REVIEW OF RESEARCH
IMPACT FACTOR : 5.7631(UIF) UGC APPROVED JOURNAL NO. 48514 ISSN: 2249-894X
VOLUME - 7 | ISSUE - 12 | SEPTEMBER - 2018

ESTIMATION OF HURST EXPONENT OF WEATHER DATA OF AURANGABAD CITY

Zakde K. R.[1], Munde S. V.[2], Khan A. R[3] and Yusuf H Shaikh[3*]
[1]Jawahrlal Nehru Engineering College Aurangabad (MS) India.
[2]PES Engineering College, Aurangabad (MS) India.
[3]Maulana Azad College, Aurangabad (MS).
[3*]Shivaji College, Kannad, Dist. Aurangabad (MS) India.

ABSTRACT:

Estimating of the Hurst exponent for experimental data plays a very vital role in the research of progressions which show properties of self-similarity. There are numerous techniques for estimating the Hurst exponent using time series. Random data share market data like BSE that possesses a high degree of unpredictability and unpredictability can be analysed using R/S analysis and Hurst exponent H. There are many techniques to predict the Hurst exponent using the time series. This paper presents comparative analysis of the statistical properties of the Hurst exponent estimators obtained by different methods using model stationary and non-stationary fractal time series. Hurst exponent is revealing of the presence of noise or fluctuations in the data which in turn gives information on the volatility of the Close value and turbulence in the close value. The results of R/S analysis as applied to weather data of Aurangabad city indicate that there is persistence in the trends and entire period studied as the value of H likes between 0.5 and 1.0. This also confirms long term memory effect that the trends tend to continue.

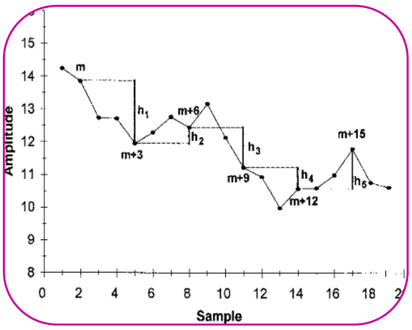
KEYWORDS – *Time series analysis, Hurst exponent, self- similarity, fractal analysis.*

1.INTRODUCTION:

The progress of mathematical models, methods, and algorithms is required for the improvement and employment of modern electronics, radio electronics, control theory and image processing problems. Currently it has been generally accepted, many stochastic processes appear in nature and engineering in long-term dependence and fractal [1-3] structure. The most useful mathematical technique for mobility and the structure of such a process is a fractal analysis.

In this paper the most commonly used method for estimating the Hurst exponents is examined. The various methods are: R / S -analysis, variance-time analysis, Detrended Fluctuation Analysis (DFA) and wavelet-based estimation [4].

Many years later, while investigating the fractal nature of financial markets noted mathematician Benoit Mandelbrot introduced to fractal geometry, in Hurst's honor, the term Generalized Hurst Exponent. The Hurst exponent is used as a measure of the long-term memory of a time series [5,6].



1.2 Estimating the Hurst exponent:

A variety of methods occur for estimating the Hurst exponent (H) and the process detailed here is both simple and highly data intensive. To estimate the Hurst exponent one must regress the rescaled range on the time span of observations. To do this, a time series of full length is divided into a number of shorter time series and the rescaled range is calculated for each of the smaller time series. A minimum length of eight is usually chosen for the length of the smallest time series. So, for example, if a time series has 128 observations it is divided into:

- two chunks of 64 observations each
- four chunks of 32 observations each
- eight chunks of 16 observations each
- 16 chunks of eight observations each

Steps for estimating the Hurst exponent after breaking the time series into chunks/masses:
For each chunk of observations, compute:

- the mean of the time series,
- a mean centered series by subtracting the mean from the series,
- the cumulative deviation of the series from the mean by summing up the mean centered values,
- the Range (R), which is the difference between the maximum value of the cumulative deviation and the minimum value of the cumulative deviation,
- the standard deviation (S) of the mean centered values, and
- the rescaled range by dividing the Range by the standard deviation.

Finally, average the rescaled range over all the chunks.

