

Hurst's Memory for Chaotic, Tree Ring, and SOI Series

Byung-Sik Kim¹, Hung-Soo Kim², Sun-Hong Min^{1*}

¹Department of Urban Environmental Prevention Engineering, School of Disaster Prevention, Kangwon National University, Gangwon, South Korea

²Department of Civil Engineering, Inha University, Incheon, South Korea

Email: hydokbs@kangwon.ac.kr, sookim1@inha.ac.kr, *sunmin@kangwon.ac.kr

Received November 13, 2013; revised December 13, 2013; accepted December 20, 2013

Copyright © 2014 Byung-Sik Kim *et al.* This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. In accordance of the Creative Commons Attribution License all Copyrights © 2014 are reserved for SCIRP and the owner of the intellectual property Byung-Sik Kim *et al.* All Copyright © 2014 are guarded by law and by SCIRP as a guardian.

ABSTRACT

Hurst's memory that roots in early work of the British hydrologist H.E. Hurst remains an open problem in stochastic hydrology. Today, the Hurst analysis is widely used for the hydrological studies for the memory and characteristics of time series and many methodologies have been developed for the analysis. So, there are many different techniques for the estimation of the Hurst exponent (H). However, the techniques can produce different characteristics for the persistence of a time series each other. This study uses several techniques such as adjusted range, rescaled range (RR) analysis, modified rescaled range (MRR) analysis, 1/f power spectral density analysis, Maximum Likelihood Estimation (MLE), detrended fluctuations analysis (DFA), and aggregated variance time (AVT) method for the Hurst exponent estimation. The generated time series from chaos and stochastic systems are analyzed for the comparative study of the techniques. Then, this study discusses the advantages and disadvantages of the techniques and also the limitations of them. We found that DFA is the most appropriate technique for the Hurst exponent estimation for both the short term memory and long term memory. We analyze the SOI (Southern Oscillations Index) and 6 tree-ring series for USA sites by means of DFA and the BDS statistic is used for nonlinearity test of the series. From the results, we found that SOI series is nonlinear time series which has a long term memory of $H = 0.92$. Contrary to earlier work, all the tree ring series are not random from our analysis. A certain tree ring series show a long term memory of $H = 0.97$ and nonlinear property. Therefore, we can say that the SOI series has the properties of long memory and nonlinearity and tree ring series could also show long memory and non-linearity.

KEYWORDS

Hurst's Memory; DFA; Chaotic Series; Tree-Ring Series; SOI; BDS Statistic

1. Introduction

Hydrologic or geophysical time series may have a certain dependent structure in itself. For example, if this month has a monthly streamflow with high level at a station, next month could also have a high streamflow. This describes that the consecutive values of hydrologic or geophysical time series show self-dependence between the values which has been known as short-range (or short-term) dependence, persistence, or memory [1]. analyzed the sequence of annual discharges from the Nile River in Egypt to estimate storage volume. The terms "Hurst phenomenon" and "Joseph effect" (due to Mandelbrot from the biblical story of the "seven years of great abundance" and the "seven years of famine") have been used as alternative names for the [2].

[3] suggested a method called the R/S analysis for detecting long memory and the method allowed the calculation of Hurst's exponent or self-similarity parameter. [3] divided the range by the standard deviation of the sample $S(n)$ and called it the "rescaled range" R/S statistic.

Short-term memory implies that the effect of an observation in a time series on the future observations be-

*Corresponding author.

comes negligible after a short period of time. Contrary to short-term memory and long-term memory implies that the effect of an observation on future observations remains significant for a long period of time. Following Hurst's approach, the detection of long memory can be done heuristically by estimating its intensity, namely the value of the parameter H , which varies between 0 and 1. The H value equal to 0.5 means absence of long memory.

Hurst exponent is being used in many fields such as physiology, electronics and computer science and used to determine the persistence of signals such as brain wave, an electrocardiogram and electronic wave. The exponent is also used to investigate the chaotic behavior or self-similarity of the system [4,5]. There are many techniques for the estimation of Hurst exponent such as 1/f power spectral density (PSD) [6], aggregated variance time (AVT) method [7-9], detrended fluctuations analysis (DFA) [10,11], maximum likelihood estimation (MLE) [12].

In this study, we investigate the estimation techniques of Hurst exponent by estimating the exponent for chaotic and stochastic time series and comparing the results for examining the advantages, disadvantages, and limitations of the techniques. We also select an appropriate technique by the examination and apply it to the SOI (southern Oscillation Index) and 6 tree-ring series.

2. Comparison of Hurst Exponent Estimation Methods

Various methods exist for the Hurst exponent estimation and this section discusses the methods by analyzing the data sets generated from the stochastic and chaotic systems. We also discuss the advantages and disadvantages of the methods and the limitations of them through the estimation of Hurst exponents for the data sets from the systems.

2.1. Data Used

This study tests the time series with the known Hurst exponents as shown in 1) to 4) in below and the series generated from chaotic system and a nonlinear stochastic model as in 5) to 7). Each series with the size of 1000 is generated from the systems and the time series plots are shown in [Figure 1](#).

1) Gaussian white noise; iid (identified independant distribution); $H \sim 0.5$

$$x_t = \varepsilon_t \sim N(0,1) \quad (1)$$

where, N : normal distribution with mean 0 and standard deviation 1

t : time

2) Fractional Gaussian noise with long range correlation (FGN); $H \sim 0.8$

$$x_i = B_H(i+1) - B_H(i) \quad i \geq 1 \quad (2)$$

where, B_H : Fractional Brownian motion.

3) Autoregressive model (AR(1)) with $\rho_1 = 0.7$

$$x_t = \rho_1 x_{t-1} + \varepsilon_t \sim N(0,1) \quad (3)$$

where, ρ_1 : lag-1 autocorrelation coefficient.

ε_t : White noise

4) Fractionally Differenced ARMA (FARIMA); $H \sim 0.8$; ARIMA (1,0.3.1)

$$\Phi(B)(1-B)^d x_t = \Theta(B) \varepsilon_t \sim N(0,1) \quad (4)$$

where, $\Phi(B)$: parameter of AR model

$\Theta(B)$: parameter of MA model

B : back shift operator

d : $H-1/2$

5) Three torus—quasi-periodic function

$$x_t = \sin\left[\frac{3t}{500}\right] + \sin\left[\frac{3\sqrt{2}t}{250}\right] + \sin\left[\frac{9\sqrt{3}t}{500}\right], \quad t \geq 1 \quad (5)$$

where, t : time

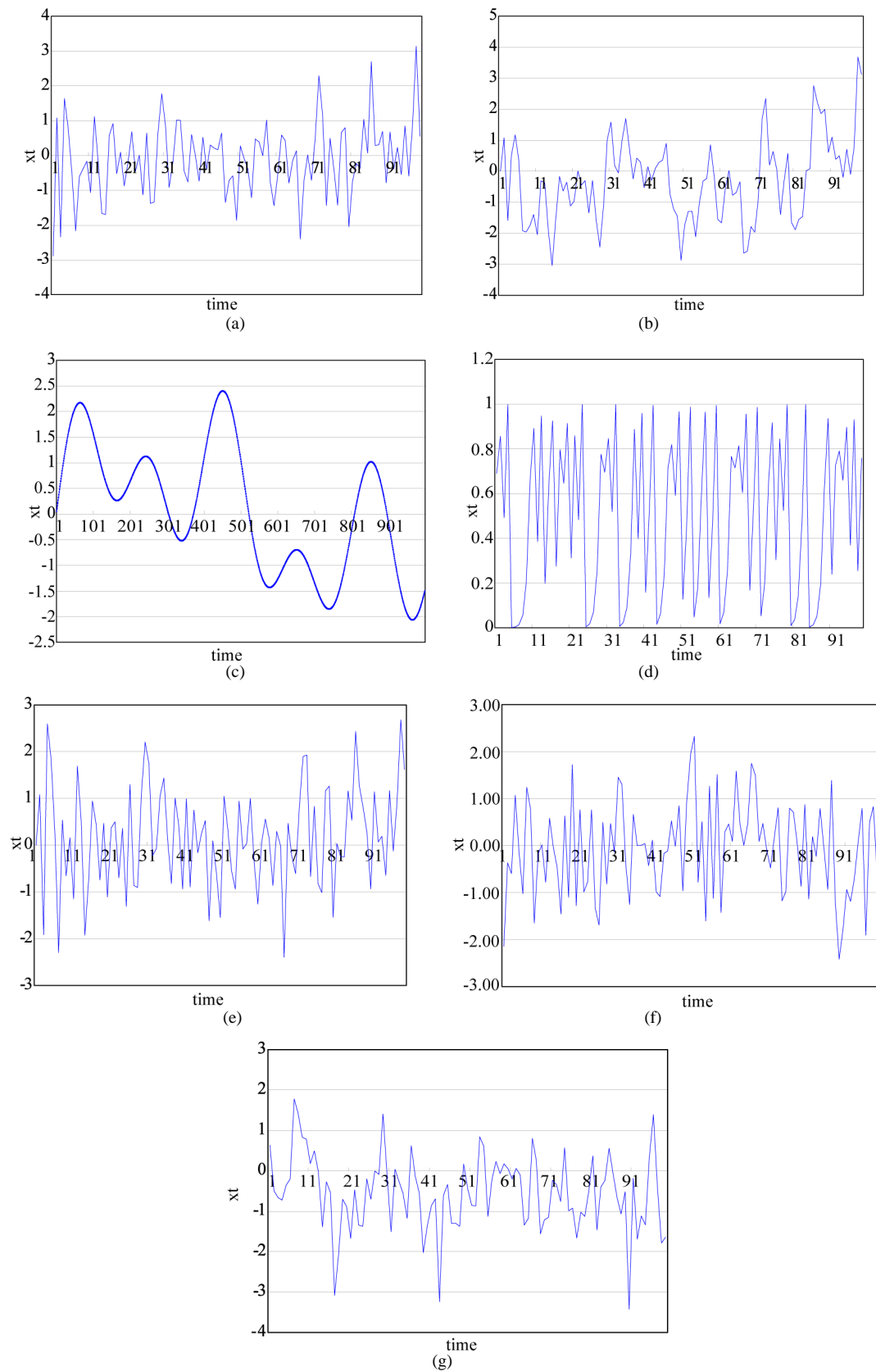


Figure 1. The generated time series from chaotic and stochastic systems. (a) white noise; (b) AR(1); (c) Three torus; (d) Logistic map; (e) TAR(2,1); (f) FARIMA; (g) FGN.

