Chapter

Fractal Scaling Properties in Rainfall Time Series: A Case of Thiruvallur District, Tamil Nadu, India

Ibrahim Lawal Kane and Venkatesan Madha Suresh

Abstract

In the present study, the features of rainfall time series (1971–2016) in 9 meteorological regions of Thiruvallur, Tamil Nadu, India that comprises Thiruvallur, Korattur_Dam, Ponneri, Poondi, Red Hills, Sholingur, Thamaraipakkam, Thiruvottiyur and Vallur Anicut were studied. The evaluation of rainfall time series is one of the approaches for efficient hydrological structure design. Characterising and identifying patterns is one of the main objectives of time series analysis. Rainfall is a complex phenomenon, and the temporal variation of this natural phenomenon has been difficult to characterise and quantify due to its randomness. Such dynamical behaviours are present in multiple domains and it is therefore essential to have tools to model them. To solve this problem, fractal analysis based on Detrended Fluctuation Analysis (DFA) and Rescaled Range (R/S) analysis were employed. The fractal analysis produces estimates of the magnitude of detrended fluctuations at different scales (window sizes) of a time series and assesses the scaling relationship between estimates and time scales. The DFA and (R/S) gives an estimate known as Hurst exponent (H) that assumes self-similarity in the time series. The results of H exponent reveals typical behaviours shown by all the rainfall time series, Thiruvallur and Sholingur rainfall region have H exponent values within $0.5 < H < 1$ which is an indication of persistent behaviour or long memory. In this case, a future data point is likely to be followed by a data point preceding it; Ponneri and Poondi have conflicting results based on the two methods, however, their H values are approximately 0.5 showing random walk behaviour in which there is no correlation between any part and a future. Thamaraipakkam, Thiruvottiyur, Vallur Anicut, Korattur Dam and Red Hills have H values less than 0.5 indicating a property called anti-persistent in which an increase will tend to be followed by a decrease or vice versa. Taking into consideration of such features in modelling, rainfall time series could be an exhaustive rainfall model. Finding appropriate models to estimate and predict future rainfalls is the core idea of this study for future research.

Keywords: Hurst Exponent, Detrended Fluctuation Analysis (DFA), Rescaled Range (R/S) Method, Fractal Analysis, Long memory

1. Introduction

It is challenging task to represent many natural phenomena such as rainfall, earthquakes, wind speed, groundwater flow which vary randomly over time with a physical model. Normally natural phenomena exhibit a high degree of randomness, which will not easily show any pattern beside seasonality and trend. Hydrological phenomena are often regarded as principal examples for non-linear systems and apprehended as complex systems. Indeed, most hydrological phenomena are considered as the outcome of simple systems with nonlinear interdependent but sensitive dependence on initial conditions [1]. Studies have shown that the so called random phenomena exhibit some correlations which can be exposed with some analytical tools. Most of these natural phenomena are not subjected to pure chance but exhibit some kind of correlation e.g. [2, 3]. Monotonous trends may lead to an uncorrelated pattern, under the impact of a trend, to look like long-term correlated pattern [4]. Furthermore, it is difficult to distinguish trends from long-term correlations, because stationary long-term correlated time series exhibit persistent behaviour or long memory and a tendency to stay close to the momentary value [5].

Rainfall is one of the challenging components of the hydrological cycle that exhibits a high non-linear and complicated phenomenon and requires standard and well detailed modelling to obtain accurate prediction. The complex nature of rainfall time series has been appreciated for decades, for example, Tiwari & Pandey [6] studied the trend of rainfall long-term record from 1851 to 2006 for seven meteorological regions of India using the methods of Linear trend analysis, innovative trend analysis, sequential Mann–Kendall test and partial cumulative deviation tests. Rakhecha [7] analysed rainwater features using descriptive statistics on seasonal features of rainstorms, areal rainfalls, quantum and rainwater variability that produced droughts and floods in West Rajasthan. He had used rainfall data of 124 years (1871–1994) in a manner that the information became useful in utilising water resources for human activities. Graham & Mishra [8] modelled with 31 years rainfall data (1985–2015) for Allahabad, Uttar Pradesh-India using Box-Jenkins Methodology. Their results indicated that the seasonal Autoregressive Integrated Moving Average model (ARIMA) model provides consistent and satisfactory predictions for rainfall parameters on monthly scale. Uba & Bakari [9] analysed 372 rainfall data observations for the period of 1981–2011 in Maiduguri-Nigeria. Their results indicated that ARIMA (1, 1, 0) provides a good fit for the rainfall data and is appropriate for short term forecast. Olatayo & Taiwo [10] presented a study that utilised emerging Fuzzy Time Series (FTS), ARIMA and the Thiel's regression methods for the analysis and forecasted the dynamical pattern of rainfall occurrences based on historic data.

Rainfalls data modelling is very essential to many hydrological issues, for example, in identification of intense, moderate and low rainfall areas; detecting areas prone to flood, drought, and other hazardous events; and for agricultural purposes [11]. However, most of the literatures deal either with linear or nonlinear modelling approaches e.g. [12, 13]; both approaches achieved successes in their domains. Nevertheless, none of approaches is found to be a common model that is suitable for all circumstances. These problems strengthen our thinking to extract more information from the available rainfall data. In fact, one of the purposes of measuring data is to learn about the mechanisms in the data themselves and to make conclusion about its present and future state.