The rescaled range and chunk size follows a power law, and the Hurst exponent is given by the exponent of this power law. For example, the 80/20 rule (20 percent of the population holds 80 percent of wealth), the winner-take-all phenomenon, friend connections in a social network and forest fires all follow power laws.

Using the Hurst exponent we can classify time series into types and gain some insight into their dynamics. Here are some types of time series and the Hurst exponents associated with each of them [3].From the Hurst exponent H of a time series, the fractal dimension of the time series can be found. When Df=1.5, there is normal scaling , when Df is between 1.5 and 2, time series is anti-persistent and when Df is between 1 and 1.5 the time series is persistent. For Df =1.5, the time series is purely random. Long term correlations of indexes in developed and emerging markets have been studied by using Hurst analysis.

1.3Fractal dimension and Hurst exponent:

Fractal analysis is prepared by conducting rescaled range (R/S) analysis of time series. The Hurst exponent can classify a given time series in terms of whether it is a random, a persistent, or an anti-persistent process. Simulation study is run to study the distribution properties of the Hurst exponent using first-order autoregressive process. If time series data are randomly generated from a normal distribution then the estimated Hurst Exponents are also normally distributed [7]. The importance of the fractal dimension of a time series lies in the fact that it recognizes that a process can be somewhere between deterministic and random [8].

The Hurst Exponent is directly related to the fractal dimension, which measures the smoothness of a surface, or, in our case, the smoothness of a time series. The relationship between the fractal dimension *D*, and the Hurst Exponent *H*, is given by:

$$D=2-H \qquad [1]$$

where, $0 \le H \le 1$. The closer the value of the Hurst Exponent to 0, the sharper will be the time series.
The Hurst Exponent is the measure of the smoothness of fractal time series based on the asymptotic behaviour of the rescaled range of the process.
The Hurst Exponent, *H*, can be estimated by:

$$H=\frac{\log(\frac{R}{S})}{\log(T)} \qquad [2]$$

where, *T* is the duration of the sample data and *R* / *S* the corresponding value of the rescaled range.

1.4 Rescaled Range Analysis:

Hurst (1965) established the rescaled range analysis, a statistical technique to analyze long records of natural phenomenon. Rescaled Range Analysis is the central tool of fractal data modeling. The two factors used in this range analysis are:

1) The difference between the maximum and the minimum cumulative values, and
2) The standard deviation from the observed values.

The *R* / *S* value scales as we increase the time increment, *T*, by a power law value which equals *H*, the Hurst Exponent. All fractals scale according to a power law. By rescaling data, we can compare diverse phenomena and time periods. Rescaled Range Analysis enables us to describe a time series that has no characteristic scale. Brownian motion is the primary model for a random walk process. Einstein (1908) found the distance a particle covers increases with respect to time according to the following relation:

$$R= T^{0.5} \qquad [3]$$

Where, *R* is the distance covered by the particle in time T [9].

The theory of the rescaled range analysis was first given by Hurst. Mandelbrot and Wallis further refined the method. Feder (1988) gives an excellent review of the analysis of data using time series [10], history, theory and applications, and adds some more statistical experiments to establish the effectiveness of this approach. The parameter H, the Hurst exponent provides insight into the trends and patterns shown by the time series in respect of as to whether the time series is random or not. It is also related to the fractal dimension, the Rescaled range analysis (R/S) approach of estimating H is helpful in distinguishing a completely random time series from a correlated time series and. The value of H reveals persistence of trends in a given time series.

1.5 Weather data of Aurangabad:

We analysed weather data of Aurangabad (Maharashtra State) using the Rescaled Range Analysis (R/S Analysis). The weather data used daily maximum and minimum temperature, wind speed, Dew point and heat index. The data used is taken from the standard information released by Automatic weather station. An automatic weather station was recently installed at MGM's Jawaharlal Nehru Engineering College (JNEC), which is keeping a tab on 15 different weather related parameters in the region. The station, which operates on solar energy besides conventional power supply, can keep a tab on weather conditions of the city and nearby areas of around 10 kilometres. The weather station can share vital information such as direction of air, rainfall received, weather forecast for 12-hours and ultraviolet rays. The automatic weather station can share information about weather systems around six hours in advance. It is linked with online resources so as to generate desired data in a user-friendly format [11].

