



# Article Uncovering Hidden Insights with Long-Memory Process Detection: An In-Depth Overview

Hossein Hassani<sup>1,\*</sup>, Masoud Yarmohammadi<sup>2</sup> and Leila Marvian Mashhad<sup>2</sup>

- <sup>1</sup> The Research Institute of Energy Management and Planning (RIEMP), University of Tehran, Tehran 19395-4697, Iran
- <sup>2</sup> Department of Statistics, Payame Noor University, Tehran 19395-4697, Iran; masyar@pnu.ac.ir (M.Y.); leila.marveian@imamreza.ac.ir (L.M.M.)
- \* Correspondence: hassani.stat@gmail.com

**Abstract:** Long-memory models are frequently used in finance and other fields to capture long-range dependence in time series data. However, correctly identifying whether a process has long memory is crucial. This paper highlights a significant limitation in using the sample autocorrelation function (ACF) to identify long-memory processes. While the ACF establishes the theoretical definition of a long-memory process, it is not possible to determine long memory by summing the sample ACFs. Hassani's -1/2 theorem demonstrates that the sum of the sample ACF is always -1/2 for any stationary time series with any length, rendering any diagnostic or analysis procedures that include this sum open to criticism. The paper presents several cases where discrepancies between the empirical and theoretical use of a long-memory process are evident, based on real and simulated time series. It is critical to be aware of this limitation when developing models and forecasting. Accurately identifying long-memory processes is essential in producing reliable predictions and avoiding incorrect model specification.

**Keywords:** long-memory process; sum of sample autocorrelation function; Hassani's -1/2 theorem; spectral density; time series; autocorrelation



Citation: Hassani, Hossein, Masoud Yarmohammadi, and Leila Marvian Mashhad. 2023. Uncovering Hidden Insights with Long-Memory Process Detection: An In-Depth Overview. *Risks* 11: 113. https://doi.org/ 10.3390/risks11060113

Academic Editors: Mogens Steffensen and Roy Cerqueti

Received: 30 March 2023 Revised: 7 June 2023 Accepted: 9 June 2023 Published: 16 June 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

## 1. Introduction

Long-memory time series are characterized by having an autocorrelation function (ACF) that decays to zero at a slow polynomial rate as the lag increases. This means that the correlations between observations at different time steps persist over a long period of time, leading to persistent patterns in the data. This property of long-memory time series is useful in various fields, as it allows researchers to model the persistence of certain patterns in the data and make better predictions based on past observations (see, for example, Doukhan et al. 2003; Beran et al. 2013; Zivot and Wang 2013). The concept of long-memory time series was first introduced by (Wei 2006) and (Tsay 2010) (for recent research, refer to Beran 1994 and Das and Bhattacharya 2021). The fractionally differenced process is a type of long-memory time series that is characterized by the fractional difference operator  $(1 - B)^{\alpha}$ , where *B* is the backshift operator and  $\alpha$  is the fractional differencing parameter; the process is defined by

$$(1-B)^{\alpha} y_t = a_t, \qquad -0.5 < \alpha < 0.5 \tag{1}$$

where  $a_t$  is a white noise series. In this equation,  $y_t$  represents the observed time series and  $a_t$  is a white noise series, which is a series of uncorrelated random variables with a mean of zero and a constant variance. The fractional difference operator  $(1 - B)^{\alpha}$  removes any short-term correlations in the time series and enhances the persistence of the long-term correlations, leading to a long-memory time series. The fractional differencing parameter  $\alpha$  lies between -0.5 and 0.5 and determines the degree of differencing to be applied to the time series. The properties of model (1) have been widely studied in the literature (e.g., Hosking 1981). We summarize some of these properties below.

- 1. When  $\alpha < 0.5$ , the long-memory process  $y_t$  is said to be weakly stationary. This means that the mean, variance, and autocovariance of the process are constant over time. Additionally, the process has an infinite moving average (MA) representation. In other words, the process can be represented as an infinite sum of past error terms, where the coefficients decay exponentially as the time lag between the error terms increases. The weakly stationary property of  $y_t$  is desirable as it simplifies the analysis and modeling of the process.
- 2. When  $\alpha > -0.5$ , the long-memory process  $y_t$  is invertible. This means that it can be transformed into a stationary process by applying a certain filter. In other words, the invertibility property ensures that the long-memory process can be represented as a finite sum of past error terms, where the coefficients decay exponentially as the time lag between the error terms increases. The invertibility property is useful in practical applications as it allows the analyst to transform the long-memory process into a stationary process, which is easier to analyze and model.
- 3. When  $-0.5 < \alpha < 0.5$ , the autocorrelation function (ACF) of the long-memory process  $y_t$  follows a certain pattern. The ACF decays at a polynomial rate of  $h^{2\alpha-1}$  as the lag h increases, leading to persistent patterns in the data and making the process a long-memory time series. This implies that the memory of the process decays very slowly, and the process can exhibit persistent trends and cycles that extend over long time horizons. The long-memory property of  $y_t$  is important as it captures the long-range dependence in the data, which is often observed in financial, economic, and environmental time series. However, accurately detecting long-memory processes can be challenging, as the sample autocorrelation function may not be a reliable measure of long-range dependence.