Thiruvallur is a region of highly variable rainfall in both spatially and temporally. Therefore, the study of rainfall variability is fundamental to examine its

impact on socio-economic activities. Rainfall in Thiruvallur is highly seasonal that is nonlinear with an organised pattern of clustered structure and may exhibits multi-scaling features. Understanding the nature of the temporal variability of rainfall is important to improve the predictability of climatic events such as floods and droughts [14]. Thus, it is essential to develop a systematic method that will capture the observed characteristics of the data. This study objectively detects the Thiruvallur rainfall patterns for better understanding of researchers in modelling rainfall time series.

2. Data used and study area

The data used in this study is from the Thiruvallur monthly rainfall records of nine locations such as; Thiruvallur, Korattur_Dam, Ponneri, Poondi, Red Hills, Sholingur, Thamaraipakkam, Thiruvottiyur and Vallur Anicut for the period of over 30 years (1971–2016) collected from Institute of Water Studies, Public Works Department, Government of India. Thiruvallur is one of the fastest developing districts in Tamil Nadu, India. it lies between 12°15′ and 13°15' North latitude and 79°15′ and 80°20′ East longitude. The district experiences semi-arid sub-tropical monsoonal climate. Thiruvallur forms part of Coromondal coastal region, topographically flat with some few hills undulated. The average maximum temperature is between 29°C to 36.6°C with the minimum within 17.3°C to 24.4°C. The average normal rainfall of the district is 1104 mm. Out of this about 50% is received during north east monsoon period and about 40% is received during south west monsoon period (http://www.tnenvis.nic.in. retrieved 02/04/2021). The geographical map of the study area is given in **Figure 1**.

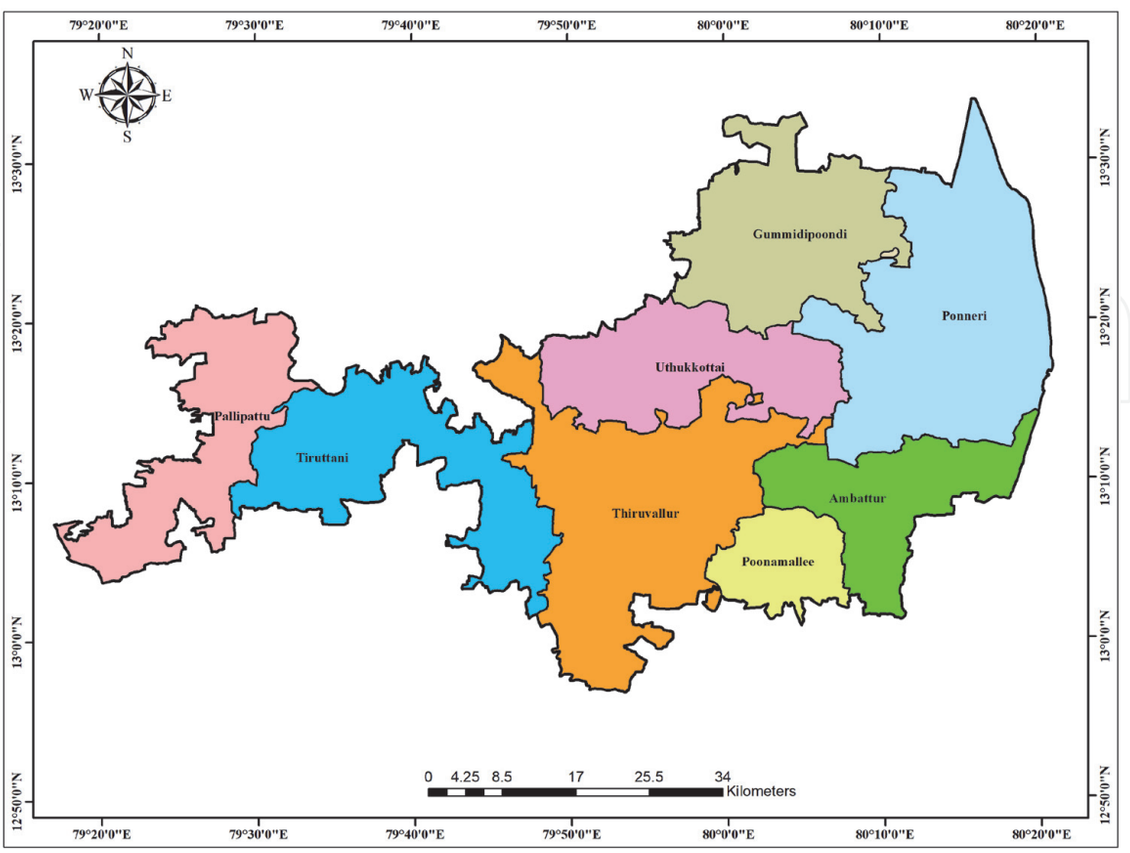


Figure 1.
Geographical map of Thiruvallur district.

3. Methodology

The analysis focus on characterising rainfall based on historical data. The descriptive statistics of the data was first discussed followed by the analysis of fractal scaling properties. The description of the methodology is given in the flowchart in **Figure 2**.