R/S analysis scans events and denotes the existence of memory through *H* exponent value. *H* provides an exact measure of randomness in any particular time series considered. *H* = 0 features an anti-persistent time series, in other words, an event that occurs subsequently after will have a 100% tendency of an opposite behaviour ; *H* = 1 indicates that the series is persistent, meaning that an event that occurs subsequently after will have a 100% tendency of maintaining the same tendency; when *H* = 1 / 2 the event is random, as the Brownian motion, meaning it is not possible to predict the behaviour of the next event based on what has occurred in the past [12].The time series shown next (Figure 1) was taken for the purpose of estimating the Hurst Exponent.

1.6 R/S analysis of temperature:

The input to the programme is a excel file containing the data (weather parameter) to be analysed in the form of continuous sequence of numbers holding the data. The first number in the file is the total number of data points that follow. If the number of data point is 256 then the first line in the data file will contain a number 256,which is followed by the actual (256)data values.

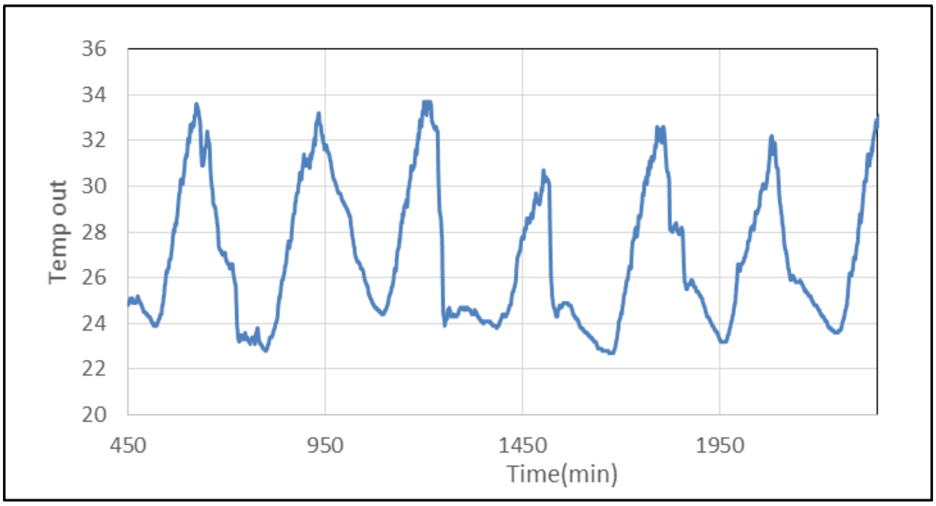


Figure 1: Time series of temperature

The average temperature out for the month of May – June 2016 is shown in figure 1. Implementation of R/S Analysis on temperature data showed that the plot of log (R/S) versus log (T/2) for temp out for the month of May –June 2016. The points are the actual values of (R/S) obtained from the calculations and the line joining the points is the least square fit to these points. The points lie along straight line (R^2 = 0.9997 indicating goodness of fit) the slope of the line (Hurst exponent) is H=0.95 and hence the fractal dimension DF=1.04. The data points from 19 May to 21 June = 304 points.

Table 1: Data points of Temp

S No	τ	τ **/2**	MEAN(R/S)	LOG (τ **/2**)	LOG(R/S)
1	4	2	0.9212665	0.30103	-0.03561
2	8	4	1.813437	0.60206	0.258503
3	16	8	3.435472	0.90309	0.535986
4	32	16	6.524828	1.20412	0.814569
5	64	32	12.61582	1.50515	1.100915
6	128	64	25.16377	1.80618	1.400776
7	256	128	51.42315	2.10721	1.711159

