quad (4)$$

3.2 Wavelet Variance Analysis (WVA) Method

This traditional process [16] [17] [18] [19] can be illustrated as follows:

In Wavelet Variance Analysis (WVA), a given time series is expanded on an orthogonal base formed by shifts and the multiresolution copies of the wavelet function. Base functions $\psi(t)$ are called wavelets if they are defined in a region of complex valued functions with restricted energy that oscillate about an abscissa axis, rapidly converge to zero and have a vanishing moment of the first order. At a given mother wavelet, ψ, the detailing coefficients $d(j,k)$ are given by:

$$d(j,k) = \int_{-\infty}^{\infty} X(t)\psi_{j,k}(t)dt \quad (5)$$

This method uses the fact that the averaged squared values of the wavelet coefficients given by:

$$E_j = \frac{1}{n}\sum_{k=1}^{n_j} |d_x(j,k)|^2 , \quad (6)$$

obey the scaling law:

$$E_j \sim 2^{(2H-1)j}, \quad (7)$$

where H is the Hurst exponent.

3.3 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 compute the fractal dimension D of a time series. This technique is proposed by Higuchi [20]. To observe the fractality, HDF is calculated [21], [22] for the discrete data point series:

$$S : S(1), S(2), S(3), \ldots, S(n) \quad (8)$$

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.

$$S_\kappa(d) : S(\kappa), S(\kappa+d), S(\kappa+2d), \ldots.. S\left(\kappa + \left\lfloor \frac{n-\kappa}{d} \right\rfloor d\right) \quad (9)$$

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

$$L_\kappa(d) = \frac{1}{d}\left\{\left(\sum_{i=1}^{\frac{n-\kappa}{d}} \left|S(\kappa+id) - S(\kappa+(i-1)d)\right|\right)\frac{n-1}{\left\lfloor \frac{n-\kappa}{d} \right\rfloor d}\right\} \quad (10)$$

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

$$L(d) = \frac{1}{d}\sum_{i=1}^{d} L_\kappa(d) \quad (11)$$

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

3.4 Detrended Fluctuation Analysis (DFA):

Detrend fluctuation analysis [23] [24] [25] is a stochastic method for determining the self-similarity of a time series signal. It is used to estimate long-range power-law correlation exponents in a time series that exhibit diverging correlation time (long memory process).

The attained exponent is analogous to the Hurst parameter but DFA may also be applied to non-stationary signals. The Cumulative sum X_t of a bounded time series x_t can be represented as

$$X_t = \sum_{i=1}^{t}(x_i - \langle x \rangle) \quad (12)$$

Where N is the length of the time series, $\langle x \rangle$ is the mean value of the time series and

$t \in \mathbb{N}$ integration or summation first converts this into an unbounded process X_t.

Then X_t is separated into time windows of length n samples each and local trend is computed by minimizing the squared errors within each time window. If Y_t is the local trend of each time window, then the fluctuation is computed as

$$F(n) = \sqrt{\frac{1}{N} \sum_{t=1}^{N} (X_t - Y_t)^2} \quad (13)$$

A log-log graph of $F(n)$ is created by reiterating the process of detrending followed by fluctuation analysis over a range of diverse window size n. A straight line generated in this graph signifies statistical self-affinity denoted as

$$F(n) \propto n^{\alpha}. \quad (14)$$

Where α is the scaling exponent represents the angle of the straight line. The scaling exponent α is the generality of Hurst parameter. The parameter value is 0.5 signifies uncorrelated white noise, the value between 0 and 1 results fractional Brownian motion.

3.5 Artificial Neural Network (ANN):

Here, as suggested in [26] a series of moving average smoothing computations has been used for extracting features from an input stochastic time series. The given input time series is passed through a series of 10 moving average smoothing computations, wherein for each stage, the power of the smoothed time series has been used as the corresponding feature. The power of a time series is given by the following:

$$P = \frac{1}{N}\sum_{i=1}^{N}|x_i|^2 \quad (15)$$

The architecture proposed in this paper is a 3 layer fully connected neural network with one output layer and two hidden layers (shown in Figure 3). For the hidden layers the activation function used is Leaky ReLU [27], which, along with introducing non-linearity to the computed hypothesis, also successfully combats the problem of vanishing gradients. The Leaky ReLU activation function is given by:

$$f(x) = \max\{(leak\_factor \times x),\ x\} \ , \quad (16)$$

The activation function used in the output layer is Sigmoid [28], which thresholds the value of the network's output between 0 and 1. The Sigmoid activation function is given by:

$$f(x) = \frac{1}{1+e^{-x}} \quad (17)$$

Finally, as the Hurst Parameter is a continuous function of the input stochastic time series, varying between 0 and 1, a regression based learning procedure has been adopted, by using a Mean Squared Error (MSE) Loss [29] as the objective function for the neural network learning. The MSE Loss function is given by the following:

$$Loss = \frac{1}{2m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2 \quad (18)$$