3.1 Fractal scaling analysis

Many geophysical fields appear geometrically complex involving high variability, intermittency and frequent occurrence of extreme values. Fractal scaling analysis, on the other hand presents variety of techniques which can quantify such properties using Hurst phenomenon. The parameter H (Hurst Exponent), display the scaling property of a time series. The Hurst exponent takes values from 0 to 1 ($0 \leq H \leq 1$). If H = 0.5, the series is a random walk (Brownian time series) and there is no correlation between any element and a future element, that is; knowing one data point does not provide insight into knowing future data points in the series. If $0.5 < H < 1$, the series indicates persistent behaviour or long memory. In this case, a future data point is likely to be a data point preceding it. If $0 < H < 0.5$, the series is called anti-persistent. In this case, an increase will tend to be followed by a decrease or vice versa [15]. Among the methods used for quantifying the H embraced in this paper are; Detrended Fluctuation Analysis (DFA) and Rescaled Range (R/S) Methods.

3.1.1 Detrended fluctuation analysis (DFA)

Consider a fluctuating time series x_i for $i = 1, 2, 3, \ldots, N$ sampled at equal time interval $i\Delta t$. Assuming that x_i are increments of a random walk process around the mean $\overline{x} = N^{-1}\sum_{i=1}^{N} x_i = 0$, the 'trajectory' of the signal following integration is given as: $y_j = \sum_{i=1}^{j} x_i$, thus, by dividing the profile into distinct segments indexed by $k = 1, 2, \ldots, \left[\frac{N}{n}\right]$, the confined trend is fitted in each segment, by the polynomial $f_k^{(p)}(j)$ of order p, and the profile is de-trended using the expression;

$$Z_j^p = y_j - f_k^{(p)}(j), \text{for } j = 1, 2, \ldots, N \tag{1}$$

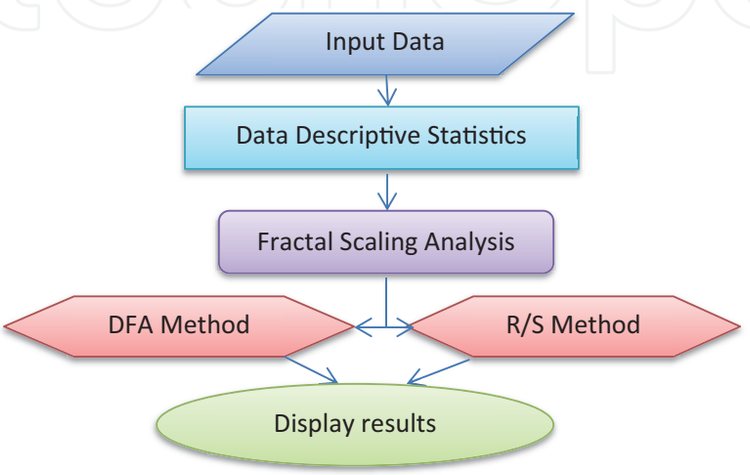


Figure 2.
Flowchart of the methodology.

Following [16], the possible fluctuations can be measured using the root mean square for a given segment of length n.

$$F_p(n) = \sqrt{\frac{1}{n[N/n]}} \sum_{j=1}^{n[N/n]} \left(Z_{\mathrm{j}}^{p}\right)^2 \tag{2}$$

A power-law relationship between $F_p(n)$ and n designates scaling with an exponent δ

$$F_p(n) \sim n^{\delta} \tag{3}$$

and such a process has a power-law autocorrelation function $C(\tau) = \langle x_j x_{\,j+r} \rangle \sim \tau^{-\alpha}$, where $0 < \alpha < 1$, and the relationship between the correlation exponents is $\alpha = 2(1-\delta)$ [17]. Hurst exponent H can be obtained directly from the scaling exponent α, that is, $\alpha = H$.

3.1.2 Rescale range method

The Rescale Range formula is given as [18]:

$$(R/S)_n = cn^H \tag{4}$$

where $(R/S)_n$ represent the rescaled range (R/S) statistic measured over a time index n, the terms c and H represents a constant and the Hurst exponent respectively. The estimation of the Hurst exponent is done by taking the logarithm of (4) to give:

$$log\,(R/S)_n = \log(c) + Hlog\,(n) \tag{5}$$

H can be estimated as the slope of log/log plot of $(R/S)_n$ versus n.

Consider a rainfall time series x_t, for $t = 1, 2, \ldots N$, the (R/S) can be defined as the series representing cumulative deviations from the mean of the rainfall series rescaled by its standard deviation. In the *(R/S)* logic, long memory or long term dependence is considered as the extended periods of whole similar behaviour with unequal duration. The methodical process to estimate $(R/S)_n$ values can be described in the steps below following [15]:

Step1: The time period of a time series of length N is grouped into m adjoining sub groups of length n such that $m \times n = N$ with sub group carrying x_{ij} where $i = 1, 2, \ldots, n$ denotes the number of terms in each sub group and $j = 1, 2, \ldots, m$ denotes the sub group index. For each sub group j the (R/S) statistic can be estimated as:

$$(R/S)_j = S_j^{-1}\left[max_{\,1 \le k \le n} \sum_{j=1}^{k}\left(x_{ij} - \overline{x}\right) - min_{\,1 \le k \le n} \sum_{j=1}^{k}\left(x_{ij} - \overline{x}\right)\right] \tag{6}$$

where S_j denote the standard deviation for each sub group. In Eq. (5), the k^{th} deviations from the sub group mean have mean equal to zero, therefore the last value of the cumulative deviations for each sub group will equally be zero. Hence, the maximum value of the cumulative deviations will be greater than or equal to 0, whereas the minimum value of the cumulative deviations will be less than or equal to 0. Thus the bracketed term, that is, the range value will be non-negative.