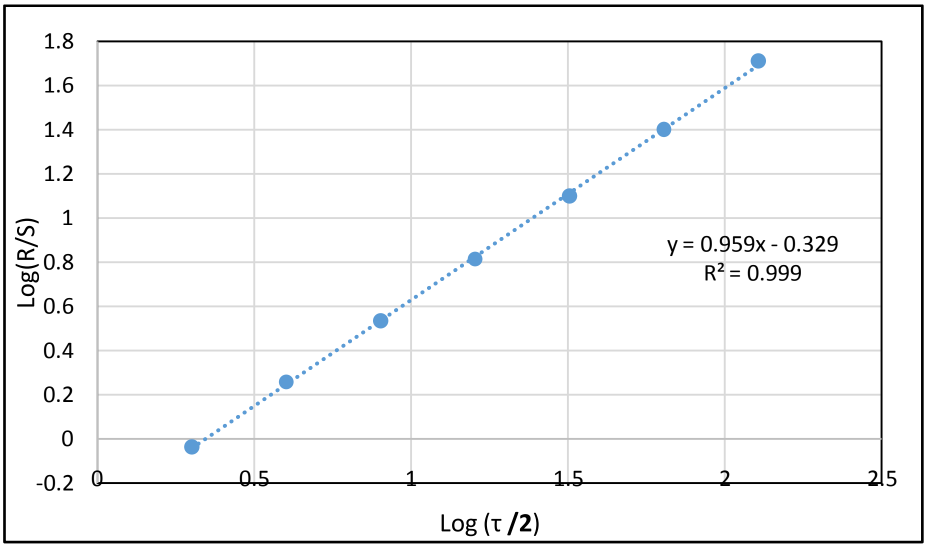


Figure 2: shows plot of log (R/S) vs log (τ /2) for temp

1.7 R/S analysis of Heat Index:

Figure 3 shows time series of Heat index. Figure 4 shows plot of log (R/S) versus log (τ **/2**) for average heat index of the month of May –June 2016.Points plotted are the actual data values of (R/S) obtained from the calculations and the line joining the points is the least square fit to these points. The points lie along a straight line (R^2=0.996 indicating goodness of fit) the slope line (Hurst exponent) is H=1.01 and hence the fractal dimension Df =0.98.

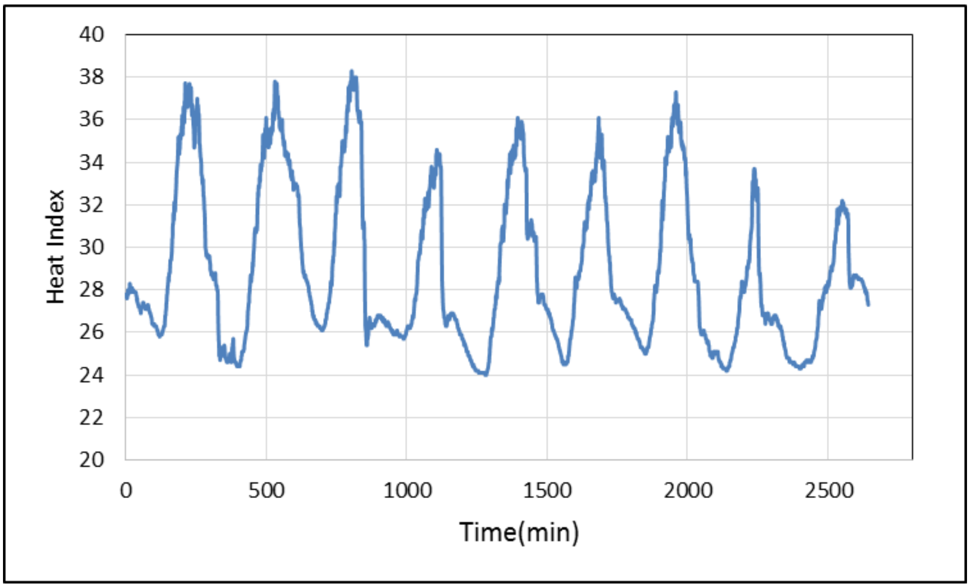


Figure 3: Time series of Heat index

Table 2 : Data points of Heat Index

S No	τ /2	τ /2	MEAN(R/S)	LOG (τ /2)	LOG(R/S)
1	4	2	0.7707458	0.30103	-0.11309
2	8	4	1.603796	0.60206	0.205149
3	16	8	3.393301	0.90309	0.530622
4	32	16	6.590572	1.20412	0.818923
5	64	32	13.59556	1.50515	1.133397
6	128	64	25.72475	1.80618	1.410351
7	256	128	52.59346	2.10721	1.720932

The table 2 shows the data points of heat index and actual values obtained and calculations.

