Long Memory Analysis of Daily Average Temperature Time Series

Ibrahim Lawal Kane
Department of Mathematics and Computer Science,
Umaru Musa Yar'adua University, 2218, Katsina State, Nigeria
Tel: +234 8038445106, +60 169669415 E-mail: ibrahimlk131@gmail.com

Dauda Usman
Department of Mathematics and Computer Science,
Umaru Musa Yar'adua University, 2218, Katsina State, Nigeria
Tel: +234 8036379069, +60 169668650 E-mail: dauusman@yahoo.com

Abstract

A time series has a long memory, in this case there is autocorrelation at long lags. If a time series display long memory, they show significant autocorrelation between observations widely separated in time. R-software has been used to analyze long memory of daily temperature series of Sokoto metropolis. The Modified Rescaled Range (R/S) statistic, the Periodogram and the Aggregated Variance Methods are used to detect long memory property of the series. Application of these tests suggests that the daily average temperature series shows evidence of long memory.

Keywords: Long memory, Hurst exponent, Aggregated Variance, Modified Rescaled Range and Periodogram.

1. Introduction

Research on long memory and fractionally integrated processes has continued at an accelerating rate since the initial publication of the work of [8], [9] and [10], which parameterized the processes of [11] on time series with hyperbolically decaying autocorrelations [2].

Long memory models address the degree of persistence in the data. In their review of long memory models, [7] defined a series as long memory based on a slowly declining autocorrelation structure, "Such autocorrelation structure suggests that the process must depend strongly upon values of time series far away from the past" [4].

The long memory or long term dependence property describes the high-order correlation structure of a series. If a series exhibits long memory, there is a persistent temporal dependence even between distant observations. Such series is characterized by distinct but non-periodic cyclical patterns. The presence of long memory dynamics causes nonlinear dependence in the first moment of the distribution and hence a potentially predictable component in the series dynamics. Fractionally integrated processes can give rise to long memory [1].

Sokoto metropolis is in the dry Sahel, surrounded by sandy Savannah and isolated hills located in the extreme northwest of Nigeria at Latitude 13^0 02 N and Longitude 05^0 15 E. Sokoto as a whole is a very hot area. The warmest months are February to April when the daytime temperature is rising. The raining season is from June to October during which showers are a daily occurrence. From late October to February, during the cold season, the climate dominated by the Hamattan wind blowing Sahara dust over the land. The dust dims the sunlight thereby lowering temperatures significantly and also leading to the inconvenience of dust everywhere in houses.

The objective of this paper is to analyze the daily average temperature record of the Sokoto metropolis for the period 1989 to 2009 to exploit the extent of long range correlations in the time series using a battery of statistical techniques which have proven useful in revealing the long memory structures of a time series. The rest of the paper is organized as follows: Section 2 contains an outline of the basic methods with long memory dynamics. Section 3 gives the estimation and discussions of the results. Section 4 contains the conclusions.

2. MATERIAL AND METHODS

LONG MEMORY PROCESSES

A stationary stochastic process $\{y_t\}$ is called a long memory process if there exist a real number H and a finite constant C such that the autocorrelation function ρ(τ) has the following rate of decay:

$$\rho(k) \sim C_{\tau}^{2H-2}, \quad \text{as } \tau \to \infty$$

The parameter H, Hurst Exponent, display the long memory property of the time series. A long memory time series is said fractionally integrated, where the fractional degree of integration d is related to the parameter H as follows:

$$d = H - 1/2$$

The Hurst exponent takes values from 0 to 1 ($0 \le H \le 1$). If H = 0.5, the series is a random walk. In a random walk there is no correlation between any element and a future element. If $0.5 < H < 1$, the series indicates persistent behavior or long memory. If there is an increase from time step t_{i-1} to t_i there will be probably be an increase from t_i to t_{i+1}. On the other hand, the same is true for decreases. A decrease will tend to follow a decrease. 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 a decrease will be followed by an increase. This behavior is sometimes called mean reversion. Brownian walks can be generated from a defined Hurst exponent. If the Hurst exponent is $0.5 < H < 1$, the random walk will be a long memory process. The time series like this is sometimes referred to as fractional Brownian motion.