Step2: The $(R/S)_n$ is computed by taking the average of $(R/S)_j$ for all the m adjoining sub groups with length n resulting to:

$$(R/S)_n = 1/m \sum_{j-1}^{k} (R/S)_j \quad (7)$$

Step 3: Note that Eq. (6) calculates the *(R/S)* value corresponding to a certain groups of length n. While applying Eq. (5), steps 1 and 2 are repeated by increasing n values until $n = N/2$. After these steps, it is apparent that the time width is contained in the *(R/S)* analysis by examining whether the range of the cumulative deviations depends on the length of the whole time period. After Eq. (7) is estimated for different n periods, the Hurst exponent can be estimated through an ordinary least square regression from Eq. (5).

4. Results

4.1 Descriptive statistics of the rainfall data sets

The descriptive statistic of the considered 9 monthly rainfall series is given in **Table 1**. It could be observed from the table that S01-S08 series follow the same statistical patterns with standard deviation less than the mean and S09 has standard deviation greater than the mean. The Coefficient of Variation (C.V) measure the distribution of data points around the mean. It symbolises the ratio of the standard deviation to the mean. Data with a C.V value less than 1 is considered to have low-variability, while that with a C.V value higher than 1 is considered to have high variability [19]. From **Table 1**, the C.V's for all the data sets are higher than 1 which indicate that the rainfall fluctuates significantly through time except that of S09 which shows negligible C.V from the mean.

4.2 Fractal scaling analysis

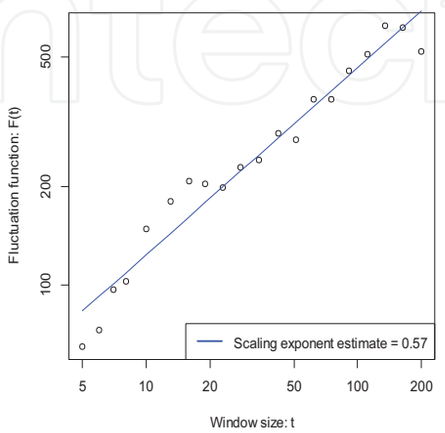
Figure 3a–i depicted the results of the (DFA) with fractal scaling properties. DFA gives estimates of the degree of detrended fluctuations at different periods (window size t) of the rainfall time series. It measures the scaling association between estimates and the window size. The estimation of the Hurst parameter H by (DFA) method shoulders self-similarity in the rainfall series. The signal is said to be self-similar if the detrended fluctuations increases as a power law function of time scale and yield a straight line on a log–log fluctuation plot as the association

Station	Station code	Mean	Std. dev.	Skewness	Kurtosis	C.V
Thiruvallur	S01	108.74	139.99	2.05	8.33	1.29
Korattur_Dam	S02	102.24	138.23	2.49	12.79	1.35
Ponneri	S03	114.63	160.33	2.23	9.52	1.39
Poondi	S04	101.11	133.30	2.89	18.93	1.32
Red Hills	S05	117.61	174.87	2.88	15.21	1.49
Sholingur	S06	73.30	94.92	2.08	9.14	1.28
Thamaraipakkam	S07	99.55	128.83	2.13	9.97	1.29
Thiruvottiyur	S08	77.85	115.49	2.54	11.38	1.48
Vallur Anicut	S09	113.64	110.19	0.47	1.77	0.97

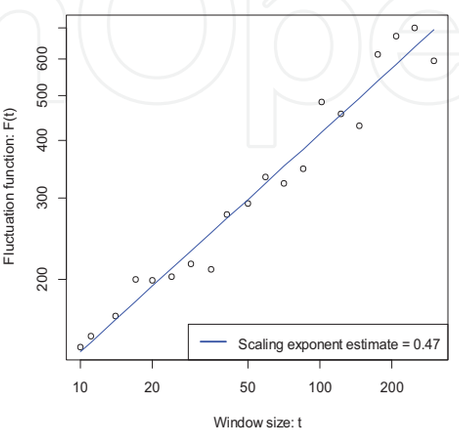
Table 1.
Summary of the descriptive statistics of the rainfall data sets.

between the estimates and window size t. The slope of the plot is the scaling exponent estimate which gives the fractal scaling property also summarised in **Table 2** for the 9 locations.