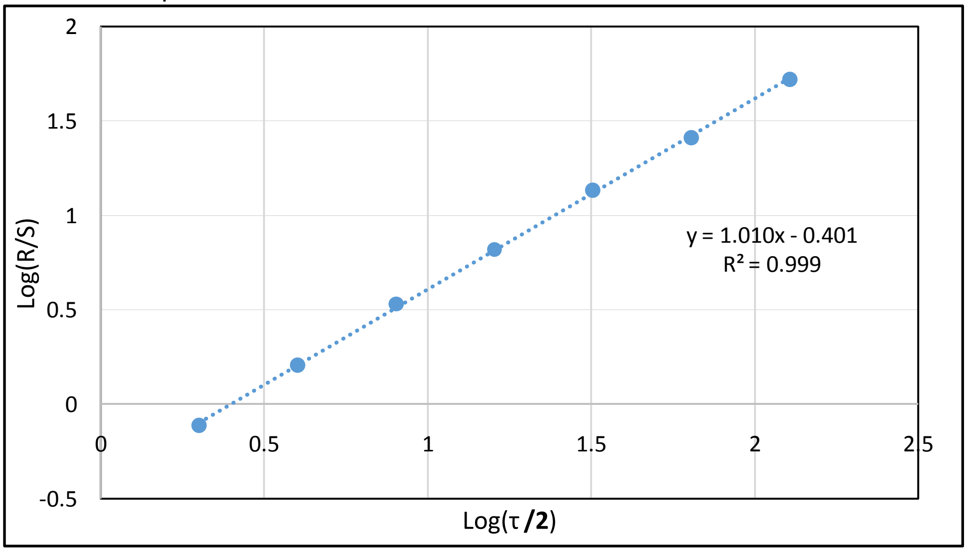


Figure 4: shows plot of log (R/S) vs log (τ /2) of heat index

1.8 R/S analysis of Dew points:

Figure 5 shows time series of Dew points for the period of May –June 2016 and table no 10 shows the actual calculations and values obtained.

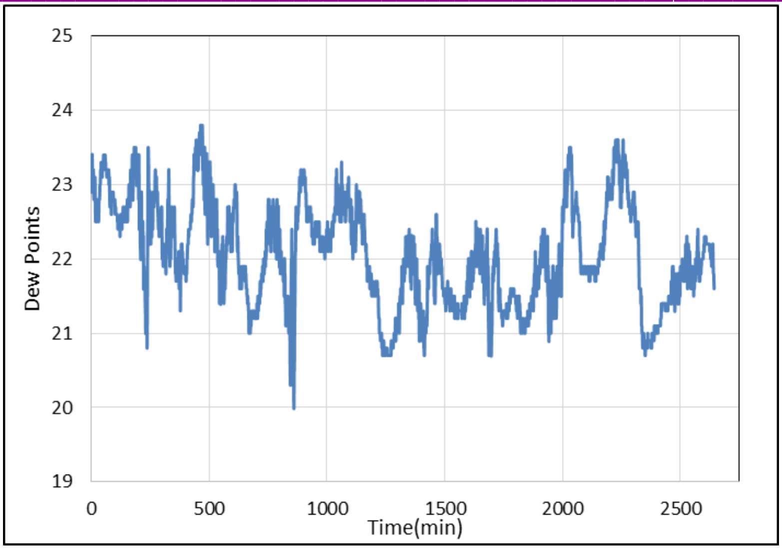


Figure 5: Time series of Dew points

Table No.3 : Data points of Dew points

S No	τ	τ /2	MEAN(R/S)	LOG (τ /2)	LOG(R/S)
1	4	2	0.7932006	0.30103	-0.10062
2	8	4	1.611153	0.60206	0.207137
3	16	8	3.170134	0.90309	0.501078
4	32	16	5.862421	1.20412	0.768077
5	64	32	12.31657	1.50515	1.09049
6	128	64	25.1715	1.80618	1.400909
7	256	128	47.86628	2.10721	1.68003