2.1 The Modified Rescaled Range (R/S) Method

The first test for long memory was used by the hydrologist [11] for the design of an optimal reservoir for the Nile river, of where flow regimes were persistent. Hurst gave the following formula:

$$(R/S)_n = cn^H \tag{1}$$

$(R/S)_n$ is the rescaled range statistic measured over a time index n, c is a constant and H the Hurst exponent. This show how the R/S statistic is scaling in time. The aim of the R/S statistic is to estimate the Hurst exponent which can characterize a series. Estimation of Hurst exponent can be done by transforming (1) to:

$$Log(R/S)_n = log(c) + H\log(n) \tag{2}$$

H can be estimated as the slope of log/log plot of $(R/S)_n$ vs. n.
For a time series $\{X_t\}$ $t = 1,\ldots, N$, the R/S statistic can be defined as the range of cumulative deviations from the mean of the series, rescaled by the standard deviation. In the R/S sense, long memory or long term dependence may be described as extended periods of similar overall behavior that are of unequal duration [12]. Within these periods, however, dependence need not exist.

2.2 Periodogram Method

In the frequency domain, analysis of time series is merely the analysis of a stationarity by means of its spectral representation. From the modification of Herglotz's theorem first by [3] followed by [5] and [13]. The Periodogram is given by

$$I_N(\lambda) = \tfrac{1}{2\pi N}\left|\sum_{j=0}^{N-1} y_t e^{j\lambda}\right|^2 \tag{3}$$

Where λ is the frequency, N is the length of the time series and y_t is the actual time series. To estimate H, one has to calculate the periodogram. Since $I_N(\lambda)$ is an estimator of the spectral density, a series with long range dependence should have periodogram which is proportional to $|\lambda|^{1-2H}$ close to the origin. Then regression of the logarithm of the periodogram on the logarithm of the frequency λ should give a coefficient of 1-2H. The slope of the fitted line is the estimate of 1-2H.

2.3 Aggregated Variance Method

The long range dependence stationary time series of length N with finite variance is characterized by a sample mean variance of order N^{2H-2} [6]. This suggests the following method of estimation. For an integer m between 2 and N/2, divide the series into blocks of length m and compute the sample average over each kth block

$$\widehat{X}_k^{(m)} = \frac{1}{m} \sum_{t=(k-1)m+1}^{km} X_t, \qquad k = 1, 2, \ldots [N/m]$$

For each m, compute the sample variance of $\widehat{X}_k^{(m)}$ across the block

$$S_m^2 = \frac{1}{([N/m]-1)} \sum_{k=1}^{[N/m]} (\widehat{X}_k^{(m)} - \widehat{X}) \qquad (4)$$

Plot log S_m^2 against log m.

For sufficiently large values of m, the points should be scattered around a straight line with slope 2H-2. In the case of short range dependence (H = 0.5), the slope is equal to -1. The estimate of H is the slope of the least squares line fit to the points of the plot.

3. RESULTS AND DISCUSSIONS

Long Memory Estimation:

3.1 Modified Rescaled Range Method

A log-log plot of the R/S statistic in equation (2) versus the number of points of the aggregated series should be a straight line with the slope being an estimation of the Hurst parameter. Figure 1 gives the estimated value of H.