6) Logistic map

$$x_{t+1} = 4x_t(1 - x_t) \quad (6)$$

where, t : time

7) Threshold autoregressive model (TAR(2,1))

$$\begin{aligned} x_t &= -0.5x_{t-1} + \varepsilon_t & \text{if } x_{t-1} < 1 \\ x_t &= +0.4x_{t-1} + \varepsilon_t & \text{if } x_{t-1} \geq 1 \end{aligned} \quad (7)$$

where, t : time.

2.2. Methods for the Estimation of Hurst Exponent

(1) Adjusted Range

[1], examined several sequences of streamflow and observed that for a sequence of n , $1 \leq n \leq N$, the adjusted range statistic, denoted by R , is defined as

$$R = \left[\max \sum_{t=1}^N (X_t - \bar{X}_n) - \min \sum_{t=1}^N (X_t - \bar{X}_n) \right] \quad (8)$$

$$R/\sigma = kN^\beta \rightarrow R/\sigma = \left(\frac{N}{2} \right)^H \quad (9)$$

$$H = \frac{\text{Log}(R_N)}{\text{Log}(N/2)} \quad (10)$$

where, X_t : time series, $t = 1, 2, 3, \dots, N$

R_N : adjusted range statistic

\bar{X}_n : mean

σ : standard deviation.

(2) Rescaled Range

The rescaled range statistic has been used extensively since its formulation by Mandelbrot. For a time series of n observations, $X_1, X_2, X_3, \dots, X_n$, the rescaled range statistic, denoted by $Q(n)$, is defined as

$$Q(n) = \frac{1}{\sigma(n)} \left[\max \sum_{t=1}^k (X_t - \bar{X}_n) - \min \sum_{t=1}^k (X_t - \bar{X}_n) \right] \quad (11)$$

$$\sigma(n) = \sqrt{\frac{\sum_{t=1}^n (X_t - \bar{X}_n)^2}{n-1}} \quad (12)$$

where, X_t : time series, $t = 1, 2, 3, \dots, N$

N : size of the entire series

n : size of the partial series ($1 \leq n \leq N$)

$\sigma(n)$: the standard deviation of the partial series.

[3], used so many annual time series to estimate the Hurst exponent and obtained the following relations with the calculation of $Q(n)$

$$\frac{Q(n)}{\sigma(n)} = \alpha n^H \quad (13)$$

where, H : Hurst exponent

n : size of the partial series ($1 \leq n \leq N$)

α : constant.

Equation (13) can be transformed with the logarithm and Hurst exponent can be estimated from the equation

$$\text{Log}(R(n)/\sigma(n)) = \text{Log}(\alpha) + H(\text{Log}(n)) \quad (14)$$

(3) Modified Rescaled Range

[13], pointed that the regression coefficients can be biased by the autocorrelation when the Hurst exponent is estimated by the regression equation from the known rescaled range. Thus, the rescaled range is a proper method for the time series of the long term memory but it is not for the short term memory. [13], developed the modified rescaled range method like the Equations (15) to (17) and [14,15] applied the method.

$$\mathcal{Q}(n, q) = \frac{1}{\sigma_n(q)} \left[\max_{t=1}^k (X_t - \overline{X_n}) - \min_{t=1}^k (X_t - \overline{X_n}) \right] \quad (15)$$

$$\sigma_n(q) = \sigma_x^2(n) + 2 \sum_{j=1}^q w_j(q) \gamma_j \quad (16)$$

$$w_j(q) = 1 - \frac{J}{q+1}, \quad q < n, \quad j = 1, 2, 3, \dots, n \quad (17)$$

where, $\sigma_n(q)$: weighted sum of autocovariances

$w_j(q)$: weighted autocovariance function

γ_j : autocovariance estimator

$\sigma_x^2(n)$: sample variance

q : Truncation lag of the weighted autocovariance function.

Here, we may carefully determine the truncation lag, q of weighted autocovariance function for the application of the modified rescaled range method. If q is so small, the effect of short range could not be considered, while q is so big the long range could be ignored [16]. Thus, the optimal truncation lag, q_{opt} should be estimated. [17], suggested the $q = n^{0.25}$ and [16] Equation (18). [15], used the modified rescaled range method with $q = N/10$. However this study will use the Equation (18) [7].

$$q_{opt} = \left(\frac{3N}{2} \right)^{1/3} \left(\frac{2\rho}{1-\rho^2} \right)^{2/3} \quad (18)$$

where, ρ : first order autocorrelation coefficient.

(4) 1/f Power Spectral Density Analysis

Power Spectral Density (PSD) has been used in the fields such as physiology, hear rate variability, and the self similarity of brain waves for the estimation of Hurst exponent. Trend in the PSD is proportional to $1/f^\alpha$ and the trend of $1/f$ is represented by the straight line with the slope of α and the $-\alpha$ can be calculated from the regression of low frequency band. The slope of α and Hurst exponent have the following relationship [6].

$$h(f) = \sigma^2 (2\pi)^{-1} \sum_{k=-\infty}^{+\infty} \rho(k) e^{ikf} \quad (19)$$

where, $h(f)$: 1/f power spectral density

$$\alpha > 1 \rightarrow H = \frac{\alpha - 1}{2} \quad (20)$$

$$\alpha < 1 \rightarrow H = \frac{\alpha + 1}{2} \quad (21)$$

(5) Detrended Fluctuations Analysis

The method of DFA has proven as a useful tool in revealing the extent of long-range correlation in time series. Firstly, the time series to be analyzed (with N samples) is integrated and the integrated time series is divided into boxes of equal length, n . In each box of length n , a least squares line is fit to the data (representing the trend in that box). The y-axis of the straight line segments is denoted, by $y_n(k)$. Next, we detrend the integrated time series, $y(k)$, by subtracting the local trend, $y_n(k)$, in each box. The root mean square fluctuation of this integrated

and detrended time series is calculated by Equations (22) and (23)

$$y(k) = \sum_{i=1}^k [X(i) - X_{\text{mean}}] \quad (22)$$

$$F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^N [y(k) - y_n(k)]^2} \quad (23)$$

This computation is repeated over all time scales (box sizes) to characterize the relationship between $F(n)$. The average fluctuation, as a function of boxes size, Typically, $F(n)$ will increase with size n . A linear relationship on a log-log plot indicates the presence of power law scaling. Under such conditions, the fluctuations can be characterized by a scaling exponent d , the slope of the linear relation of $\log F(n)$ vs. $\log n$.

The d has the following relationship and this relationship is represented by Equation (24). Then, The Hurst exponent is estimated by Equations (20) and (21).

$$\alpha = 2d - 1 \quad (24)$$

(6) Aggregated Variance Time Metho

Firstly, the mean value is obtained with the block of N/m by using Equations (25) and (26) for the time series X_i $i = 1, 2, \dots, N$ (see the [Figure 2](#)).

$$X^{(m)}(k) = \frac{1}{m} \sum_{i=(k-1)m+1}^{km} X_i, \quad k = 1, 2, 3, \dots, [N/m] \quad (25)$$

$$\text{Var}X^{(m)} = \frac{1}{N/m} \left(X^{(m)}(k) - \bar{X} \right)^2 \quad (26)$$

For successive values of m , the sample variance of the aggregated series is plotted by m vs. log-log plot. The result should be a straight line with a slope of β and β has the following relationship:

$$H = 1 - \beta/2 \quad (27)$$

In practical, the slope is estimated by fitting a least-squares line to the points of the plot.

(7) Maximum Likelihood Estimation(MLE)

Here the d is estimated by the S-MLE function of S-Plus [18], and the d is related to the Hurst exponent as follows;

$$d = H - 0.5 \quad (28)$$

2.3. Applications of the Hurst Exponent Estimation Methods

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar: This section is to apply previous mentioned techniques for the estimation of Hurst exponent to the time series in a Section 2.1. [Figure 3](#) shows the autocorrelation functions (ACF) of the time series. If we see the ACFs, the white noise and logistic map show the short term memory and others show the long term memory. Theoretical results for normal independent processes [3], indicated that asymptotically $h = 1/2$. One interpretation of the Hurst phenomenon has been to associate $h = 1/2$ with short memory models possessing short-term dependence structure, and $h > 1/2$ with long memory models possessing long-term dependence [19].

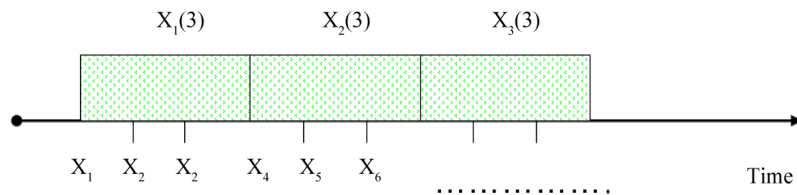


Figure 2. Aggregation of time series.