RMSProp [30] is the optimization algorithm used to minimize the loss and to learn the optimal network parameters. The update rule for RMSProp is given by:

$$E[g^2]_t = 0.9E[g^2]_{t-1} + 0.1g_t^2 \quad (19)$$

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{E[g^2]_t + \epsilon}} g_t \ , \quad (20)$$

Once trained, the neural network can compute the Hurst parameter at least fourteen times faster than the other conventional statistical signal processing techniques.

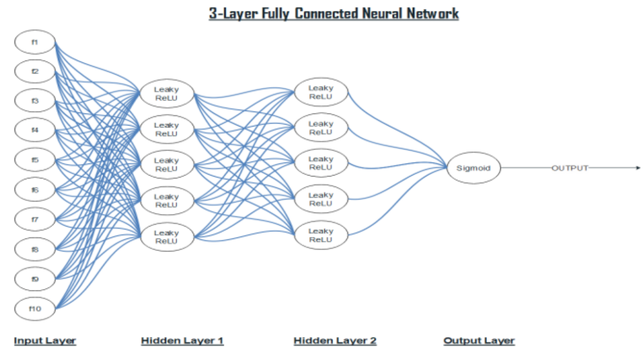

Figure 3. Neural Network Architecture

4. Detection of Stationarity/ Non-Stationarity

4.1 Augmented Dickey Fuller (ADF) Test:

Dickey Fuller Test [31] was evolved in 1979 by two statisticians- David Dickey and Wayne Fuller. In 1984, for assisting more complicated models having unspecified orders, they extended their elementary autoregressive unit root test. The Augmented Dickey Fuller test, in comparison to Dickey fuller test, is the one that is used for testing a unit root in a model of time series. Augmented Dickey Fuller statistic that is being used in the ADF test is a negative number. The more negative the statistic will be, the stronger will be the probability of the hypothesis rejection that there is a unit root. Here, it is essential to check whether the time series data set has the existence of stationary property for each variable. ADF test [32] is a unit root test which helps in time series analysis to test stationarity. ADF test uses a parametric autoregressive structure to obtain the presence of any serial correlation. For a substantial and complex set of models of time series it can be identified as an augmented version of the Dickey - Fuller test.

The testing process for the ADF test is being applied to the model [33]

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \cdots + \delta_{p-1}\Delta y_{t-p+1} + \varepsilon_t \quad (21)$$

Hypothesis testing involves the watchful construction of two statements- alternative hypothesis and null hypothesis. The alternative and null hypotheses are two mutually exclusive statements about a population. The null hypothesis of the Augmented Dickey Fuller t-test is that H_0 implies the existence of a unit root and signifies the non-stationarity of the time series data. It needs to be differenced to make it stationary and the alternative, H_A signifies that the data is stationary.

If data is potentially slow-turning around a trend line and there is a trend visible in the time series, then the equation used has an intercept term as well as a time trend,

$$\Delta z_t = \alpha_0 + \theta z_{t-1} + \gamma t + \alpha_1 \Delta z_{t-1} + \alpha_2 \Delta z_{t-2} + \cdots + \alpha_p \Delta z_{t-p} + \alpha_t \quad (22)$$

By minimizing the Schwartz Bayesian information criterion, the number of augmenting lags (p) can be established. If ‘p’ is too large, the power of the test will suffer and if ‘p’ is too small then we will get a biased test.

4.2 Kwiatkowski-Phillips-Schmidt-Shin (KPSS) Test

Kwiatkowski–Phillips–Schmidt–Shin *(KPSS) tests* [34] are *used* for *testing* a null hypothesis to check whether the observable time series is *stationary* or termed stationary or is non-stationary. This test is used as a complement to the standard tests in analyzing time series properties. The KPSS test is based on linear regression. The time series is broken down into three parts: a random walk (r_t), a stationary error (ε_t), and a deterministic trend (βt) with the regression equation:

$$X_t = r_t + \beta t + \varepsilon_t \quad (23)$$

If the data is stationary, it will have a fixed element for an interceptor the series will be stationary around a fixed level [35]. The test uses Ordinary Least Square (OLS) to_find the equation, which differs slightly depending on whether you want to test for level stationarity or trend stationarity. A simplified version is used to test level stationarity without the time trend component.