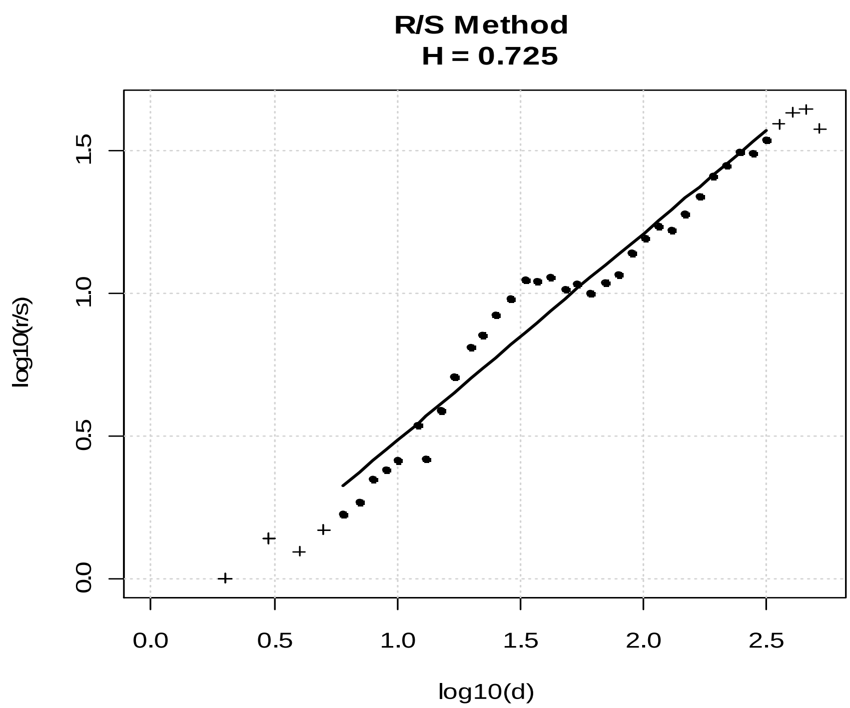

Figure 1: The Pox plot of the Rescaled Range method with Hurst estimate

Fig. 1 shows that, for increasing k, the value of R/S statistics are scattered around a least square fitted line with a slope of approximately equal to 0.725. Therefore estimated H is 0.725. This value of H belongs to ½<H<1 clearly indicates that daily average temperature data has a strong long memory.

3.2 Periodogram Method

This method plots the logarithm of the spectral density of a time series in equation (3) versus the logarithm of the frequencies. The slope of the fitted line is the estimate of 1-2H. The slope provides an estimate of H. Figure 2 gives the estimated value of H.