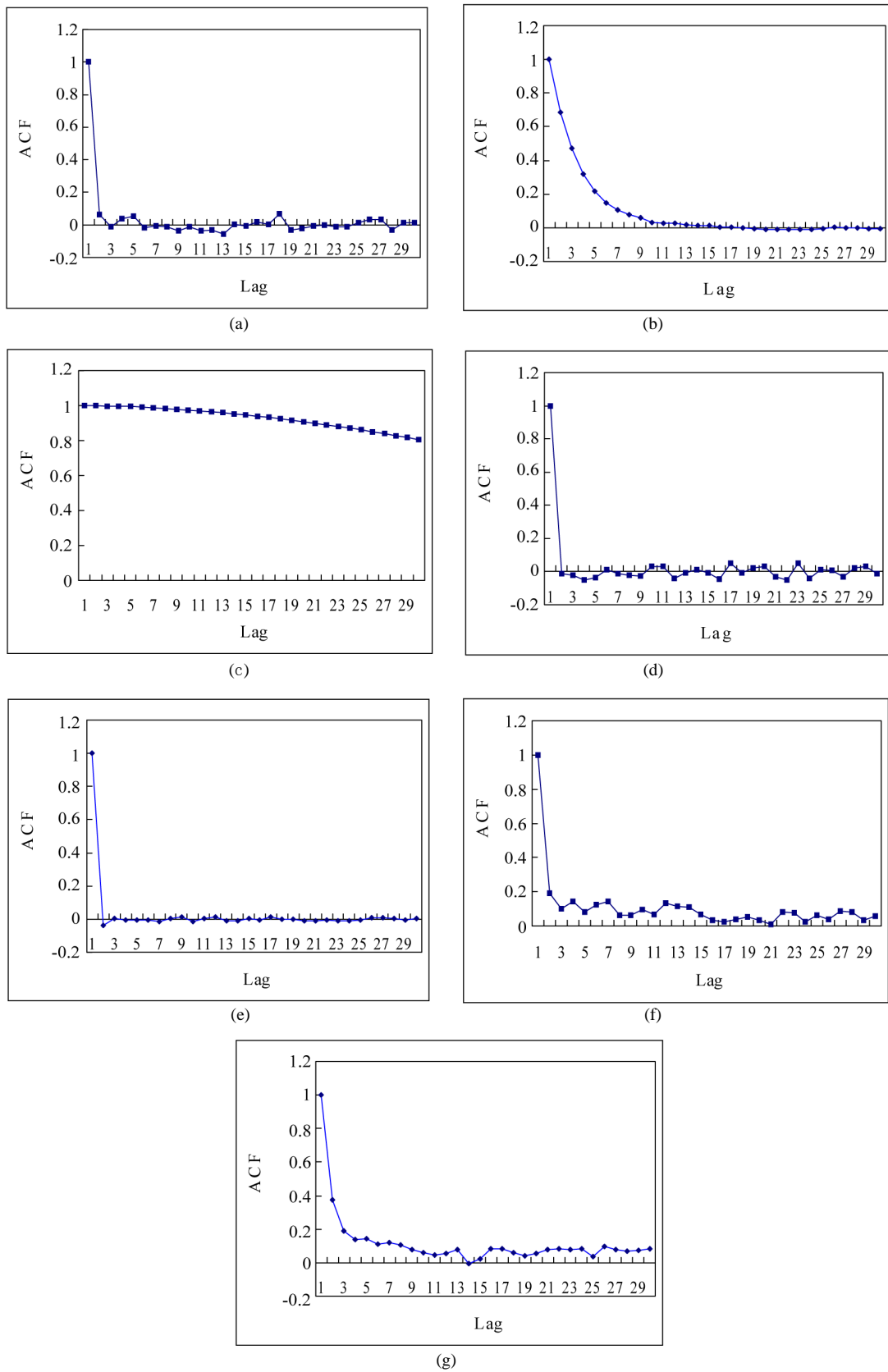


Figure 3. Autocorrelation functions of time series from chaotic and stochastic systems. (a) White noise; (b) AR(1); (c) Three torus; (d) Logistic map; (e) TAR(2,1); (f) FARIMA; (g) FGN.

(1) Rescaled Range and Modified Rescaled Range

The rescaled range and the modified rescaled range methods are applied to the time series for the error estimation of the methods. **Figure 4** shows the comparison of Hurst exponents estimated by the rescaled and modified rescaled range methods. As we can see in **Figure 4**, the slopes estimated by the rescaled and modified rescaled range methods for the time series of short term memory characteristic are similar. However, the slopes for the time series of long term memory show relatively large differences.

If the sample size is small, the error is small but the error is increased as the sample size is increased as shown in **Figure 5**. Even though the white noise is a random series, the Hurst exponent is estimated as 0.604 which represents the long term memory by the rescaled range method. However, if we use the modified rescaled range method the exponent is estimated as 0.55 which represents the short term memory (also see **Figure 6**).

(2) 1/f Power Spectral Density Analysis

We perform 1/f power spectral density (PSD or periodogram) analysis for the estimation of α and obtain the Hurst exponents by the Equations (20) and (21) for the time series mentioned in Section 2.1. **Figure 7** shows the results of PSD analysis for each time series and this method shows relatively reasonable Hurst exponents as we can see in **Table 1**. However, this PSD method may have a problem for the strong persistence system such as the three torus time series which is a quasi-periodic (see **Table 1**).

(3) Detrended Fluctuation Analysis

This section estimates the Hurst exponents by using DFA method. The Hurst exponent is estimated by the Equations of (20) and (21). **Table 2** shows the results of d and Hurst exponents. The DFA method shows the most reasonable Hurst exponent estimates for all-time series as shown in **Table 2**.

(4) Aggregated Variance Time Method

In this section, we estimate Hurst exponent of time series by means of the aggregated variance time method (AVM). As shown in **Figures 8** and **9**, the AVM shows the property which the plotted points are scattered for the partial series is over a certain size, the slopes of the straight lines were obtained by the regression, and the Hurst exponents are estimated by Equation (27). However, the exponents are different from the known values. If we estimated the Hurst exponents after removing the scattered points the estimated exponents were similar with the known values (**Tables 3** and **4**).

Case I: Hurst exponent estimation by the regression with all points;

Case II: Hurst exponent estimation by the regression without the scattered points.

(5) Maximum Likelihood Estimation(MLE)

We estimate the Hurst exponents for the time series by using the statistical package of S-MLE in S-Plus. The results are shown in **Table 5** and we can know that the MLE method gives very reasonable values for each time series. Maximum likelihood estimation for FARIMA models can be performed in several ways [20]. We applied here an approximation in the spectral domain of the Gaussian maximum likelihood function, which was first proposed by [21], for short-memory models.

2.4 Results and Discussions

From the analysis of results, the Hurst exponent has estimated appropriately for the long term memory series by the adjusted and rescaled range methods but it was not for the short term memory like a white noise. The modified rescaled range method estimated the Hurst exponent properly for the short term memory series but it was not proper for the long term memory series. The 1/f PSD method estimated the Hurst exponents properly for the series except for the three torus case. The DFA and MLE methods have shown the reasonable results for the short and long term memory series. Especially, the DFA method is more convenient for the application than the MLE. **Figure 6** shows the comparison for the Hurst exponent estimation results.

3. Application of DFA for Tree-Ring and SOI Series

We found that the DFA is the most appropriate technique for the Hurst exponent estimation for both the short term memory and long term memory. In this section, we analyze 6 tree-ring series at USA sites and the SOI (Southern Oscillations Index) by means of DFA and the BDS statistic is used for nonlinearity test of the series. Especially, the BDS statistic is used for the nonlinearity of tree ring series tested by [14], who showed the tree ring series has random characteristics. The random characteristic means short term memory and may represent linear stochasticity of the series. Therefore we will examine the memory of tree ring series in the following sections.

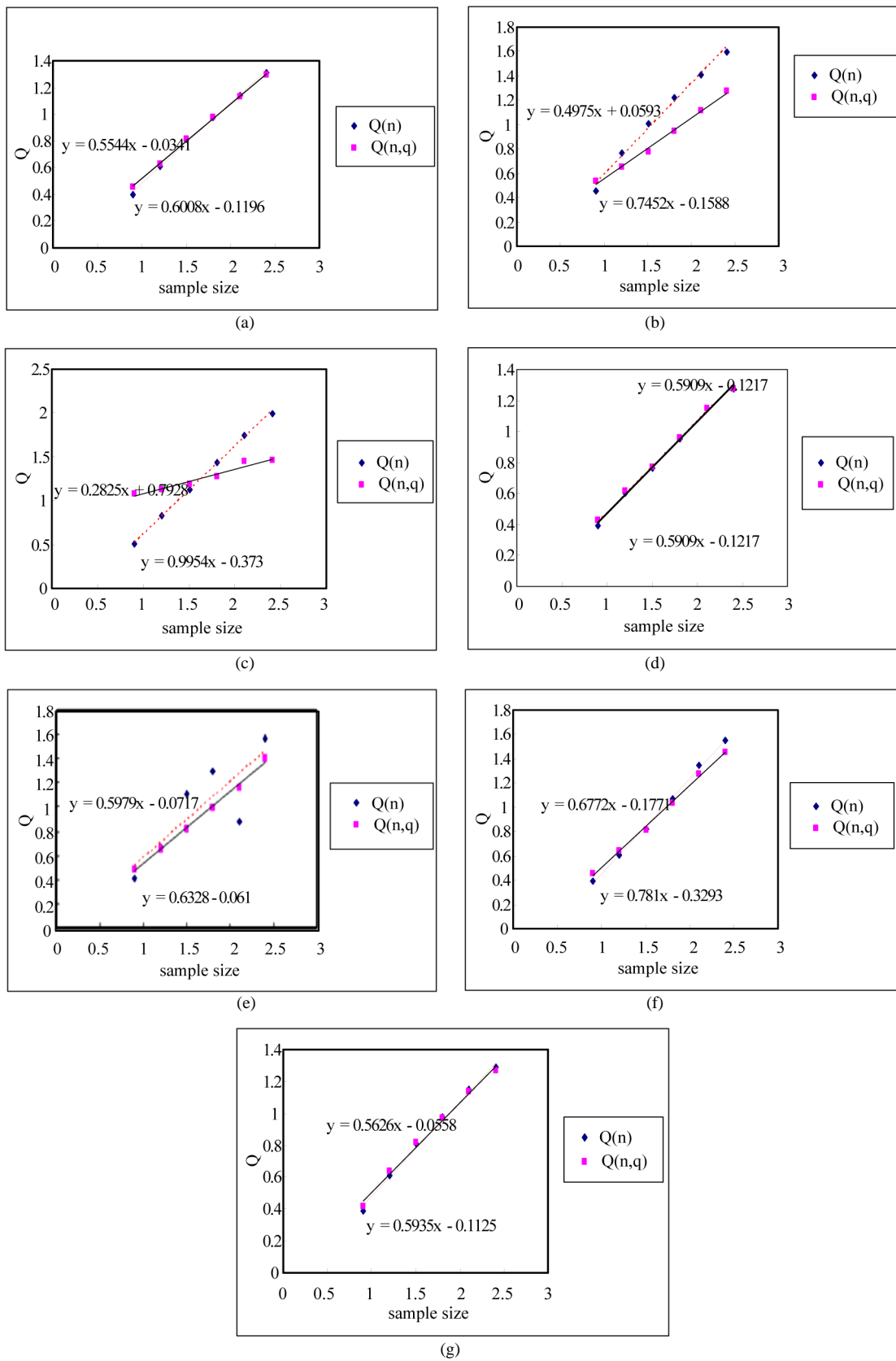


Figure 4. Estimation of hurst exponent by rescaled range and modified rescaled range methods. (a) White noise; (b) AR(1); (c) Three torus; (d) Logistic map; (e) FGN; (f) FARIMA; (g) TAR(2,1).