4.3 Smoothed Pseudo Wigner-Ville Distribution (SPWVD) Method:

The Wigner Ville Distribution method mainly highlights the energy distribution of the signal. The main goal of distributing energy of the signal over frequency and time is to increase the firmness of the time-frequency spectrum.
Wigner-Ville Distribution is described as:

$$W_x(t,f) = \int_{-\infty}^{\infty} z\left(t + {}^{\gamma}/_{2}\right) z^*\left(t - {}^{\gamma}/_{2}\right) e^{-j2\pi f\gamma} d\gamma. \quad (24)$$

Where z is an analytic signal which is derived from real signalling system.
Now, in term of spectrum WVD is defined as

$$W_x(t,f) = \int_{-\infty}^{\infty} Z\left(f + {}^{\varepsilon}/_{2}\right) Z^*\left(\omega - {}^{\varepsilon}/_{2}\right) e^{-j2\pi\varepsilon\gamma} d\gamma \quad (25)$$

Here * indicates a complex conjugate.

Now to eliminate the interference which represents cross terms or the frequency components which do not exist Cohen’s class Distribution Method [32] is to be applied.

Class distribution of The Cohen’s method is defined as,

$$C_x(t,v;f) = \iiint f(\epsilon,\gamma)\, e^{-j2\pi\varepsilon(m-t)} x\left(m + {}^{\gamma}/_{2}\right) x^*\left(m - {}^{\gamma}/_{2}\right) e^{-j2\pi v\gamma} d\varepsilon\, dm\, d\gamma. \quad (26)$$

Here $f(\epsilon,\gamma)$ is called ‘kernel’ because it is a two- variables parameterization function.
The equation 15 can be inscribed as

$$C_x(t,v,f) = \iint f(\varepsilon,\gamma)\, A(\epsilon,\gamma) e^{-j2\pi}(v\gamma + \epsilon t) d\varepsilon\, d\gamma \quad (27)$$

Where the ambiguity function can be defined as

$$A(\epsilon,\gamma) = \int x\left(m + {}^{\gamma}/_{2}\right) x^*\left(m - {}^{\gamma}/_{2}\right) dm \quad (28)$$

If $f(\varepsilon,\gamma) = 1$, we get the Wigner-Ville distribution.

$$C_z(t,f) = \int z\left(t + {}^{\gamma}/_{2}\right) z^*\left(t + {}^{\gamma}/_{2}\right) e^{-j2\pi f\gamma} d\gamma \quad (29)$$

where z denotes the analytic part of the original signal.
The Fourier transform of the ‘kernel’ $f(\varepsilon,\gamma) = G(\varepsilon)h(\gamma)$ is given by

$$F(t,v) = FT[f(\varepsilon,\gamma)] = g(t)H(v). \quad (30)$$

The function $h(\gamma)$ denotes the frequency smoothing and the $g(t)$ indicates the temporal smoothing.
SPWVD can be found by replacing the kernel function in the general expression

$$SPWVD = \int h(\gamma) \left[\int g(m-t) x\left(m + {}^{\gamma}/_{2}\right) x^*\left(m - {}^{\gamma}/_{2}\right) dm\right] e^{-j2\pi v\gamma} d\gamma \quad (31)$$

The following kernel function is to be chosen to filter out the interferences independently in the two axis directions and provides more flexibility to the smoothing operation.

$$f(\varepsilon,\gamma) = e^{-\left(\frac{\pi\varepsilon}{\sqrt{2}\sigma_\varepsilon}\right)^2} e^{-\left(\frac{\pi\gamma}{\sqrt{2}\sigma_\gamma}\right)^2} \quad (32)$$

Now substituting this expression of kernel in equation (26) we get

$$SPWVD(t,v) = \sqrt{\frac{2}{\pi}} \int e^{-\left(\frac{\pi\varepsilon}{\sqrt{2}\sigma_\varepsilon}\right)^2} B\, e^{-j2\pi v\gamma} d\gamma \quad (33)$$

Therefore, SPWVD is represented by its the discrete form

$$SPWVD(r,s) = 2\sqrt{\frac{2}{\pi}}\sum_{k=-h}^{h} e^{-\left(\frac{2\pi k}{\sqrt{2}\sigma_{\gamma}}\right)} \sum_{u=-g}^{g} \sigma_{\varepsilon}\, e^{-2(\sigma_{\varepsilon}u)^2}\, x(r+u+k)x(r+u-k)e^{-j4\pi sk} \quad (34)$$

IV. RESULTS AND DISCUSSION

The Hurst exponent values of specified profile, which are computed by the specified methods are being represented below in Table 2.