If the sample autocorrelation function (ACF) of a time series decays slowly, it can indicate that the series has long memory. In this case, the ACF will not quickly approach zero as the lag increases, but will instead show persistent patterns in the data. However, it is important to keep in mind that other factors, such as the size of the sample or the presence of outliers, may also affect the behavior of the ACF and should be considered when making this determination. Ultimately, statistical tests and model-based approaches are often used to formally test for the presence of long memory in a time series. There are several methods that can be used to estimate the fractional differencing parameter  $\alpha$  in the fractionally differenced model defined by Equation (1).

- 1. Maximum likelihood method: This method involves maximizing the likelihood function of the fractionally differenced model given the observed data, and then using the resulting estimates of the parameters to estimate  $\alpha$ .
- 2. Regression method with logged periodogram at lower frequencies: This method involves regressing the log of the periodogram (a measure of the spectral density of the time series) at lower frequencies on the log of the frequency, and then using the slope of the regression line to estimate  $\alpha$ .

These methods can provide a good starting point in estimating d in practice, although the choice of method will depend on the specifics of the problem at hand. For example, the maximum likelihood method may be preferred when the data are well behaved, while the regression method may be preferred when the data are noisy or contain outliers. Ultimately, the choice of method will also depend on the specific software or package being used to analyze the data.

The sum of the sample autocorrelation function ( $S_{ACF}$ ) has often been used as a diagnostic for long memory, as long-memory processes are characterized by the non-summability of their theoretical autocovariance function. However, this may not always be the case in practice, as the sample ACF may behave differently from the theoretical autocovariance. This is why it is important to consider the implications of using the sample sum of the ACF as a diagnostic for long memory (Hurst 1951).

This study highlights the limitations of relying solely on the sample autocorrelations to diagnose the presence of long memory in a time series. The results demonstrate that

some time series processes can exhibit misleading features in their sample autocorrelations, making it difficult to accurately identify the presence of long memory. These findings have significant implications for the routine use of the sum of the sample autocorrelations in practice and emphasize the need for more robust methods to detect long memory.

The structure of this paper is as follows. In Section 2, we examine the definition of longmemory processes and the sum of the sample autocorrelations. This section also includes a discussion of the characteristics and relevant definitions of long memory. Section 3 focuses on well-known long-memory processes and provides an overview of key theoretical results, along with an examination of the implications of using the sample autocorrelation function. To provide further clarification, an illustrative example is included in Section 3. Section 4 presents a comprehensive comparison of various approaches to long-memory detection. It examines the strengths and limitations of each approach and provides insights into their effectiveness in capturing long-range dependence in time series data. In Section 5, a detailed discussion expands on the findings and implications of the comparative analysis. It delves into the nuances of each approach, highlighting their theoretical foundations, practical considerations, and potential areas for improvement. The discussion critically evaluates the suitability of the examined approaches in real-world scenarios and highlights open research questions and challenges that need to be addressed.

Finally, Section 6 presents the conclusions of the paper. It summarizes the main findings and contributions of the research, emphasizing the significance of the comparative analysis and the implications for long-memory detection in time series data. The section also outlines potential directions for future research, highlighting areas where further advancements are needed.

#### 2. ACF and Long-Term Memory Process

Autocovariance and autocorrelation functions are two fundamental concepts in time series analysis. Autocovariance measures the linear dependence between two observations of a time series at different time lags. Autocorrelation, on the other hand, measures the linear relationship between a time series and a delayed copy of itself. In this subsection, we provide a brief introduction to autocovariance and autocorrelation functions, and we define their mathematical properties.

In practice, we do not have access to the entire population of a time series, but rather to a sample of observations from the series. Therefore, we need to estimate the autocovariance and autocorrelation functions from the available data. In this subsection, we present an estimator for the autocovariance function based on a sample of observations from a stationary time series. We also discuss an alternative estimator and its properties. Finally, we define the autocovariance function and provide an estimator for it based on the estimated autocovariance function.