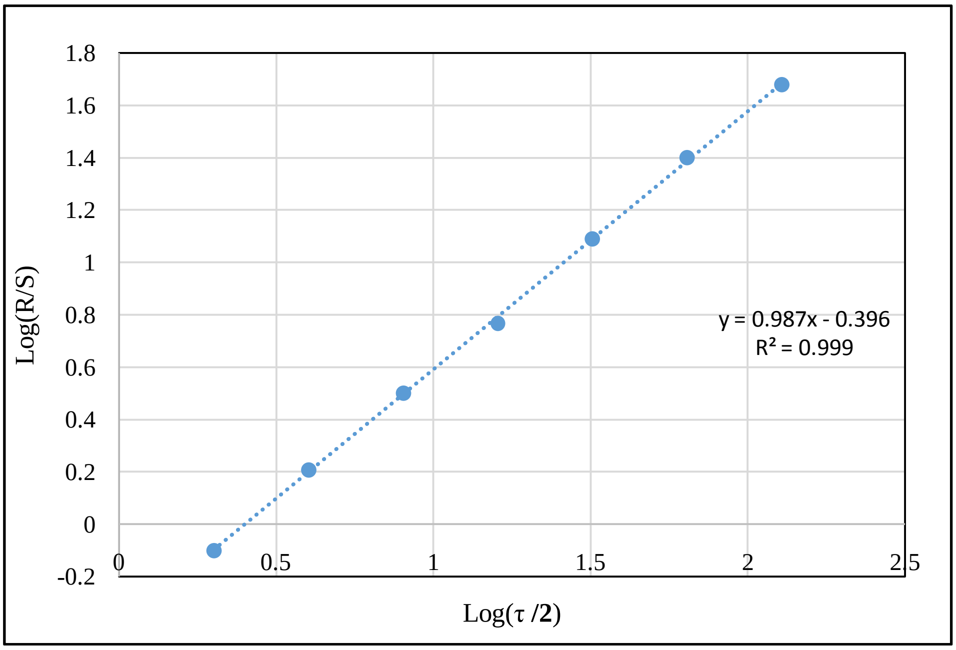


Figure 6: shows plot of log(R/S) vs log (τ /2)

Figure 6 shows plot of log(R/S) versus log (τ **/2**) for dew points of the month of May-June 2016. Points are the actual values of (R/S) obtained from the calculations and the line joining the point is the least square fit to these points. The points lie along a straight line (R^2=0.9996 indicating goodness of fit) the slope of the line (Hurst exponent) is H=0.98 and hence the fractal dimension Df =1.01.

1.9 R/S analysis of Wind speed:

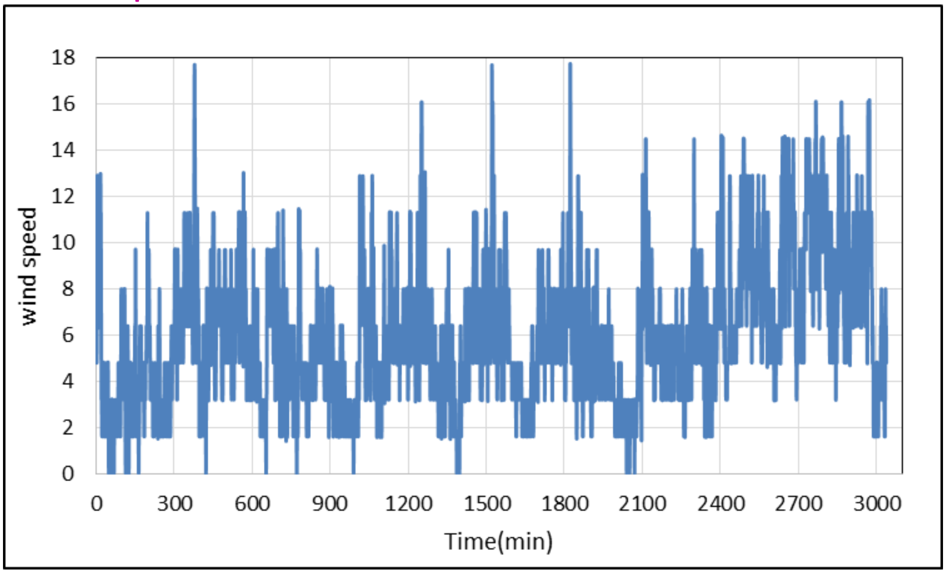


Figure 7: Time series shows Wind Speed

Above figure 5.19 shows time series of wind speed for the month May –June 2016.