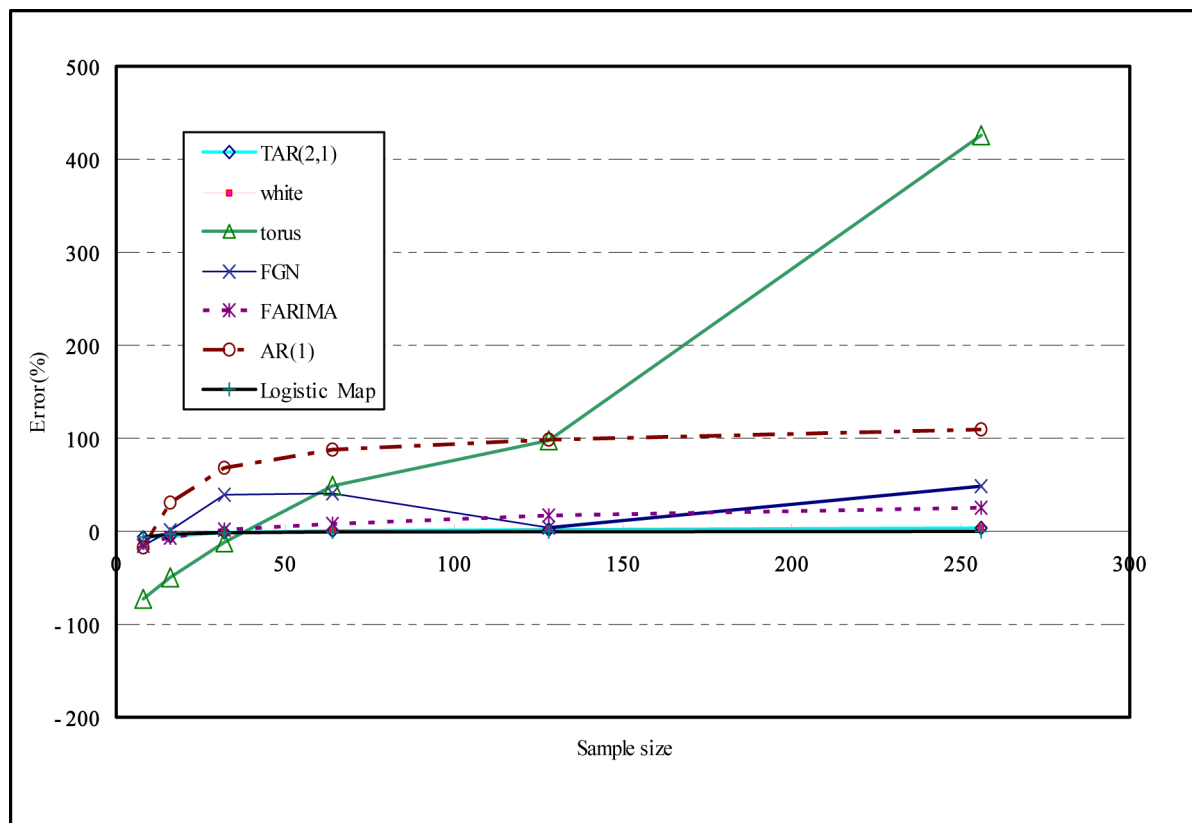


Figure 5. Errors for the rescaled range and modified rescaled range methods.

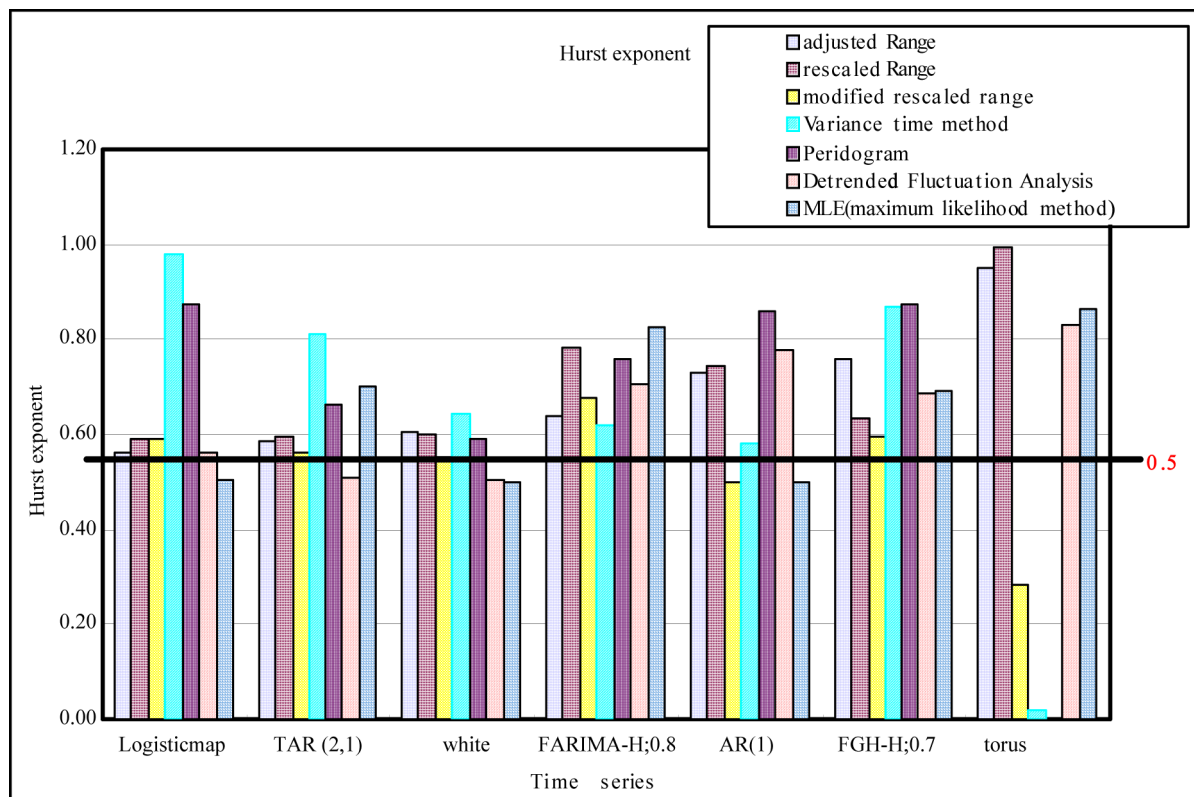


Figure 6. Comparison of the Hurst exponent estimates from various methods.

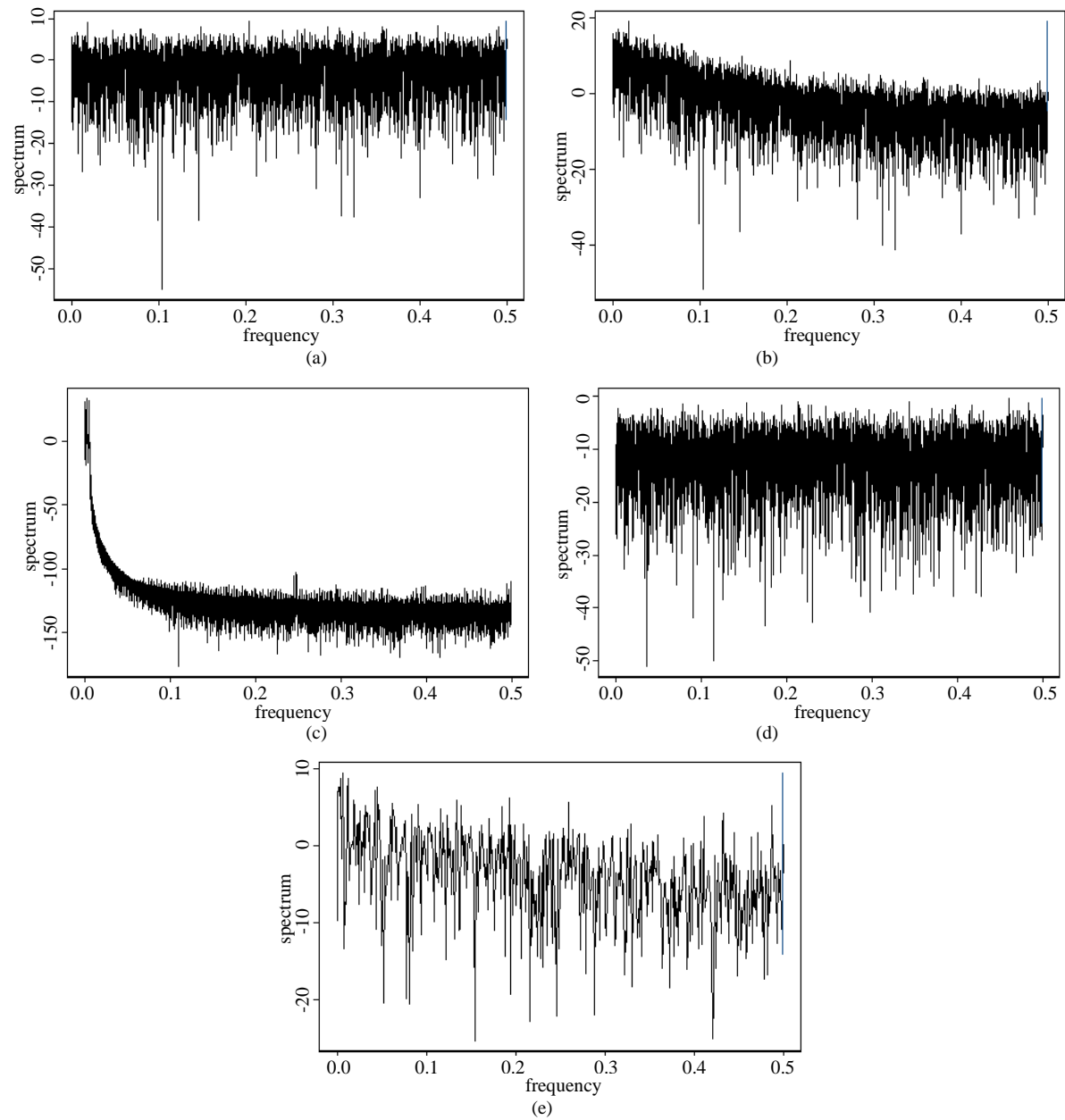


Figure 7. $1/f$ Power spectral density analysis. (a) White noise; (b) AR(1); (c) TAR(2,1); (d) FARIMA; (e) FGN.

Table 1. $1/f$ spectral slope and hurst exponent.

	White noise	AR(1)	Three torus	Logistic map	TAR(2,1)	FARIMA	FGN
α	0.180	0.716	5.287	0.745	0.323	0.519	0.746
H	0.590	0.858	∞	0.873	0.661	0.760	0.873

Table 2. d slope and hurst exponent.