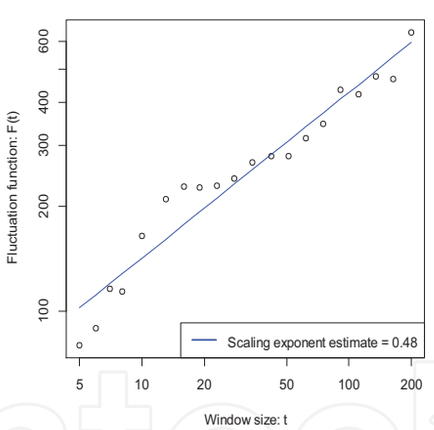
The statistical technique based on (R/S) is designed to assess the nature and extent of variability in data over a time period with the purpose of providing an assessment of how the apparent variability of the rainfall series changes with the length of the time-period 1971 to 2016. (R/S) reveals whether or not a time series exhibits persistence or anti-persistence. A log–log plot of the (R/S) statistic versus the number of points of the aggregated series (**Figure 4a–i**) formed a straight line



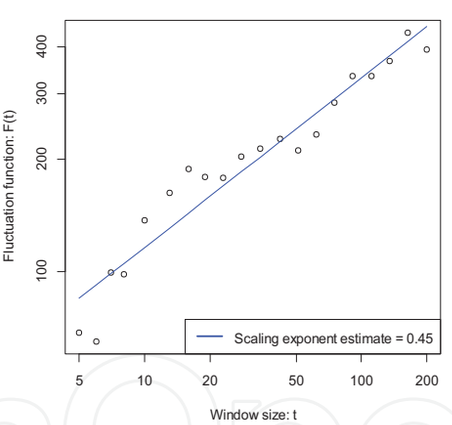
(a)



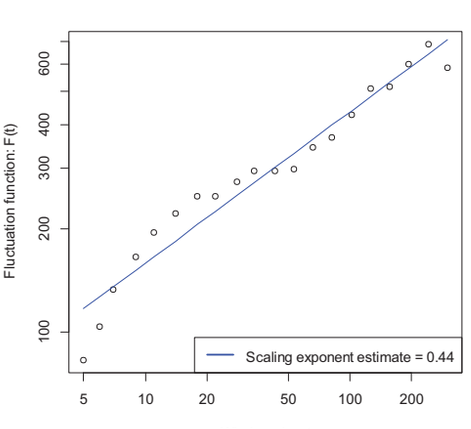
(b)



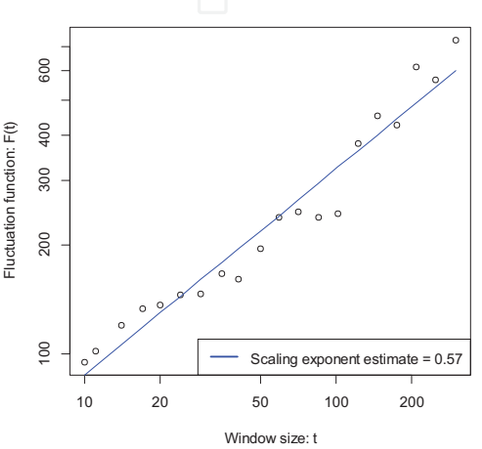
(c)



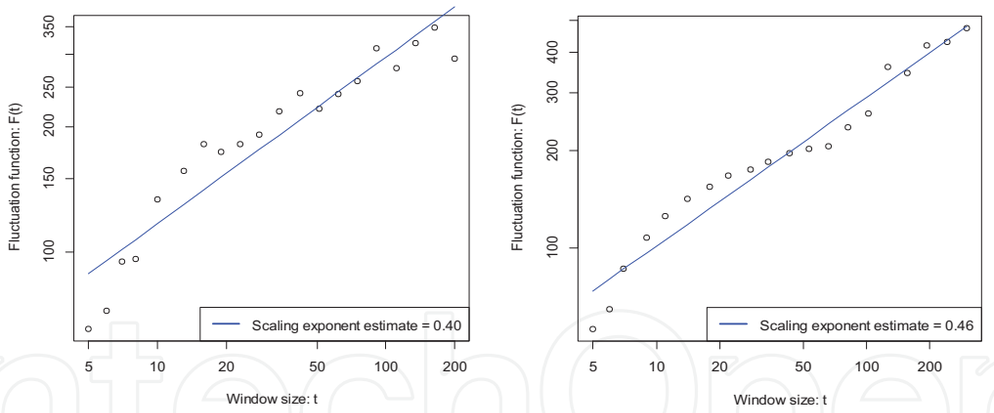
(d)



(e)



(f)



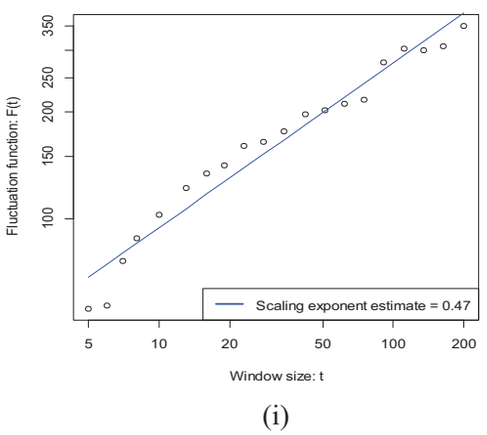


Figure 3.
Hurst exponents results based on DFA method for (a) Thiruvallur, (b) Korattur Dam, (c) Ponneri, (d) Poondi (e) Red Hills, (f) Sholingur, (g) Thamaraipakkam, (h) Thiruvottiyur, and (i) Vallur Anicut.