Table 4: Data points of wind speed

S No	τ	τ **/2**	MEAN(R/S)	LOG (τ **/2**)	LOG(R/S)
1	4	2	0.7693841	0.30103	-0.11386
2	8	4	1.640837	0.60206	0.215066
3	16	8	3.061585	0.90309	0.485946
4	32	16	5.599591	1.20412	0.748156
5	64	32	10.9598	1.50515	1.039803
6	128	64	19.71742	1.80618	1.29485
7	256	128	35.83566	2.10721	1.554315

Table 4 shows actual calculated data points of wind speed and the actual calculated values of (R/S).

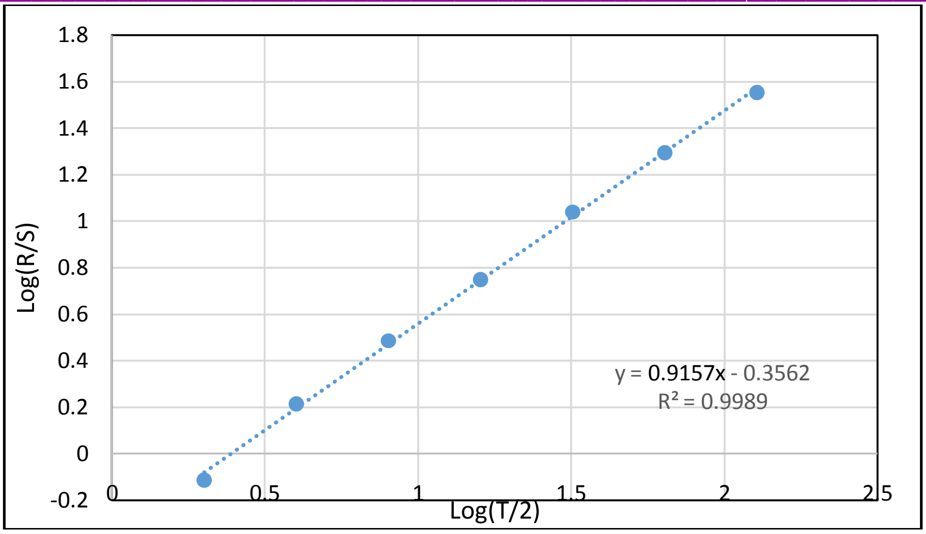


Figure 8: shows plot of log (R/S) vs log (T/2)

Figure 5.20 shows plot of log (R/S) versus log (T/2) of wind speed for the month of May –June 2016. Points plotted are the actual values of (R/S) obtained from the calculations and line joining the points is least square fit to these points . The points lie along a straight line (R^2= 0.996 indicating goodness of fit) the slope of the line (Hurst exponent) is H= 0.91 and hence the fractal dimension Df = 1.08.

Table 5: Gives summery of the estimated values of Hurst exponent and the Fractal Dimensions

Sr. No.	Set	Hurst Exponent H	Fractal Dimension Df	R^2
1	I-Temp out	0.9598	1.0402	0.9997
2	II-Dew point	1.0103	0.9897	0.9996
3	III-Heat index	0.987	1.013	0.9996
4	IV-Wind speed	0.9157	1.0843	0.9996

Table 5 shows the summery of the estimated values of Hurst exponent and the Fractal Dimensions. The values of H of all sets is between 0.5 and 1, this shows the trend is persistent which indicates long memory effect. This means that the increasing trend in the past implies increasing trend in the future also or decreasing trend in past implies decreasing trend in future also.

Thus, *R/S* analysis scans events and denotes the existence of memory through *H* exponent value. *H* provides exact measure of randomness in any specific time series considered. *H* = 0 features an anti-persistent time series, in other words, an event that arises subsequently after will have a 100% tendency of an opposite behaviour; *H* = 1 point toward that the series is persistent, meaning that an event that occurs subsequently after will have a 100% tendency of maintaining the same tendency; when *H* = 1 /2 the event is random, as the Brownian motion, meaning it is not possible to predict the behaviour of the next event based on what has occurred in the past.[13-20]

1.10 CONCLUSION:

The summery of weather data, of four different parameters that is Time, Heat index, Dew points and wind speed the values of Hurst exponent and the Fractal Dimensions has been calculated. The values of H of all sets is between 0.5 and 1, this shows the trend is persistent which indicates long memory effect.