Table 2: Hurst parameter values for magnitude of Earthquake

METHODS	HURST EXPONENT(H)
VGA	0.3631
WVA	0.3776
HFD	0.2990
DFA	0.3551
ANN	0.2665

It represents the value of Hurst exponents is less than 0.5 which indicates the anti-persistency of the specified earthquake time series data. It also indicates that the upcoming future values of that time series have the inclination to return to the value on long-term basis. As there is a probability for the profile to back again to its respective mean, it has been observed that some encouraging forces must be required which evoke the profile close to its mean when the profile depart from the mean. This signifies that some respond system must be functioning negatively which endlessly attempt to stabilize the systems. However, low values of H denote short-range dependency (SRD) which indicates the fractal behaviour in short range basis for that specified profile.

A scaling distribution of minor fault-slip regions usually triggers short-period seismic waves which is recognized from the people and seismogram records where each zone radiates short-period waves. Even though plate tectonics demonstrated clear context for explaining the long-range progressions, but the short-range activation of the processes are still ambiguous, which makes earthquakes most unpredictable.

Table 3: Non-stationarity test for magnitude of Earth Quake data

Stationarity/Non-Stationarity	
Method used	Magnitude of Earthquake
ADF	Logical 0
KPSS	Logical 1

After carrying out the ADF test for Magnitude of Earthquake shows the logical value 0 and the p-value is 0.8870 and the test statistic is 0.8132. After executing the KPSS test for same profile, result shows logical value 1 and the p-value is 0.01 and the test statistic is 13.7610. In case of ADF test, if the p-value is less than 0.05, and in case of KPSS test, the series is said to be stationary when the p-value is greater than 0.05. Result which has been derived from the both the methods may infer that the magnitude of occurred Earth Quake belongs to non-stationarity. A non-stationary time series will contain trends and seasonality that will influence the values of the time series at various times.

Figure 2 shows the result of Time Frequency Representation (TFR) based method SPWVD which has been applied on that specified data set for getting unarguable conclusion regarding the testing of stationarity/non-stationarity behaviour of that specified profile. It undoubtedly indicates that the magnitude of Earthquake's frequency is changing with respect to the time. So, the magnitude of Earthquake is non-stationary in nature.

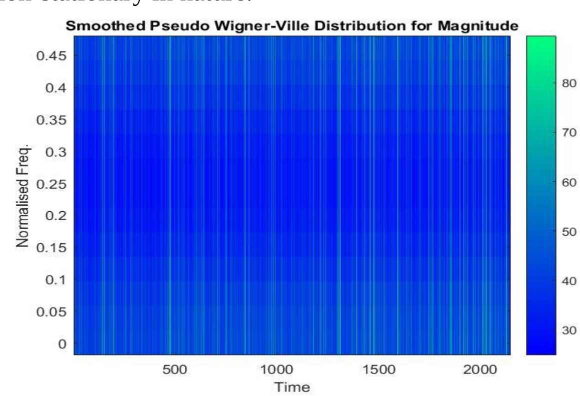

Figure 4: SPWVD for Magnitude of Earthquake.

The frequency aspects which are the functions of time and having a dependency on the oscillation of the time is referred as non-stationarity of the signal. So, it can be concluded that the magnitude of occurred earthquake per unit time is dependent on time. It was found that the stress growth rate on earth crust is inconsistent over time. Big earthquake occurred on an area of a fault, alters the pressure in the nearby areas and increases / decreases the seismic activity, based on the fault geometry. The strength of the shell is variable over time. Moreover, earth's shell is significantly weakened by migrating fluids with altering times which accelerates earthquake. The stress fall at the timing of earthquakes also fluctuates from occurrence to occurrence.