We also explore the concept of long-term memory processes, which can be defined in various ways. One of the most commonly used definitions is based on the sum of the autocorrelation function, while others rely on the hyperbolic decay of the autocovariances or the power-law decay of the spectral density function. These definitions can provide insights into the behavior of time series data and help to identify long-memory patterns. Additionally, the Wold decomposition of a process can provide an alternative definition, emphasizing the role of past shocks or innovations in influencing the process's behavior over long periods. It is worth noting that these definitions are not necessarily equivalent and can be useful in different contexts. In the following sections, we delve into each of these definitions and their implications for long-term memory processes.

#### 2.1. Sum of the Sample Autocorrelation Function

The autocovariance function of a wide sense stationary process  $\{Y_t\}$  at lag *h* is

$$R(h) = E[(Y_{t+h} - \mu_Y)(Y_t - \mu_Y)], \quad h \in \mathbb{Z}$$

$$\tag{2}$$

where *E* is the expected value operator;  $\mu_Y$  is the expected value of the variable *Y*. In practical problems, we only have a set of data  $Y_T = (y_1, \dots, y_T)$ ; the following estimator can be considered as an estimate of  $\tilde{R}(h)$ :

$$\tilde{R}(h) = \begin{cases} \frac{\sum_{t=1}^{T-|h|} (y_{t+|h|} - \bar{y})(y_t - \bar{y})}{T-|h|}, & h = 0, \pm 1, \cdots, \pm (T-1) \\ 0, & |h| \ge T \end{cases}$$
(3)

where  $\overline{y} = \frac{1}{T} \sum_{t=1}^{T} y_t$  is the sample mean, which is an unbiased estimator of  $\mu$ . There is an alternative estimate of  $\tilde{R}(h)$ :

$$\hat{\gamma}(h) = \begin{cases} \frac{\sum_{t=1}^{T-|h|} (y_{t+|h|} - \overline{y})(y_t - \overline{y})}{T}, & h = 0, \pm 1, \cdots, \pm (T-1) \\ 0 & |h| \ge T \end{cases}$$
(4)

The autocovariance  $\hat{\lambda}(h)$  is biased on the use of the divisor *T* rather than T - |h| and also has larger bias than  $\tilde{R}(h)$ . The autocorrelation function, AFC, is given by

$$\rho(h) = \frac{R(h)}{R(0)} \quad h \in \mathbb{Z}$$
(5)

and an estimate of  $\rho(h)$  is

$$\hat{\rho}(h) = \frac{\hat{\lambda}(h)}{\hat{\lambda}(0)} \quad 0, \pm 1, \cdots, \pm (T-1)$$
(6)

**Theorem 1.** The sum of the sample ACF,  $S_{ACF}$ , with lag  $h \ge 1$  is always  $\frac{-1}{2}$  for any stationary time series with arbitrary length  $T \ge 2$ :

$$S_{ACF} = \sum_{h=1}^{T-1} \hat{\rho}_h = \frac{-1}{2}$$
(7)

Proof. (Hassani 2009)

The  $S_{ACF}$  has the following properties:

- 1. It does not depend on the time series length T;  $S_{ACF} = \frac{-1}{2}$  for  $T \ge 2$ . This property is interesting because it implies that the overall level of autocorrelation in a stationary time series, as measured by the sum of the ACF values, is not affected by the length of the time series. This means that even if we have a very long or a very short time series, the overall degree of temporal dependence in the data remains the same. This property can be useful in comparing the overall level of temporal dependence between different time series of varying lengths.
- 2. The value of  $S_{ACF}$  is equal to  $\frac{-1}{2}$  for any stationary time series. Thus, for example,  $S_{ACF}$  for ARMA(p,q) of any order (p,q) is equal to a Gaussian white noise process and both are equal to  $\frac{-1}{2}$ . The second property of the theorem states that for any stationary time series, the value of  $S_{ACF}$  is always equal to  $\frac{-1}{2}$ . This means that the sum of the sample ACF at each lag is always a constant, regardless of the length of the time series. For example, the sample ACF of an ARMA(p,q) process of any order (p,q) is equal to a Gaussian white noise process, and both have a value of  $\frac{-1}{2}$  for  $S_{ACF}$ . This result has important implications for autoregressive model building and forecasting. If we use the sample ACF to detect the parameters of an autoregressive model, we might yield the improper detection of the order, since the ACF values are not informative of the order.