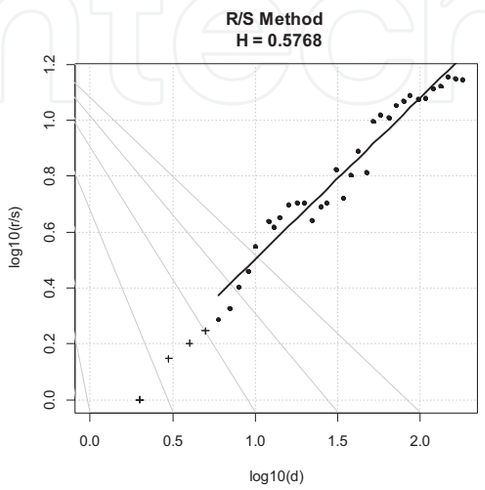
Station	Station code	DFA	(R/S)
Thiruvallur	S01	0.57	0.58
Korattur_Dam	S02	0.47	0.44
Ponneri	S03	0.48	0.52
Poondi	S04	0.45	0.51
Red Hills	S05	0.44	0.42
Sholingur	S06	0.57	0.56
Thamaraipakkam	S07	0.40	0.40
Thiruvottiyur	S08	0.46	0.45
Vallur Anicut	S09	0.47	0.48

Table 2.
Summary results for Hurst exponents based on detrended fluctuation analysis and rescaled range methods.

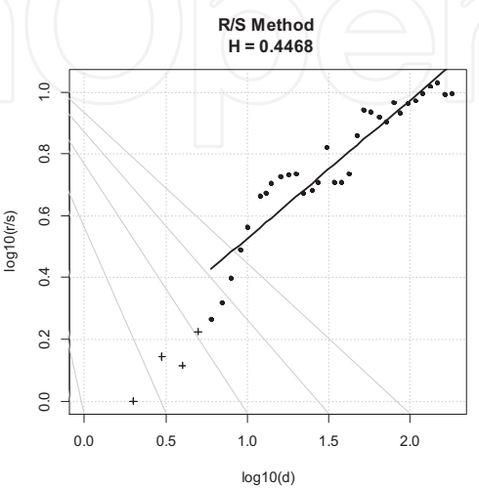
with the slope being an estimate of the Hurst parameter value. The (R/S) results are summarised in **Table 2**.

The evidences in **Figures 3** and **4** for *(R/S)* and DFA methods clearly shows that monthly rainfall time series of Thiruvallur and Sholingur have Hurst exponent

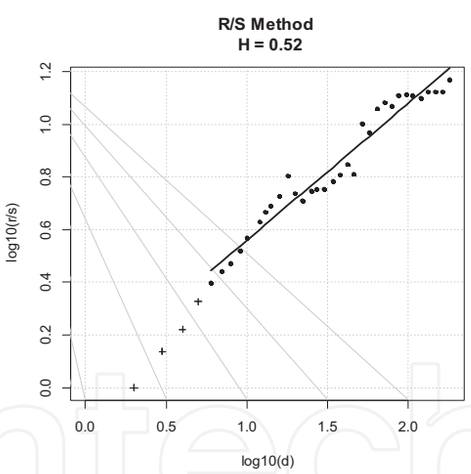
values within $0.5 < H < 1$ which is an indication of strong persistence or long memory, as such the series have a predictable component. Thamaraipakkam, Thiruvottiyur, Vallur Anicut, Korattur Dam and Red Hills have H values less than 0.5, indicating a property called anti-persistent where in an increase will tend to be followed by a decrease or vice versa [14]. Ponneri and Poondi have conflicting results based on the two methods, however, their H values are approximately 0.5 showing a random walk behaviour, and in this, there is no correlation between any part of the data. That is, knowledge of one data point does not provide insight to predict future data points in the series.



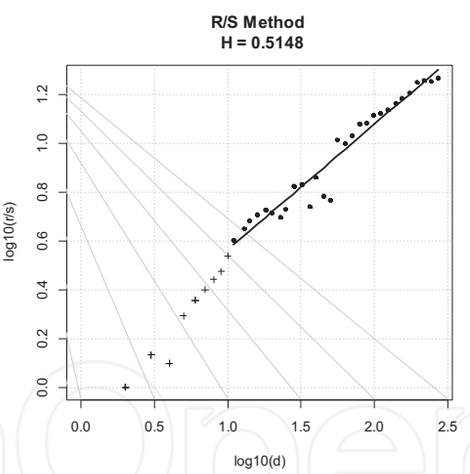
(a)



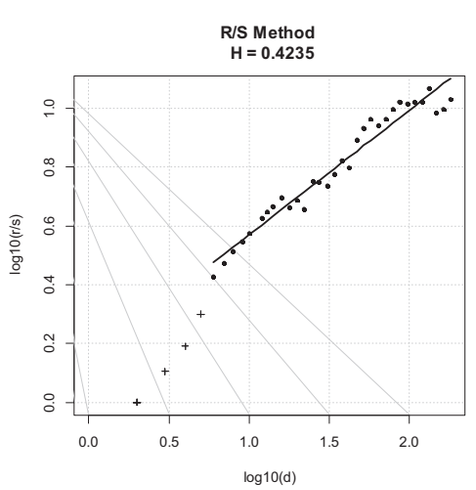
(b)