V. CONCLUSION & FUTURE SCOPE

It is believed that in a nonlinear system, Hurst Parameter is always belonging to greater than ½ or less than ½ but not equal to 1. As the Hurst parameter values of the magnitude of occurred earthquake are below ½, it can be resolved they are nonlinear in nature. Both binary based ADF, KPSS test and TFR based SPWVD test reflects non-stationarity nature of the signal. Moreover, the anti-persistent behaviour gives outline of some non-positive response system which desires to be discovered further in the consequent works. So, it can be decided that the earthquake is not an arbitrary spectacle rather it is much more composite, non-linear and steady process. As process is long-term and consistent, but substantial progressive disparities of seismicity arise, which makes exact forecast of earthquakes very problematic. From the view of geophysics and economy problems a time series is more important. Future scope is to have a noticeable understanding of earthquakes and develop several other advanced methods for predicting and forecasting the magnitude of forthcoming earthquakes.

REFERENCES

[1] E. E. B. Hiroo Kanamori, "The Physics of Earthquakes," Physics Today, p. 34, June, 2001.
[2] E. Priyadarshini, "An Analysis of the Persistence of Earthquakes in Indonesia using Rescaled Range," Indian Journal of Science and Technology, vol. 9, no. 21, pp. 1-, 2016.
[3] R. Yulmetyev, F. Gafarov, P. Hanggi, R. Nigmatullin and S. Kayumov, "Possibility between earthquake and explosion seismogram differentiation by discrete stochastic non-Markov processes and local Hurst exponent analysis," PHYSICAL REVIEW E, vol. 64, no. 066132, pp. 1-14, 2001.
[4] Y. Ogata, "A Prospect of Earthquake Prediction Research," Statistical Science, vol. 28, no. 4, pp. 521-541, 2013.
[5] G. Preethi and B. Santhi, "Study on Techniques of Earthquake Prediction," International Journal of Computer Applications, vol. 29, no. 4, pp. 0975 - 8887, 2011.
[6] G. Michas, P. Sammonds and F. Vallianatos, "Dynamic Multifractality in Earthquake Time Series: Insights from the Corinth Rift, Greece," Pure and Applied Geophysics, vol. 172, no. 7, p. 1909–1921, 2015.
[7] S. Fong and Z. Nannan, "Towards an Adaptive Forecasting of Earthquake Time Series from Decomposable and Salient Characteristics," in PATTERNS 2011 : The Third International Conferences on Pervasive Patterns and Applications, 2011.
[8] J. B. Panduyos, F. P. Villanueva and R. N. Padua, "Fitting a Fractal Distribution on Philippine Seismic Data: 2011," SDSSU Multidisciplinary Research Journal, vol. 1, no. 1, pp. 50-58, 2013.
[9] P. k. Dutta, O. P. Mishra and M. K. Naskar, "A Review of Operational Earthquake Forecasting Methodologies Using Linguistic Fuzzy Rule-Based Models From Imprecise Data With Weighted Regression Approach," Journal of Sustainability Science and Managemen, vol. 8, no. 2, pp. 220-235, 2013.
[10] B. Enescu, Z. R. Struzik and K. Ito, "Wavelet-Based Multifractal Analysis Of Real And Simulated Time Series Of Earthquakes," Annuals of Disas. Prev. Res. Inst., Kyoto Univ., pp. 1-14, 2004.
[11] K. S. K. Mohankumar, "A Study on Earthquake Prediction Using Neural Network Algorithms," International Journal of Computer Sciences and Engineering, vol. 6, no. 10, pp. 200 - 204, 2018.
[12] C. Shaganpreet kaur, "Comparative Analysis of Data Mining," International Journal of Computer Science and Engineering , vol. 6, no. 4, pp. 301-304, 2018.
[13] S. Mukherjee, R. Ray, R. Samanta, K. H. Moffazal and G. Sanyal, "Characterisation of wireless network traffic: Fractality and stationarity," in ICRCICN 17, IEEE, Kolkata, 2017.
[14] L. Lacasa, B. Luque, J. Luque and J. Nuno, "The visibility graph: Aa new method for estimating the Hurst exponent of fractional Brownian motion," EPL (Europhysics Letters), vol. 30001, pp. 1-5, 2009.
[15] R. Albert and A. Barabasi, "Statistical mechanics of complex networks.," Rev Mod Phys, vol. 74(1), pp. 47-97, 2002.
[16] D. Percival and P. Guttorp, "Long Memory Process,the Allan Variance and Wavelets," pp. 1-15, 1994.
[17] D. Percival, "Estimation of wavelet variance," pp. 619-631, 1995.
[18] D. Percival and D. Mondal, "M-estimation of wavelet variance," Elsevier, pp. 623-657, 2012.
[19] R. Ray, M. H. Khondekar, K. Ghosh and A. K. Bhattacharjee, "Memory persistency and nonlinearity in daily mean dew point," 2015.
[20] T. Higuchi, "Approach to an irregular time series on the basis of the fractal theory," Physica D: Nonlinear Phenomena, vol. 31, pp. 277-283, 1988.
[21] C. Gomez, A. Mediavilla, R. Hornero, D. Abasolo and A. Fernandez, "Use of the Higuchi's fractal dimension for the analysis of MEG recordings from Alzheimer's disease patients.," Medical engineering & physics, vol. 31, no. 3, pp. 306-13, 4 2009.
[22] M. J. Wairimu, "Features Affecting Hurst Exponent estimation on time series," in Jomo Kenyatta University Of Agriculture and Technology., Juja, 2013.
[23] K. Hu, P. C. Ivanov, Z. Chen, P. Carpena and H. Eugene Stanley, "Effect of trends on detrended fluctuation analysis," Phys. Rev. E, vol. 64, no. 1, p. 011114, 6 2001.
[24] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley and A. L. Goldberger, "Mosaic organization of DNA nucleotides," Phys. Rev. E, vol. 49, no. 2, pp. 1685-1689, 2 1994.
[25] J. W. Kantelhardt, E. Koscielny-Bunde, H. H. A. Rego, S. Havlin and A. Bunde, "Detecting long-range correlations with detrended fluctuation analysis," Physica A: Statistical Mechanics and its Applications, vol. 295, pp. 441-454, 2001.
[26] S..Ledesma-Orozco, J. Ruiz-Pinales, G. García-Hernández, G. Cerda-Villafaña and D. Hernández-Fusilier, "Hurst Parameter Estimation Using Artificial Neural Networks," Journal of Applied Research and Technology, vol. 9, p. 15, 2011.
[27] A. L. Maas, A. Y. Hannun and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," in ICML, 2013.
[28] T. Kocak, "Sigmoid Functions and Their Usage in Artificial Neural Networks," https://excel.ucf.edu, 2007.
[29] W. contributors, "Mean squared error," Wikipedia, The Free Encyclopedia, 2018.
[30] T. Tieleman and G. Hinton, "Lecture 6e rmsprop: Divide the gradient by a running average of its recent magnitude," Coursera, 2012.