	White noise	AR(1)	Three torus	Logistic map	TAR(2,1)	FARIMA	FGN
d	0.506	0.78	1.832	0.438	0.510	0.703	0.684
Hurst exponent	0.506	0.78	0.832	0.562	0.510	0.703	0.684

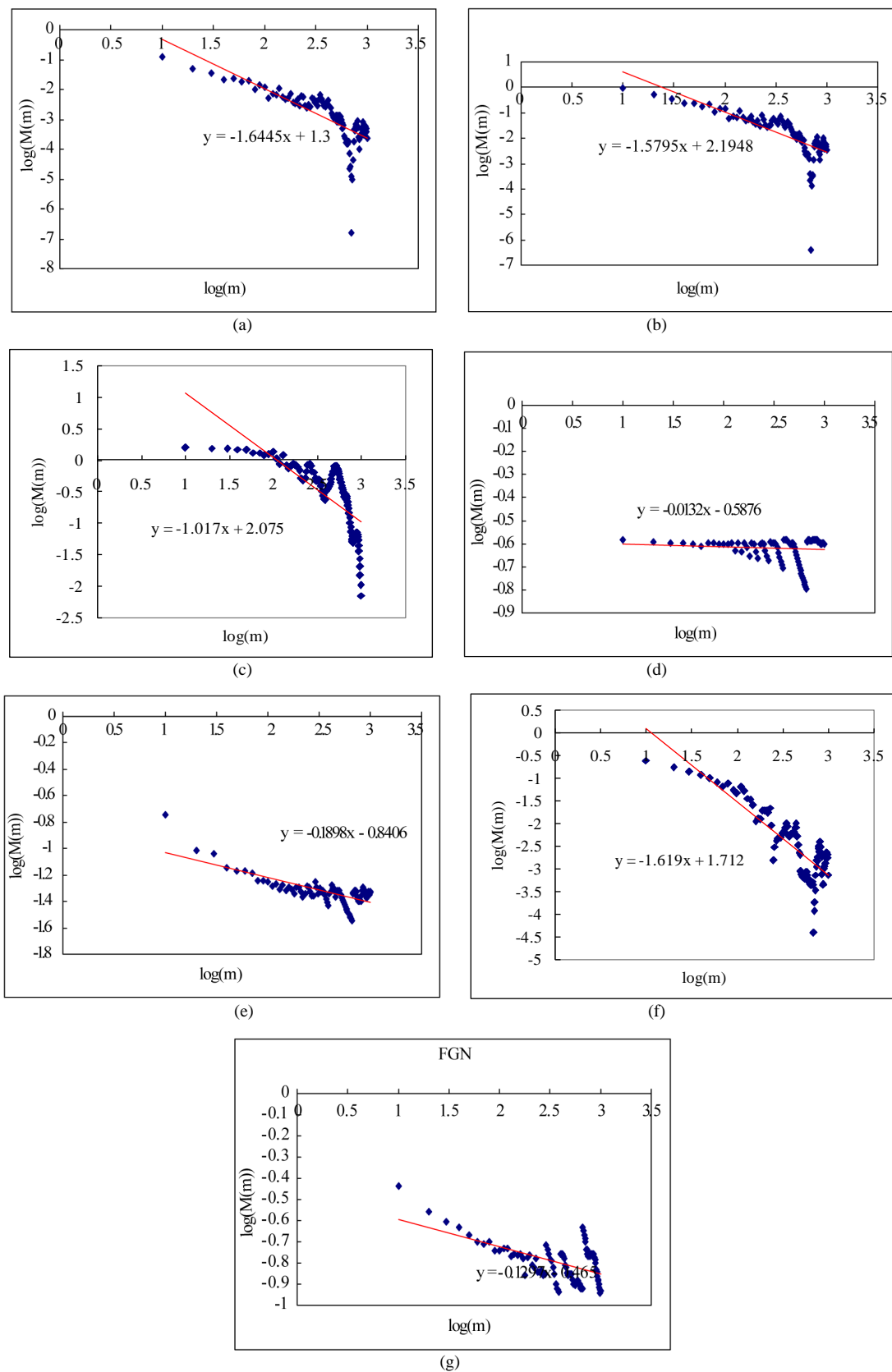


Figure 8. Estimation of Hurst exponent by Aggregated Variance Time method (case-I). (a) White noise; (b) AR(1); (c) Three torus; (d) Logistic map; (e) TAR(2,1); (f) FARIMA; (g) FGN.

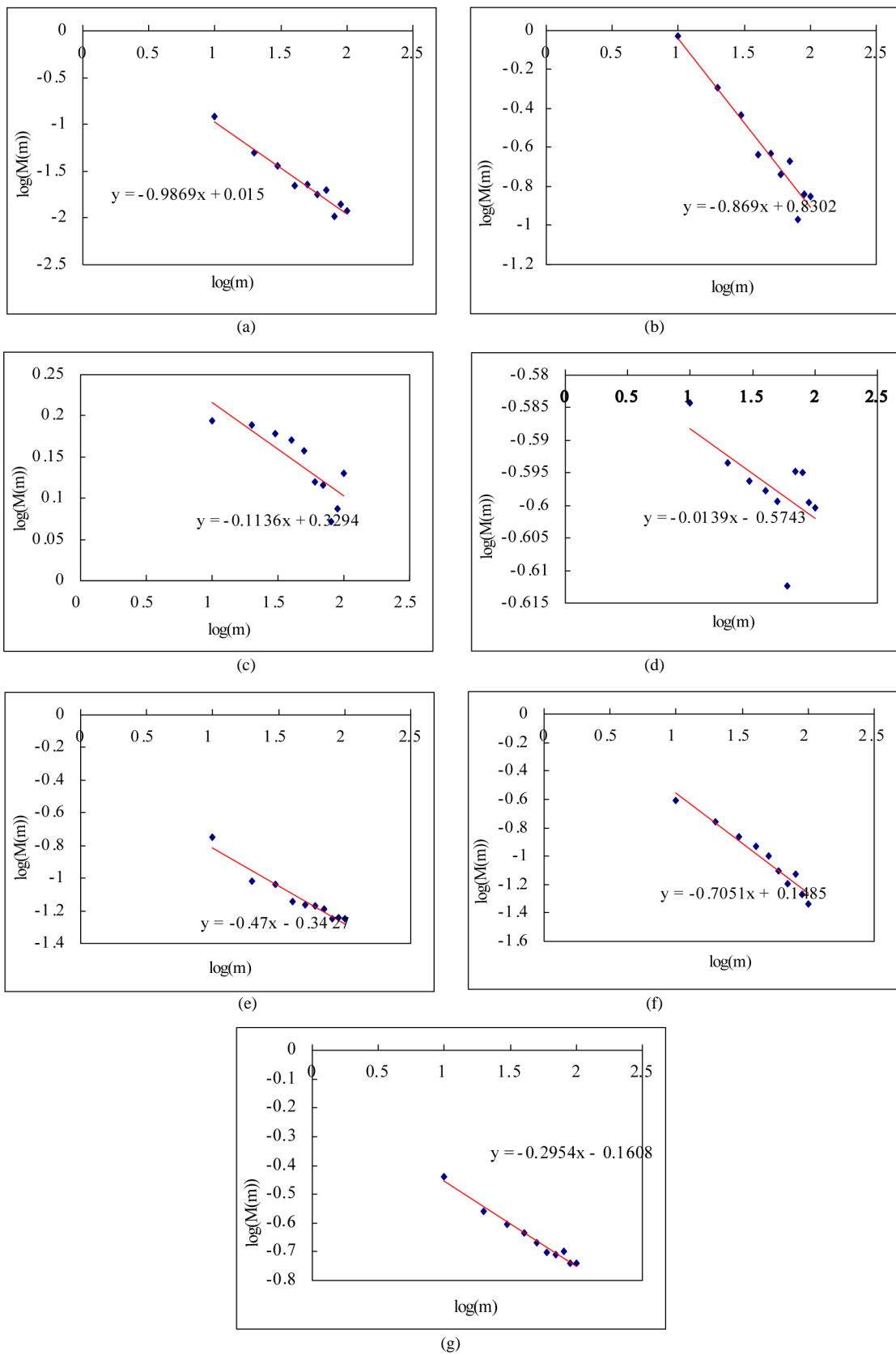


Figure 9. Estimation of Hurst exponent by aggregated variance time method (case-II). (a) White noise; (b) AR(1); (c) Three torus; (d) Logistic map; (e) FARIMA; (f) FARIMA; (g) FGN.

Table 3. β slope and hurst exponent (case-I).

	White noise	AR(1)	Three torus	Logistic map	TAR(2,1)	FARIMA	FGN
β	1.645	1.580	1.017	0.013	0.189	1.619	0.129
Hurst exponent	0.170	0.210	0.491	0.934	0.905	0.190	0.930

$$H = 1 - 2\beta.$$

Table 4. β slope and hurst exponent (case-II).

	White noise	AR(1)	Three torus	Logistic map	TAR(2,1)	FARIMA	FGN
β	0.987	0.869	0.114	0.014	0.470	0.705	0.295
Hurst exponent	0.506	0.566	0.943	0.993	0.765	0.648	0.852

$$H = 1 - 2\beta.$$

Table 5. d slope and hurst exponent.

	White noise	AR(1)	Three torus	Logistic map	TAR(2,1)	FARIMA	FGN
d	0.005	0.0004	0.364	0.023	0.199	0.326	0.193
Hurst exponent	0.5005	0.5004	0.864	0.502	0.699	0.826	0.693

3.1. Descriptions of Tree-Ring and SOI Series

The climatic characteristics of an area are reflected in the growth of the trees. Consequently, the characteristics of tree are studied to determine the long-term climate behavior of a region. In this section, 6 tree ring series from USA are analyzed. The data sets of tree ring series consist of the period of 1892 to 1980 and sites SACNDX, Spring, Freder, Calam, Hager, Dalton in California and Arizona. Tree-ring data are recorded as local climate properties and annual time scale. Each time series plot is shown in [Figure 10](#).

The SOI is defined as the normalized pressure difference between Tahiti and Darwin. SOI values are calculated using the monthly mean sea level pressure (MSLP) data at papeete, Tahiti (149.6°W, 17.5°S) and Darwin, Australia (130.9°W, 12.4°S). In this study, we use the monthly SOI data from January of 1951 to December of 1999. [Figure 11](#) shows the time series plot of SOI.