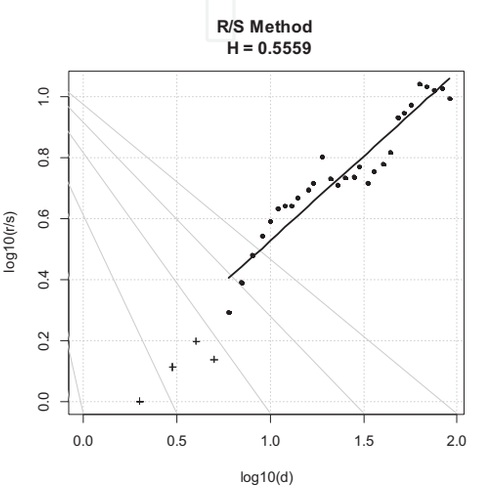
(c)



(d)



(e)



(f)

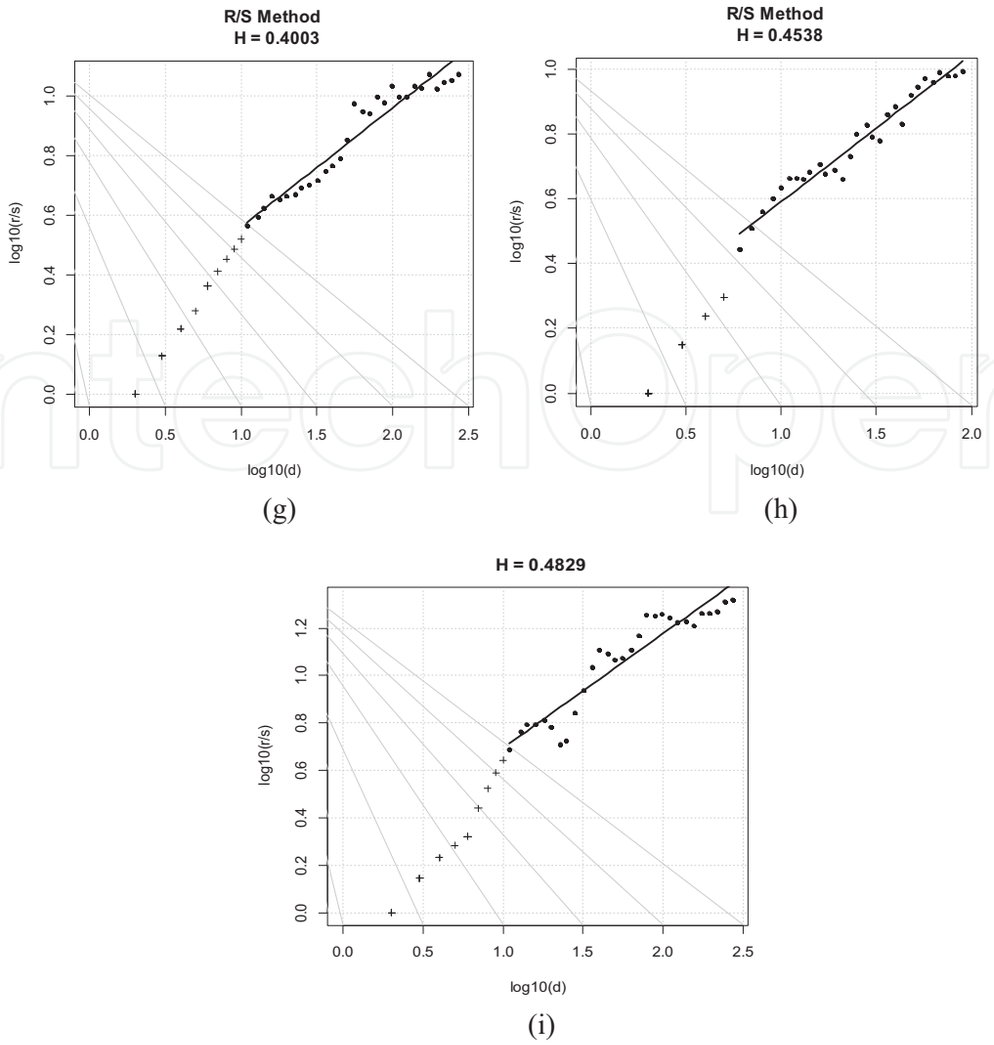


Figure 4.
Plots of (R/S) Hurst estimates for (a) Thiruvallur, (b) Korattur Dam, (c) Ponneri, (d) Poondi (e) Red Hills, (f) Sholingur, (g) Thamaraipakkam, (h) Thiruvottiyur, and (i) Vallur Anicut.

5. Conclusion

Monthly rainfalls in different locations of Thiruvallur district exhibit a tendency of randomness in the long run. The presence of a changing deterministic pattern was examined through a method that allows detecting both apparent and hidden features in rainfall time series. The fractal Scaling analysis based on DFA and the (R/S) methods reveals typical behaviours shown by all the rainfall time series, some are persistent and purely random, some behaves as random walk and some have anti-persistent behaviour. This shows that there is no universal model for predicting rainfall in Thiruvallur district. Rather, rainfall in a location need to be treated based on its associated features. Non consideration of fractal features in hydrological variable modelling may lead to spurious estimates. Finding appropriate model to estimate and predict future rainfalls with consideration to the observed characteristic would be a subject for future research.