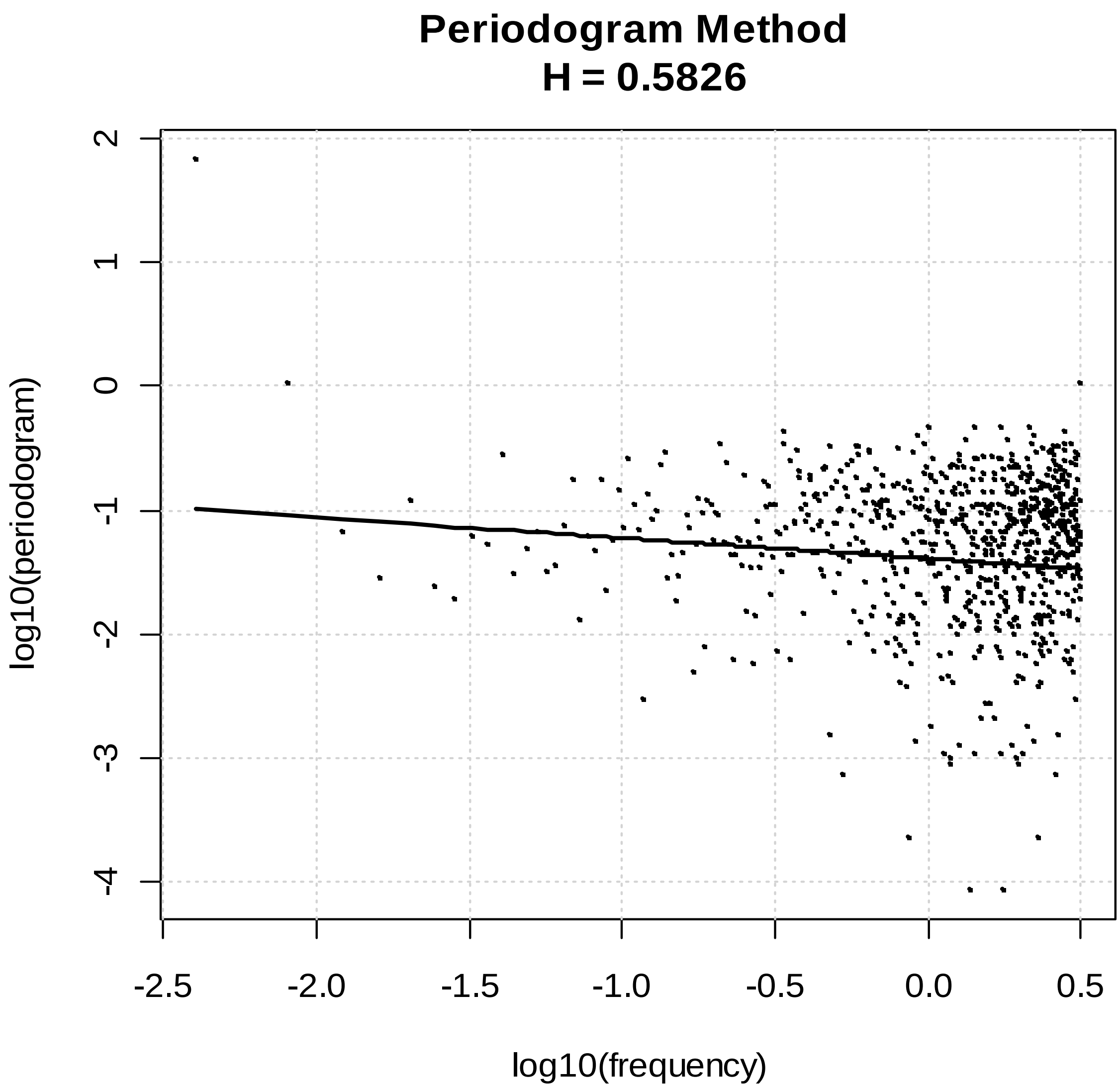

Figure 2: The Periodogram plot with Hurst estimate

The points in Fig. 2 are randomly scattered around a least square fitted line that has a slope of -0.1651603. Therefore estimated H for daily average temperature is 0.5826. This H which belongs to ½<H<1 clearly indicates that the daily average temperature data of Sokoto have a long memory.

3.3 Aggregated Variance Method

The application of this method allows us to estimate the value of the Hurst exponent H which characterized the series. For sufficiently large values of m the plot of log of the equation (4) against log m, the points should be scattered around a straight line with slope 2H-2. The estimate of H is the slope of the least squares line fit to the points of the plot. The estimate of H is given in figure 3.