3. The values of  $\hat{\rho}(h)$  are linearly dependent:

$$\hat{\rho}(i) = \frac{-1}{2} - \sum_{j \neq i=1}^{T-1} \hat{\rho}(i) \quad i = 1, \cdots, T-1$$
(8)

This equation shows that the value of  $\hat{\rho}(i)$  can be expressed as a linear combination of the other sample ACF values, with a constant term of  $\frac{-1}{2}$ . In other words, the ACF values are not independent of each other, but, rather, they are related to each other in a systematic way. This property is a consequence of the fact that the ACF values depend only on the time lag between observations, and not on the specific values of the observations themselves. Therefore, once the ACF values for some lags are known, the values for other lags can be determined using this linear relationship.

4. There is at least one negative  $\hat{\rho}(h)$  for any stationary time series, even for AR(p) with a positive ACF (Hassani 2010).

This property states that for any stationary time series, there is at least one negative sample autocorrelation function (ACF) value, even for autoregressive (AR) models with positive ACF values.

An AR model is a popular class of linear models for time series data, where the value of a variable at time t depends linearly on its own past values, up to a certain number of lagged observations (specified by the model order p). When the AR model is fitted to a stationary time series, the resulting ACF values are typically positive for the first few lags, indicating some degree of autocorrelation in the data. However, this property states that there must always be at least one negative ACF value, even for AR models with positive ACF values.

This property can be understood as follows: although an AR model may capture some of the temporal dependencies in the data, it is unlikely to capture all of them perfectly. In other words, there are likely to be some patterns in the data that are not fully explained by the AR model. These unexplained patterns can lead to negative ACF values, indicating a lack of autocorrelation at certain lags. Therefore, even for stationary time series that exhibit positive autocorrelation overall, there will always be some degree of randomness or unpredictability in the data, resulting in at least one negative ACF value.

The property of  $S_{ACF}$  being constant and equal to  $\frac{-1}{2}$  for any stationary time series has important implications for time series analysis and modeling (see, for example, Hassani and Silva 2015 and Hassani et al. 2021).

#### 2.2. Long-Term Memory Process

The concept of long-memory processes can be defined in different ways. One common definition is based on the sum of the autocorrelation function, as shown in Equation (9),

$$\sum_{h=-\infty}^{\infty} |\rho(h)| = \infty$$
(9)

which states that the sum of the absolute values of the autocorrelation coefficients is infinite. However, there are alternative definitions that can also capture long-memory behavior, such as the hyperbolic decay of the autocovariances, as shown in Equation (10):

$$\rho(h) = h^{2\alpha - 1} l_1(h) \tag{10}$$

In this case, the autocovariances decrease at a rate of  $h^{2\alpha-1}$  as h approaches infinity, where  $\alpha$  is the long-memory parameter and  $l_1(0)$  is a slowly varying function. Another approach to characterizing long-memory processes is through their spectral density function. This definition describes a spectral density function that exhibits a power-law decay for small frequencies, with  $l_2(0)$  being a slowly varying function. These different definitions

can be useful in different contexts and can help to identify long-memory behavior in time series data.

Another definition of strong dependence in the frequency domain is based on the spectral density function  $f(\lambda)$ , which can be expressed as  $|\lambda|^{-2\alpha}l_2(1/|\lambda|)$  for  $\lambda$  near zero, where  $\alpha$  is the long-memory parameter and  $l_2(0)$  is a slowly varying function. This definition highlights the relationship between the behavior of the process and the power of the spectral density function at low frequencies.

Additionally, the Wold decomposition of a process can provide an alternative definition of long-memory behavior. This definition characterizes the process as having a slow decay in its Wold representation, which is a linear combination of past innovations with decreasing weights. This definition emphasizes the role of past shocks or innovations in influencing the behavior of the process over long periods of time:

$$y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} = \psi(B) \varepsilon_t \tag{11}$$

where  $\psi_0 = 1$ ,  $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ ,  $\{\varepsilon_t\}$  is a white noise sequence with variance  $\sigma^2$ .

$$\psi_j \sim j^{\alpha - 1} l_3(j) \tag{12}$$

for j > 0, where  $l_3(0)$  is a slowly varying function. These definitions are not necessarily equivalent; see (Ding et al. 1993; Doukhan et al. 2003).

### 3. Empirical versus Theoretical Results

Let us first briefly consider some widely used long-term memory models.

#### 3.1. Selected Long-Term Memory Models

 GARMA(p,q), which stands for Generalized Autoregressive Moving Average with Conditional Heteroscedasticity, is a type of time series model that combines the features of both the ARMA and GARCH models. This model is suitable for analyzing time series data with a non-constant mean and variance.

The GARMA(p,q) model includes both autoregressive and moving average components as well as a conditional heteroscedasticity term, which captures the time-varying volatility in the data. This allows for better modeling and forecasting of time series data that exhibit changes in volatility over time.