3.2. Hurst's Memory for Tree-Ring and SOI Data by DFA

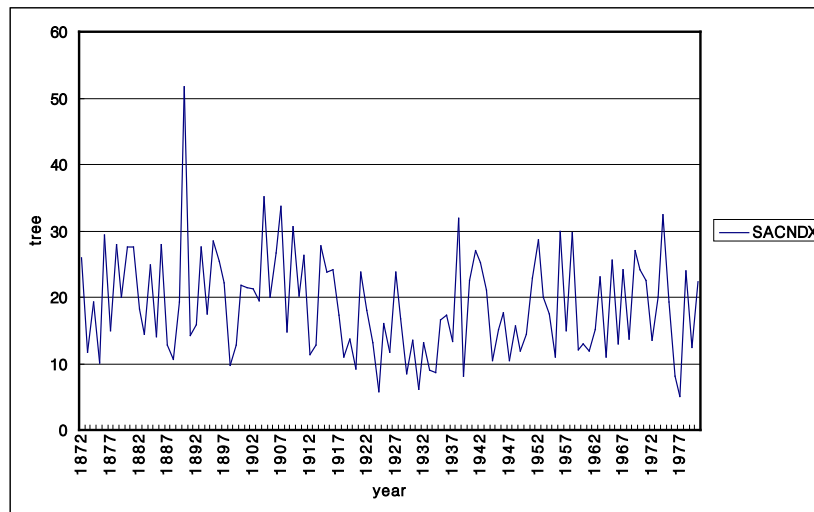
[Figure 12](#) shows the comparison of regression plots and results are shown in [Table 6](#). [Figure 13](#) shows the comparison of the autocorrelation (ACF) for tree-ring and SOI data. From the results, we found that the SOI series is time series which has a long term memory of $H = 0.92$. However, contrary to earlier work of [14], all the tree-ring series are not random and some of them show strong persistence from our analysis. A certain tree-ring series shows a long term memory of $H = 0.97$. Therefore, we can say that the tree ring and SOI series may show long term memory. And if the tree ring series are random as tested in [14], it may have linear stochasticity but the tree ring series which shows long term memory may have its nonlinearity and this could be modeled by the nonlinear stochastic models. Therefore, we would like to examine the nonlinearity of the tree ring and SOI series.

3.3. Nonlinearity for Tree-Ring and SOI Series

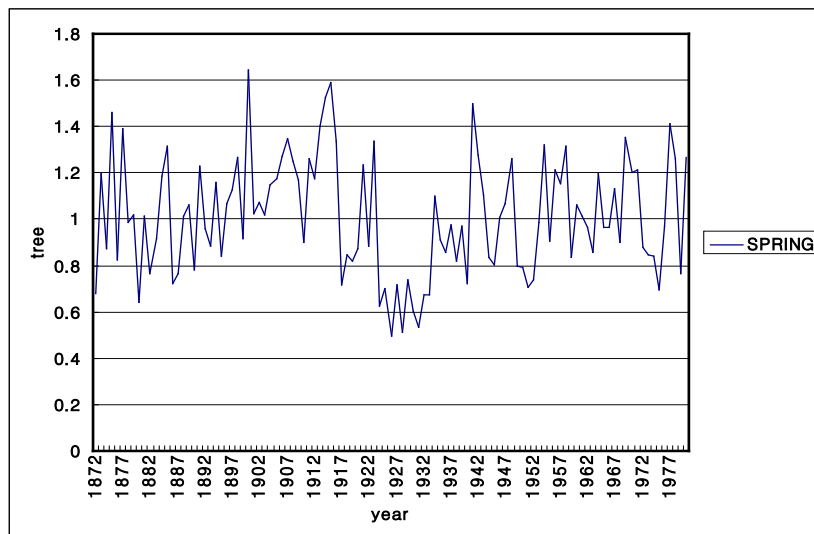
The BDS statistic is derived from the correlation integral and has its origins in the recent work on deterministic nonlinear dynamics and chaos theory. The method of delays can be used to embed a scalar time series $\{x_i\}, i = 1, 2, \dots, N$ into an m -dimensional space as follows

$$\mathbf{x}_i = (x_i, x_{i+t}, \dots, x_{i+(m-1)t}), \quad \mathbf{x}_i \in R^m. \quad (29)$$

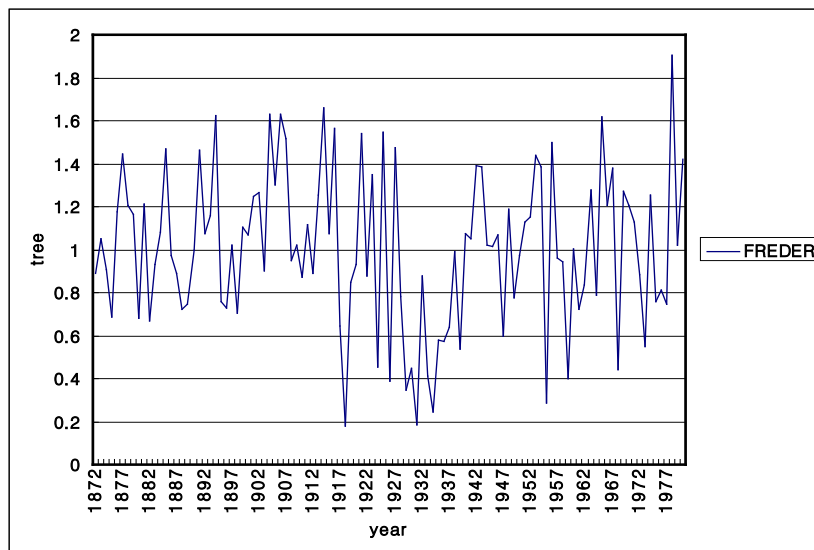
where t is the index lag. The correlation integral at embedding dimension m is given by



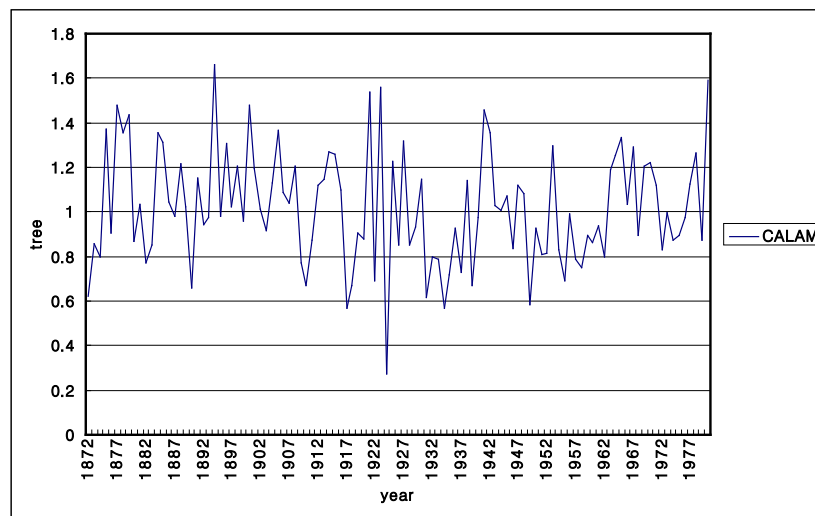
(a)



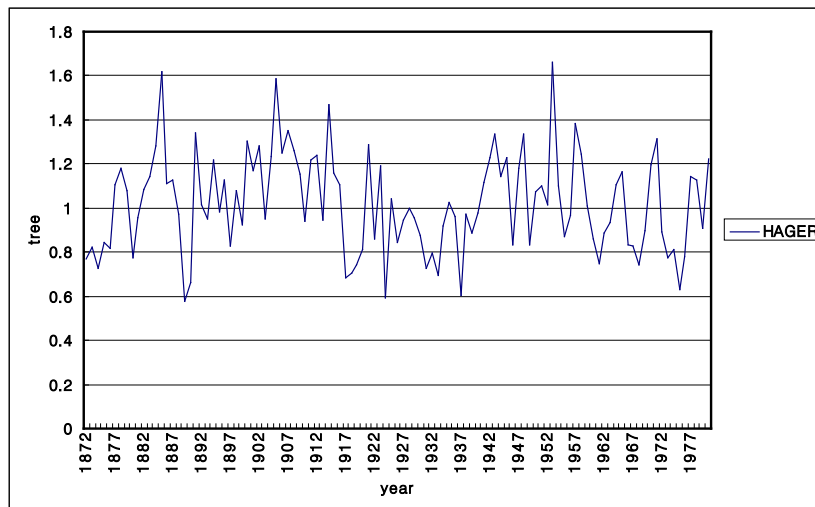
(b)



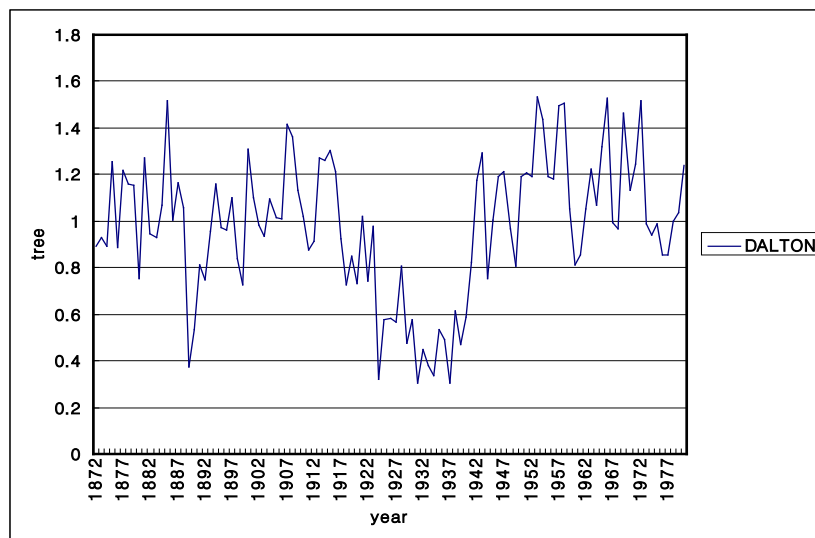
(c)



(d)



(e)



(f)

Figure 10. Time series plots of the tree-rings. (a) SACNDX; (b) Spring; (c) Freder; (d) Calam; (e) Hager; (f) Dalton.