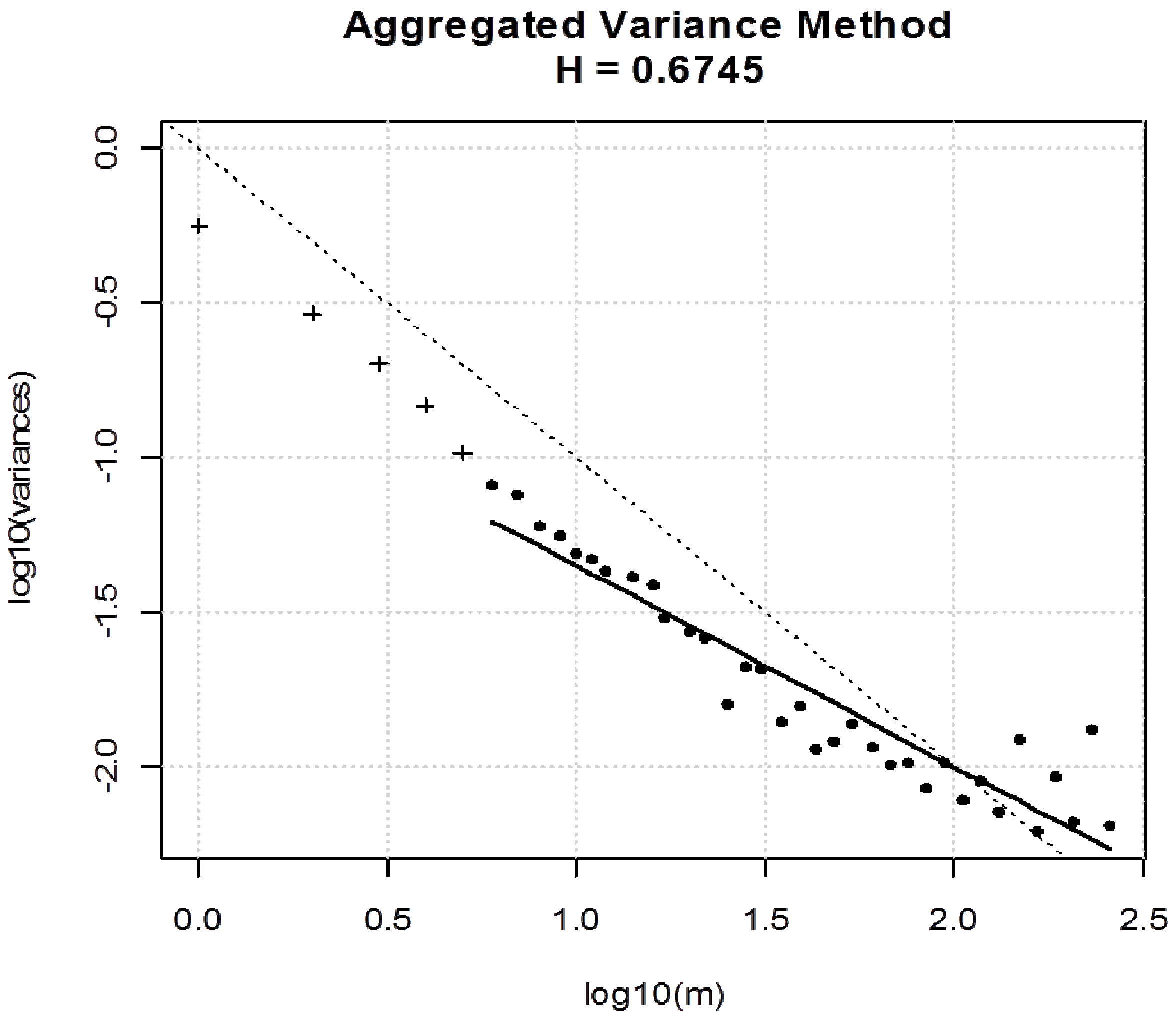

Figure 3: Aggregated Variance plot with Hurst estimate

Fig. 3 gives the Variance time plot of average temperature data set, the points in the plot are scattered around a straight line with a slope of -0.6510185 indicating the presence of long memory among the observation. In this case, the estimated value of H is 0.6749. The estimated H falls within the range 0.5 < H < 1, therefore indicating the existence of long memory.

4. Conclusion

The estimated values of H are approximately in the range of [0.58, 0.72] which clearly indicate the presence of long memory in the temperature series. The variation in H values may be brought about due to the fact that there is variation in the cut off points while fitting a straight line. The consistency of the estimators suggests that all the methods are working rather well.

Reference:

[1] Alptekin, N. (2006), Long memory analysis of usd/trl exchange rate. *International Journal of Social Sciences*, 111-112.

[2] Baillie, R.T. and Kapetanios, G. (2007), Semi parametric estimation of long memory: The Holy grail or poisoned chalice. *International Symposium on Financial Econometrics*, 1-2.

[3] Bloomfield, P. (1992), Trends in Global Temperature. *Climatic Change*, **21**, 1-6.

[4] Box-Steffensmeier, J.M. and Tomlinson, A.R. (1999), Fractional integration methods in political science. *ICPSR*, 1.

[5] Brockwell, P. and Davis, R. (1996), “An introduction to time series and forecasting”. Springer - Verlag, New York, 16.

[6] Chelwoo, P., Felix, H., Juhyun, P., Long, L., Marron, J., Michelle, T., Richard, L., Vladas, P., Zhengyunan Z. (2004), Long range dependence analysis of internet traffic. LRD webpage, 1.

[7] Granger, C. and Ding, Z. (1996), Varieties of long memory models. *Journal of Econometrics*, **73**, 61 – 77.

[8] Granger. C. and Joyeaux (1980), "An Introduction to long memory time series models and fractional differencing". *Journal of Time Series Analysis* **1**, 15-19

[9] Granger C.W., (1980), Long memory relationships and the aggregation of dynamical models. *Journal Econometrics*, **14**, 227–238

[10] Hosking, J. R. (1981), "Fractional differencing". Biometrika**, 68**, 165-178.

[11] Hurst, H. E. (1951), “Long term storage capacity for reservoirs,” Transactions of the American Society of Civil Engingeers**, 116,** 770-799.

[12] Mandelbrot, B. (1972), “Statistical methodology for non-periodic cycles: from the covariance to R/S analysis”. *Annals of Economic and Social Measurement*, **1**, 112.

[13] Sarker, M. M. (2007), Estimation of the self-similarity parameter in long memory processes. *Journal of mechanical Eng.*, **38**, 35.