[31] "The Augmented Dickey-Fuller Test," [Online]. Available: https://www.thoughtco.com/the-augmented-dickey-fuller-test-1145985.
[32] H. B. Nielsen, "Non-Stationary Time Series and Unit Root Tests," [Online]. Available: http://www.econ.ku.dk/metrics/econometrics2_05_ii/slides/08_unitroottests_2pp.pdf.
[33] "Augmented Dickey-Fuller Unit Root Tests," [Online]. Available: http://faculty.smu.edu/tfomby/eco6375/bj%20notes/adf%20notes.pdf.
[34] D. Kwiatkowski, P. C. B. Phillips, P. Schmidt and Y. Shin, "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?," Journal of Econometrics, vol. 54, pp. 159-178, 1992.
[35] W.Wang, Nonlinearity and Forecasting of Streamflow Processes., Stochasticity, 2006.

Authors Profile

Mr. Bikash Sadhukhan pursed Bachelor of Technology from University of Kalyani, West Bengal in 2004 and Master of Technology from West Bengal University of Technology in the year 2008. He is currently working as Assistant Professor in Department of Computer Science and Engineering, Techno International New Town (Formerly known as Techno India College of Technology), Kolkata since 2007. He has 13 years of teaching experience. His research interest include Statistical Signal processing, Data analytics, Wireless Networks, IoT and its Securty,

Dr. Somentah Mukherjee pursed Bachelor of Engineering from Burdwan University, West Bengal in 2006 and Master of Technology from National Institute of Technology, Durgapur, West Bengal in the year 2010. He has completed his Ph.D from National Institute of Technology, Durgapur, West Bengal in the year 2018. He is currently working as Assistant Professor in Department of Computer Science and Engineering, Techno International New Town (Formerly known as Techno India College of Technology), since 2017. His research interest include Wireless Network, Statistical Signal processing, Data Science etc.