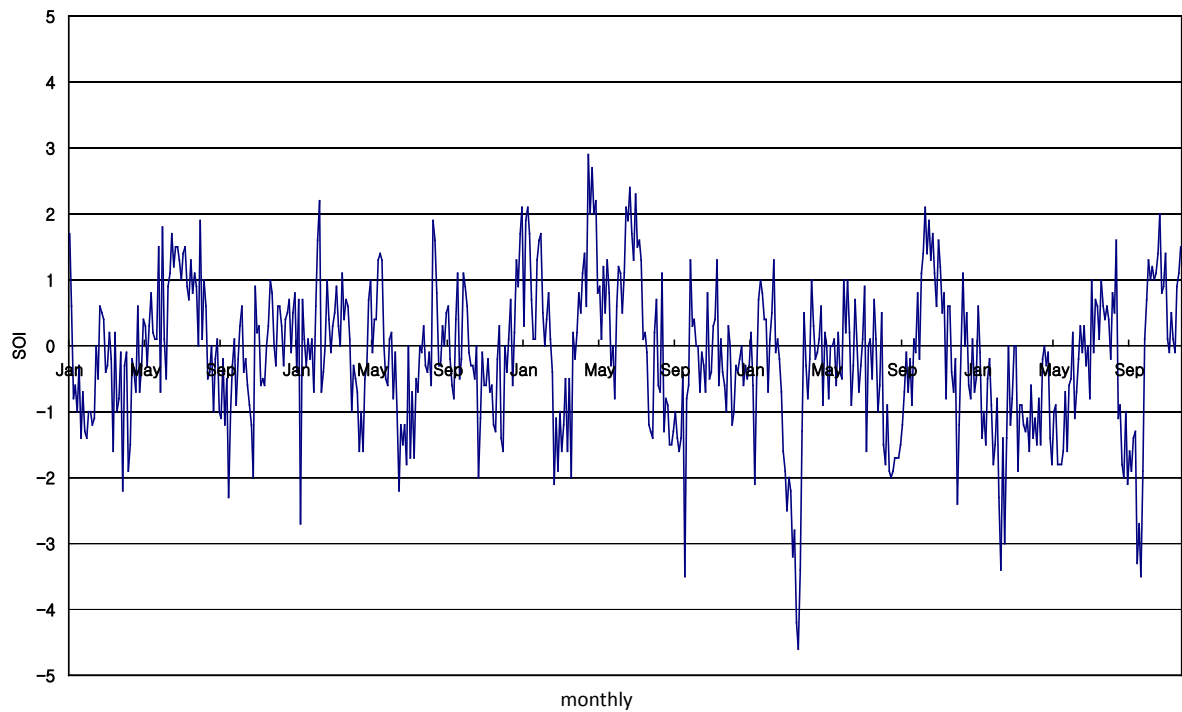


Figure 11. Time series plots of SOI.

Table 6. Estimations of the Hurst exponent by DFA for the tree rings and SOI series.

ataD	malaC	notlaD	rederF	regaH	xdncaS	gnirpS	IOS
H	0.58	0.91	0.67	0.67	0.53	0.77	0.9

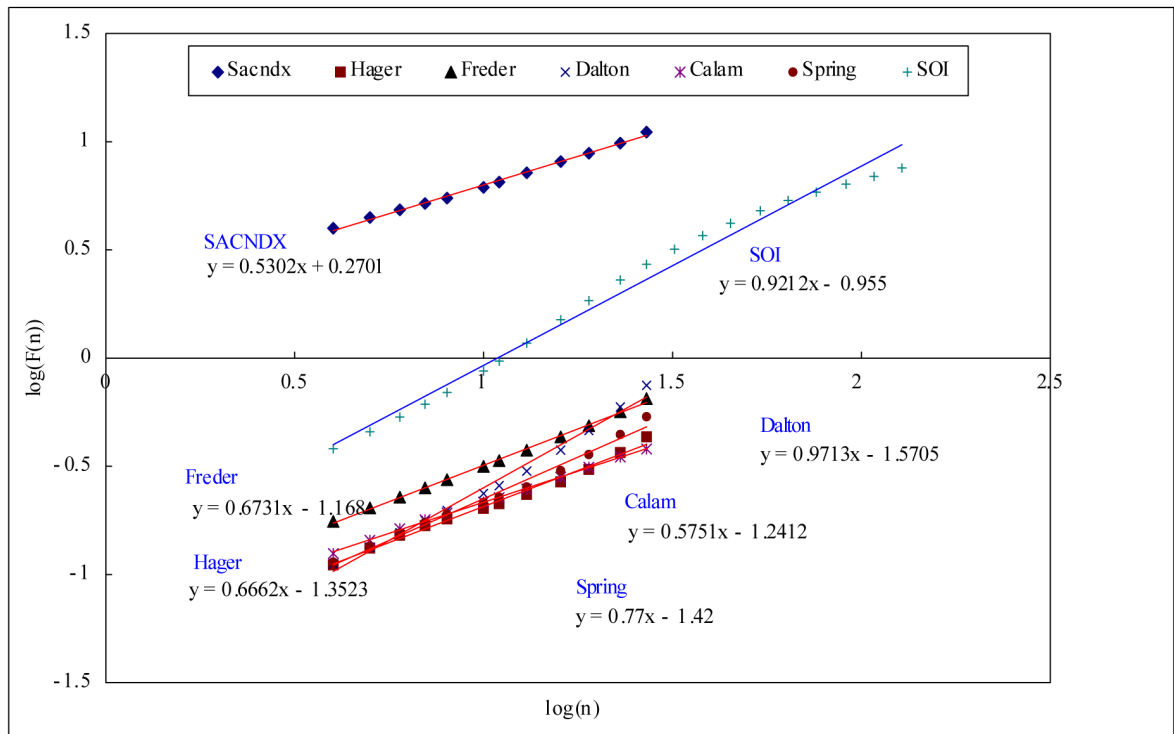


Figure 12. Comparison for regression plots by DFA.

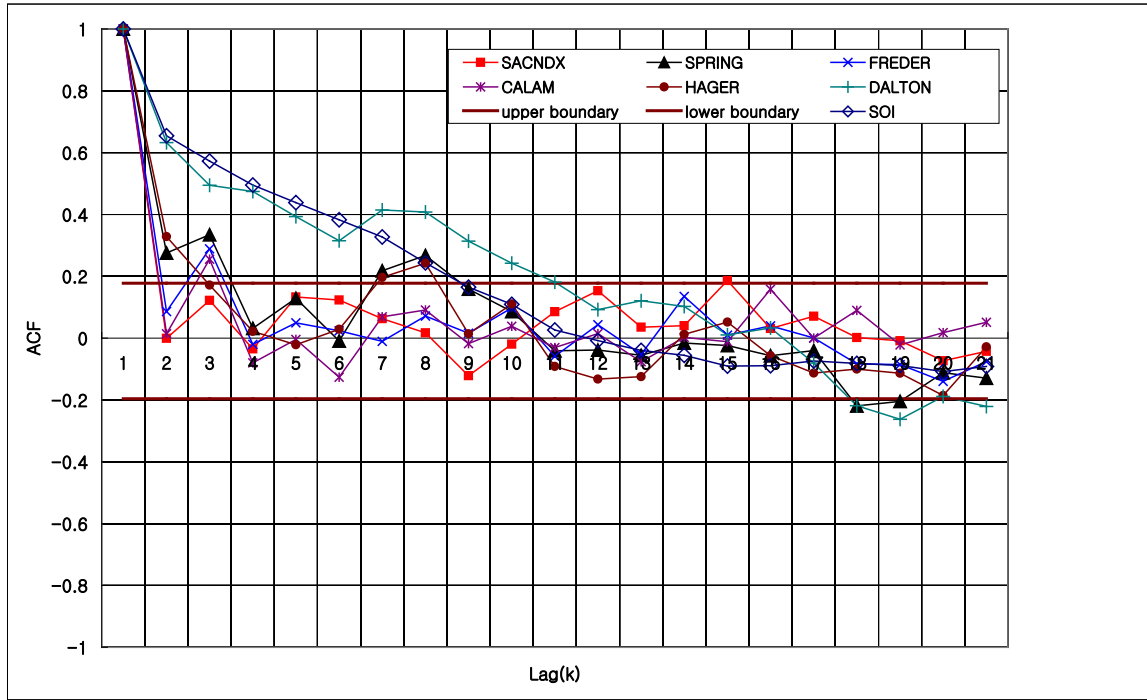


Figure 13. Comparison of the autocorrelation for tree-ring and SOI series.

$$C(m, N, r) = \frac{2}{M(M-1)} \sum_{1 \leq i < j \leq M} \Theta(r - \|x_i - x_j\|), \quad r > 0. \quad (30)$$

$$\begin{aligned} \Theta(a) &= 0, & \text{if } a \leq 0 \\ \Theta(a) &= 1, & \text{if } a \geq 0. \end{aligned}$$

where N is the size of the data sets, $M = N - (m-1)t$ is the number of embedded points in m -dimensional space, and $\|\cdot\|$ denotes the sup-norm. $C(m, N, r)$ measures the fraction of the pairs of points x_i , $i = 1, 2, \dots, M$, whose sup-norm separation is no greater than r . If the limit of $C(m, N, r)$ as $N \rightarrow \infty$ exists for each r , we write the fraction of all state vector points that are within r of each other as $C(m, r) = \lim_{N \rightarrow \infty} C(m, N, r)$.

If the data is generated by a strictly stationary stochastic process which is absolutely regular, then this limit exists. In this case the limit is as follows

$$C(m, r) = \int \int_{\mathcal{X}} \Theta(r - \|x - y\|) dF(x) dF(y), \quad r > 0. \quad (31)$$

When the process is IID, and since $\Theta(r - \|x - y\|) = \prod_{k=1}^m \Theta(r - |x_k - y_k|)$, Equation (31) implies that $C(m, r) = C^m(1, r)$. Also $C(m, r) - C^m(1, r)$ has asymptotic normal distribution, with zero mean and variance as follows

$$\begin{aligned} \sigma^2(m, M, r)/4 &= m(m-1)C^{2(m-1)}(K - C^2) + K^m - C^{2m} \\ &+ 2 \sum_{i=1}^{m-1} \left[C^{2i} (K^{m-i} - C^{2(m-i)}) - mC^{2(m-i)} (K - C^2) \right], \end{aligned} \quad (32)$$

We can consistently estimate the constants C by $C(1, r)$ and K by

$$K(m, M, r) = \frac{6}{M(M-1)(M-2)} \times \sum_{1 \leq i < j \leq M} \left[\Theta(r - \|x_i - x_j\|) \Theta(r - \|x_i - x_j\|) \right]. \quad (33)$$

Under the IID hypothesis, the BDS statistic for $m > 1$ is defined as

$$\text{BDS}(m, M, r) = \frac{\sqrt{M}}{\sigma} [C(m, r) - C^m(1, r)]. \quad (34)$$

has a limiting standard normal distribution under the null hypothesis of IID as $M \rightarrow \infty$ and obtain its critical values using the standard normal distribution.

Before applying the BDS statistic, the first addressed issue is what region of “ r ” yields BDS statistic that are well approximated by the asymptotic distribution. As the sample size is increased, the distribution of the BDS statistic becomes more normal. So the minimal number of data must be provided. Next, the region of embedding dimension “ m ” should be suggested. If the sample size is fixed, we expect the finite sample property to worsen as “ m ” increases. This study follows the recommendation of [22] for selecting the ranges of m , r , N . Therefore, 500 or more observations are prepared and the embedding dimension m is used in the range of $2 \leq m \leq 5$. Then, the value of “ r ” is selected as the half standard deviations of the data sets. The BDS statistic is a powerful tool for distinguishing random time series from the time series generated by nonlinear systems [22].