The Hurst exponent is a useful statistical method for inferring the properties of a time series without making assumptions about stationarity. It is most useful when used in conjunction with other techniques, and has been applied in a wide range of industries. For example the Hurst exponent is paired with technical indicators to make decisions about trading securities in financial markets; and it is used extensively in the healthcare industry, where it is paired with machine-learning techniques to monitor EEG signals. The Hurst exponent can even be applied in ecology, where it is used to model populations.

REFERENCES:

1. Zakde K. R, Munde S.V, Yusuf H.Shaikh, Journal of Chemical, Biological and Physical Sciences,An International Peer Review E-3 Journal of Sciences,May 2014 – July 2014, Vol. 4, No. 3; 3587-3592. E-ISSN: 2249 –1929
2. Zakde Kranti, A.R. Khan, Gulam Rabbani, Yusuf H. Shaikh, Characterisation of Irregular Shapes of Biological Images using concept of Fractals, Bionano frontier Vol.10(2)April 2017, Print ISSN 0974-0678
3. Zakade K.R, Gulam Rabbani, Khan A.R, Yusuf H Shaikh, Temporal Evolution of Viscous Fingering in Hele Shaw Cell: A Fractal Approach, International Journal of Scientific & Engineering Research, Volume 6, Issue 9, September-2015, ISSN:2229-5518.
4. Ludmila Kirichenko, Tamara Radivilova, ZhannaDeineko, International Journal "Information Technologies & Knowledge" Vol.5 / 2011,pp 371
5. Mu Sigma , White paper , Predictability of Time Series, July 2012
6. Zakde Kranti, Talal Ahmed Saleh Khamis, Yusuf H. Shaikh, Fractal Analysis of Time series and Estimation of Hurst Exponent in BSE, International Journal of Universal Print,Vol.No.4,Issue No.8,ISSN: 2454-7263,pg.464-470, March 2018
7. MalharKale,FerryButar,Journal of Mathematical Sciences & Mathematics Education, Vol. 5 No. 1
8. Peters, E. E. (1996). Chaos and Order in the Capital Markets (2nd ed.), John Wiley & Sons, Inc. Peters, E. E. (1994). Fractal Market Analysis: Applying Chaos Theory to Investment and Economics, John Wiley &Sons, Inc.
9. Peters E. E. (1994). Fractal Market Analysis: Applying Chaos Theory to Investment and Economics, John Wiley &Sons, Inc.
10. Kantelhardt, B. E Koscielny., H. H Rego A, S Havlin, Bunde A, Physica A 295, 441(2001).19. T.C. Mills, 'Statistical analysis of daily gold price data' PhysicaA 338 (3–4), 559–566(2004).
11. MGM's JNEC weather station, ,Aurangabad
12. Maurea N. Flynn, William R. L. S. Pereira (2013), Biomatematica 23, 1–14, ISSN 1679-365X
13. Matteo T. Di, Aste T.,Michel M. Dacorogna (2005), Elsevier:Journal of Banking & Finance,29,827.
14. H.E Hurst., 'Long-term storage capacity of reservoirs' *Trans. Am. Soc. Civil Eng*'116, 770 (1951).
15. X. Zhou, N. Persaud, H. Wang. 'Scale invariance of daily runoff time series in agricultural watersheds '*Hydrology and Earth System Sciences* 10, 79(2006).
16. A.R Rao., D. Bhattacharya, 'Comparison of Hurst Exponent Estimates in Hydrometeorological Time Series' *J. of Hydrologic Eng*, 4, 3, 225,(1999).
17. Granger, C. W. J, 'Some properies of Time Series Data and Their Use in Econometric Specification'. *Journal of Econometrics*; 16: 121-130(1981).
18. Dr. S. Kaliyamoorthy, 'Relationship of Gold Market and Stock Market: An Analysis' *International Journal of Business and Management Tomorrow* Vol. 2 No. 6 (2012).
19. Capie, F., Terence, C.M., Wood, G., 'Gold as a hedge against the dollar' 'International Financial Markets,Institution and Money 15, 343-352(2005).
20. Zakde K. R., Ph.D thesis submitted to Dr. Babasaheb Ambedkar Marathwada University

Zakde K. R.
Jawahrlal Nehru Engineering College Aurangabad (MS) India.