Conflict of interest

The authors declare no conflict of interest.

DOI: http://dx.doi.org/10.5772/intechopen.100631

Author details

Ibrahim Lawal Kane[1*] and Venkatesan Madha Suresh[2]

1 Department of Mathematics and Statistics, Umaru Musa Yar'adua University, Katsina, Nigeria

2 Director, G.S.Gill Research Institute and Professor and Head (former) Centre for Natural Hazards and Disaster Studies, University of Madras, Chennai, India

*Address all correspondence to: ibrahim.lawal@umyu.edu.ng

References

[1] Sivakumar, B. (2004). Chaos theory in geophysics: past, present and future. Chaos, Solitons and Fractals. 19, 441-462.

[2] Gil-Alana, Luis A. (2012). U.K. Rainfall Data: A Long-Term Persistence Approach. J. Appl. Meteor. Climatol. 51, 1904–1913.

[3] Lovallo, M., Shaban, A., Darwich, T. and Telesca, L. (2013). Investigating the time dynamics of monthly rainfall time series observed in northern Lebanon by means of the detrended fluctuation analysis and the Fisher-Shannon method. Acta Geophysica, DOI 10.2478/s11600-012-0094-9.

[4] Nnaji C. C. (2011). Time Series Analysis of Monthly Rainfall In Nigeria With Emphasis on Self-Organized Criticality. Journal of Science and Technology.Vol. 31, No. 1, pp 139

[5] Koscielny-Bunde, E. Kantelhardt, J. W. Braun, P. Bunde, A. and S. Havlin, S. (2006). Long-term persistence and multifractality of river runoff records: Detrended fluctuation studies. J. Hydrol. 322, 120– 137.

[6] Tiwari, H., and Pandey, B.K. (2019) Non-parametric characterization of long-term rainfall time series. Meteorol Atmos Phys 131, 627–637. https://doi.org/10.1007/s00703-018-0592-7

[7] Rakhecha PR. (2018). Hydrometeorology of the Rajasthan desert rainfall. Int J Hydro.; 2(6):688–694. DOI: 10.15406/ijh.2018.2.00144

[8] Graham A, & Mishra E. P. (2017). Time series analysis model to forecast rainfall for Allahabad region. Journal of Pharmacognosy and Phytochemistry; 6(5): 1418-1421. E-ISSN: 2278-4136.

[9] Uba E. S & Bakari H. R. An Application of Time Series Analysis in Modeling Monthly Rainfall Data for Maiduguri, North Eastern Nigeria. Mathematical Theory and Modeling, 2015, Vol.5, No.11, Pp 24-33. ISSN 2225-0522 (Online), www.iiste.org.

[10] Olatayo T. O, & Taiwo A. I. (2004). Statistical Modelling and Prediction of Rainfall Time Series. Global Journal of Computer Science and Technology, 2014, G Interdisciplinary Volume 14 Issue 1 Version1.0. (USA) Online ISSN: 0975-4172 & Print ISSN: 0975-4350

[11] Patakamuri S.K, Krishnaveni M. and Venkataramana S. (2020). Long-Term Homogeneity, Trend, and Change-Point Analysis of Rainfall in the Arid District of Ananthapuramu, Andhra Pradesh State, India. Water, 12, 211; doi:10.3390/w12010211.

[12] Kokilavani S, Pangayarselvi R, Ramanathan SP, Dheebakaran Ga, Sathyamoorthy NK, Maragatham N & Gowtham R.(2020). SARIMA Modelling and Forecasting of Monthly Rainfall Patterns for Coimbatore, Tamil Nadu, India. Current Journal of Applied Science and Technology, 2020; 39(8): 69-76, Article no.CJAST.55802.

[13] Csaba I, Endre T. and Péter S. (2018). Examination of rainfall data for 110 years using spectral and wavelet analysis, Central European Geology, Vol. 61, 1, 1–15. DOI: 10.1556/24.61.2018.01

[14] Morata A, Martı'n M L, Luna M Y, & Valero F. Self-similarity patterns of precipitation in the Iberian Peninsula. Theoretical and Applied Climatology. 2006, 85, 41–59. DOI 10.1007/s00704-005-0175-7

[15] Alptekin, N. (2006). Long memory analysis of USD/trial exchange rate. Int. J. Soc. Sci., 1, 111–112.

[16] Peng, C.K., Buldyrev, S. V. and Havlin, S. (1994). Mosaic organization

of DNA nucleotides. Phys. Rev. E. 49 (2), 1685 – 1689.

[17] Koscielny-Bunde, E., A., Bunde, S., Havlin, H.E., Roman, Y., Goldreich, H. and Schellnhuber, J. (1998). Indication of a universal persistence law governing atmospheric variability. Physics Review Letters. 81, 729-732.

[18] Talkner, P. and Weber, R.O. (2000). Power spectrum and Detrended fluctuation analysis: application to daily temperatures. Phys Rev. E62, 150–160.

[19] Chandramouli, A. (2013). The future of U.S manufacturing: a change manifesto. Retrieved on November 23, 2013, from www.cognizant.com, Last access: 23/11/2013.