We use the BDS statistic for testing randomness of tree-ring and SOI data series and the results are shown in Table 7. From the results, the SOI and certain tree-ring are representing their nonlinearities and thus could be modeled by the nonlinear type models [23].

4. Summary and Conclusions

This study has used the seven methods for the estimation of Hurst exponent and compared the results by applying each method to the generated time series from chaos and stochastic systems which have different characteristics. Then, this study discusses the advantages and disadvantages of the techniques and also the limitations of them.

Table 7. The BDS statistics for tree-ring and SOI.

r	Embedding dimension (m)	Tree-ring data						SOI
		Calam	Dalton	Freder	Hager	Sacndx	Spring	
0.5σ	2	−1.44(A)	8.07(R)	−1.08(A)	2.27(R)	−1.44(A)	3.01(R)	24.42(R)
	3	−1.05(A)	6.54(R)	−1.08(A)	2.48(R)	−2.24(R)	2.76(R)	26.75(R)
	4	−1.8(A)	5.21(R)	0.37(A)	0.67(A)	1.76(A)	1.61(A)	28.97(R)
	5	−2.64(R)	4.37(R)	3.68(R)	−3.73(R)	1.34(A)	0.98(A)	31.59(R)
1.0σ	2	1.6(A)	9.68(R)	0.07(A)	0.73(A)	−1.78(A)	3.19(R)	23.33(R)
	3	1.32(A)	9.07(R)	0.47(A)	0.67(A)	−2.18(R)	3.76(R)	25.09(R)
	4	1.03(A)	8.38(R)	1.67(A)	0.37(A)	−0.13(A)	3.42(R)	26.7(R)
	5	0.92(A)	8.61(R)	2.41(R)	0.49(A)	0.15(A)	3.99(R)	28.39(R)
1.5σ	2	1.7(A)	8.87(R)	0.74(A)	1.78(A)	−3.59(R)	3.5(R)	21.84(R)
	3	1.79(A)	8.7(R)	1.4(A)	1.03(A)	−3.12(R)	4.55(R)	22.86(R)
	4	1.55(A)	8.46(R)	2.2(R)	0.38(A)	−2.51(R)	4.56(R)	23.33(R)
	5	1.91(A)	8.56(R)	2.64(R)	0.51(A)	−2.37(R)	4.59(R)	23.61(R)
2.0σ	2	2.86(R)	9.86(R)	0.32(A)	0.93(A)	−3.11(R)	4.00(R)	21.37(R)
	3	2.87(R)	9.81(R)	0.91(A)	0.06(A)	−2.55(R)	4.97(R)	22.27(R)
	4	2.53(R)	9.54(R)	1.92(A)	−0.58(A)	−2.21(R)	4.78(R)	22.1(R)
	5	2.39(R)	9.52(R)	2.46(R)	−0.68(A)	−1.9(A)	4.56(R)	21.67(R)

Confidence interval: −1.96, +1.96 ($\alpha = 0.05$); A: accepted; R: rejected.

1) The adjusted and rescaled range methods were proper for the long-term memory but the modified rescaled range method was okay for the short-term memory. These methods have shown that the error was increased as the sample size was increased. Especially the error is increased for the series which the ACF is large.

2) The aggregated variance time method has shown that the plotted points on the proper scale were scattered. And so the Hurst exponent was appropriately estimated when we used the regression after removing the scattered points.

3) The $1/f$ PSD method has shown that it is a reasonable technique for the Hurst exponent estimation but it was not proper for a very strongly auto correlated series like a three tours system.

4) The DFA and MLE methods have shown that they are the most appropriate techniques for the Hurst exponent estimation for both the short term memory and long term memory.

From these results, we found that the DFA is the most appropriate technique for the Hurst exponent estimation for both the short term memory and long term memory. We analyzed the 6 tree-ring and SOI series for their memory tests by means of DFA and then the BDS statistic has used for nonlinearity test of the series. Our analysis has obtained the following conclusions

5) We found that SOI series is nonlinear time series which has a long term memory of $H = 0.92$. Contrary to earlier work of [14], all the tree-ring series are not random from our analysis. A certain tree ring series showed a long term memory of $H = 0.97$ and nonlinear property. Therefore, we can say that the SOI and tree-ring series may show long term memory and nonlinearity.

REFERENCES

- [1] H. Hurst, "Long-Term Storage Capacity of Reservoirs," *Translation of the American Society of Civil Engineer*, Vol. 116, 1951, pp. 770-799.
- [2] D. Koutsoyiannis and A. Efstratiadis, "Climate Change Certainty versus Climate Uncertainty and Inferences in Hydrological Studies and Water Resources Management," *1st General Assembly of the European Geosciences Union, Geophysical Research Abstracts*, European Geosciences Union, Austria, Vol. 6, 2004.
- [3] B. Mandelbrot and J. Wallis, "Robustness of the Rescaled Range R/S in the Measurement of Noncyclic Long Run Statistical Dependence," *Water Resources Research*, Vol. 5, No. 5, 1969, pp. 967-988. <http://dx.doi.org/10.1029/WR005i005p00967>
- [4] P. Herman, "Physiological Time Series: Distinguishing Fractal Noises from Motions," *European Journal of Physiology*, Vol. 439, No. 4, 1999, pp. 403-415.
- [5] F. Pallikari, "Rescaled Range Analysis of Random Events," *Journal of Scientific Exploration*, Vol. 13, No. 1, 1999, pp. 25-40.
- [6] T. Bigger, C. Richard, A. B. Steinman, M. Linda, L. Joseph and J. Richard, "Power Law Behavior of RR-Interval Variability in Healthy Middle-Aged Persons, Patients with Recent Acute Myocardial Infarction, and Patients with Heart Transplants," *Circulation*, Vol. 93, No. 12, 1996, pp. 2142-2151. <http://dx.doi.org/10.1161/01.CIR.93.12.2142>
- [7] M. S. Taqqu and V. Teverovsky, "Estimation for Long Range Dependence: An Empirical Study," *Fractals*, Vol. 3, No. 4, 1995 pp. 785-798. <http://dx.doi.org/10.1142/S0218348X95000692>
- [8] Y. Zhao, "Self-Similarity in High Performance Network Analysis," University of Missouri-Columbia, Columbia, 1998.
- [9] S. Zafer and T. Sirin, "Traffic Engineering for Multimedia Networks; Data Collection on the Internet, Extensions to Wireless," NSF industry/University Co-Operative Research Center for Digital Video & Media, 1999.
- [10] M. A. Ausloos, "Statistical Physics in Foreign Exchange Currency and Stock Markets," *Physica A*, Vol. 285, No. 1-2, 2000, pp. 48-65. [http://dx.doi.org/10.1016/S0378-4371\(00\)00271-5](http://dx.doi.org/10.1016/S0378-4371(00)00271-5)
- [11] J. W. Kantelhardt, B. E. Koscielny, H. A. Rego, S. Havlin and A. Bunde, "Detecting Long Range Correlations with Detrended Fluctuation Analysis," *Physica A*, Vol. 295, No. 3-4, 2001, pp. 441-454. [http://dx.doi.org/10.1016/S0378-4371\(01\)00144-3](http://dx.doi.org/10.1016/S0378-4371(01)00144-3)
- [12] C. M. Kendzioriski, "Evaluation Maximum Likelihood Estimation Methods to Determine the Hurst Coefficient," *Physica A*, Vol. 273, No. 3-4, 1999, pp. 439-451. [http://dx.doi.org/10.1016/S0378-4371\(99\)00268-X](http://dx.doi.org/10.1016/S0378-4371(99)00268-X)
- [13] A. W. Lo, "Long Term Memory in Stock Market Prices," *Econometrica*, Vol. 59, No. 5, 1991, pp. 1279-1313. <http://dx.doi.org/10.2307/2938368>
- [14] A. R. Rao and D. Bhattacharya, "Hypothesis Testing for Long Term Memory in Hydrologic Series," *Journal of Hydrology*, Vol. 216, No. 3-4, 1999, pp. 183-196. [http://dx.doi.org/10.1016/S0022-1694\(99\)00005-0](http://dx.doi.org/10.1016/S0022-1694(99)00005-0)
- [15] A. R. Rao and D. Bhattacharya, "Effect of Short-Term Memory on Hurst Phenomenon," *Journal of Hydrologic Engineering*, Vol. 6, No. 2, 2001, pp. 125-131. [http://dx.doi.org/10.1061/\(ASCE\)1084-0699\(2001\)6:2\(125\)](http://dx.doi.org/10.1061/(ASCE)1084-0699(2001)6:2(125))
- [16] V. Teverovsky, M. S. Taqqu and W. Willinger, "Acritical Look at Lo's Modified R/S Statistic," *Journal of Statistical Planning and Inference*, Vol. 80, No. 1-2, 1998, pp. 211-227.
- [17] P. Philips, "Time Series Regression with a Unit Roof," *Econometrica*, Vol. 55, No. 2, 1987, pp. 703-705.

<http://dx.doi.org/10.2307/1913237>

- [18] MathSoft, "S-Plus2000; Guide to Statistical and Mathematical Analysis," StatSci Division, Seattle, 2000.
- [19] J. D. Salas and R. Pielke, "Stochastic Characteristics and Modeling of Hydroclimatic Processes," In: T. D. Potter and B. Colman, Eds., *Handbook of Weather, Climate, and Water*, Chapter 32, John Wiley & Sons, New York, 2002, pp. 585-603.
- [20] J. Beran, "Statistics for Long-Memory Processes," Chapman and Hall, New York, 1994.
- [21] P. Whittle, "Estimation and Information in Stationary Time Series," *Arkiv för Matematik*, Vol. 2, No. 5, 1953 pp. 423-434.
<http://dx.doi.org/10.1007/BF02590998>
- [22] H. S. Kim, D. S. Kang and J. H. Kim, "The BDS Statistic and Residual Test," *Stochastic Environmental Research and Risk Assessment*, Vol. 17, No. 1-2, 2003, pp. 104-115. <http://dx.doi.org/10.1007/s00477-002-0118-0>
- [23] J. H. Ahn and H. S. Kim, "Nonlinear Modeling of El Nino/Southern Oscillation Index," *Journal of Hydrologic Engineering-ASCE*, Vol. 10, No. 1, 2005, pp. 8-15. [http://dx.doi.org/10.1061/\(ASCE\)1084-0699\(2005\)10:1\(8\)](http://dx.doi.org/10.1061/(ASCE)1084-0699(2005)10:1(8))