GARMA(p,q) has been applied in various fields for the modeling and forecasting of time series data with changing volatility patterns. For example, it has been used to model stock market returns, exchange rates, and weather data. The GARMA(p,q) model is defined as follows

$$g(\mu_t) = Z'_t \beta = X'_t \beta + \tau_t \tag{13}$$

with

$$\tau_t = \sum_{j=1}^p \phi_j A(y_{t-j}, X_{t-j}, \beta) + \sum_{j=1}^q \theta_j M(y_{t-j}, M_{t-j})$$
(14)

 $\tau_t$ : AR and MA component;

- *A*: the function that represents an autoregressive form;
- *M*: the function that represents a moving average form;
- $\phi_i$ : autoregressive parameter at *j*;
- $\theta_i$ : moving average parameter at *j*.

The above GARMA(p,q), as defined by Equation (13), specifies a linear regression of a function  $g(\mu_t)$  on a set of predictor variables  $X_t$  and a set of unknown parameters  $\beta$ . The error term  $\tau_t$  in Equation (13) is decomposed into an autoregressive

(AR) component and a moving average (MA) component, which are specified in Equation (14).

The function  $A(y_{t-j}, X_{t-j}, \beta)$  in Equation (14) represents the autoregressive component of the model, where  $y_{t-j}$  is the value of the time series at lag j, and  $X_{t-j}$  is the corresponding vector of exogenous variables. The autoregressive parameter at lag j is denoted by  $\phi_j$ , which represents the impact of the lagged value of the time series on the current value, conditional on the values of the exogenous variables.

The function  $M(y_{t-j}, M_{t-j})$  in Equation (14) represents the moving average component of the model, where  $M_{t-j}$  is the set of lagged moving average errors. The moving average parameter at lag *j* is denoted by  $\theta_j$ , which represents the impact of the lagged moving average error on the current value of the time series.

• Integrated GARCH (IGARCH) is a type of time series model that is widely used to model financial and economic data. It is an extension of the GARCH model that accounts for the persistence of shocks in financial markets.

In the IGARCH model, the past conditional variances of the series are included as predictors of the current conditional variance. This allows the model to capture the long-memory effect, where shocks have a persistent effect on future variance.

IGARCH has been applied in various fields for the modeling and forecasting of time series data with persistence in volatility. For example, it has been used to model stock market returns, exchange rates, interest rates, and commodity prices. The IGARCH model is particularly useful for risk management, portfolio optimization, and option pricing. The IGARCH(1,1) model is given by

$$a_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \alpha_0 + \beta_1 \sigma_{t-1}^2 + (1 - \beta_1) a_{t-1}^2, \quad 0 < \beta_1 < 1$$
(15)

ARCH(∞) is a type of time series model that extends the ARCH model to include an infinite number of lags in the conditional variance equation. This allows the model to capture long memory in the volatility of financial and economic time series data.

The ARCH( $\infty$ ) model is based on the idea that past shocks can have a persistent effect on the variance of the series over an infinite time horizon. The model can be estimated using maximum likelihood methods and has been shown to provide a better fit to financial data than finite-order ARCH models.

The ARCH( $\infty$ ) model has been applied in various fields for the modeling and forecasting of time series data with long memory in volatility. For example, it has been used to model stock market returns, exchange rates, and interest rates. The model is particularly useful in finance for risk management, portfolio optimization, and option pricing. However, the estimation of the model can be computationally intensive, and the interpretation of the infinite number of parameters can be challenging. The process { $\varepsilon_t$ } is said to be an ARCH( $\infty$ ), whenever

$$E(\varepsilon_t \mid \psi_{t-1}) = 0 \quad and \quad var(\varepsilon_t \mid \psi_{t-1}) = h_t$$
(16)

with

$$h_t = \alpha_0 + \sum_{i=1}^{\infty} \alpha_i \varepsilon_{t-i}^2 \tag{17}$$

where  $\alpha_i \ge 0$  and  $\psi_t$  represents the information set of all information up to time *t*, i.e.,  $\psi_t = \sigma\{\cdots, \varepsilon_{t-2}, \varepsilon_{t-1}, \varepsilon_t\}$ .

Similarly, IARCH( $\infty$ ) is a type of time series model that extends the integrated GARCH (IGARCH) model to include an infinite number of lags in the conditional variance equation. This model captures the long-memory effect in the volatility of time series data and accounts for the persistence of shocks in financial markets. It should be noted that the estimation of the model can be computationally intensive, and the interpretation of the infinite number of parameters can be challenging.

• The Fractionally Integrated GARCH with a Vector of Conditional Variances (FI-GARCHv) model is a type of time series model that extends the IGARCH model

to allow multiple volatility series to be modeled simultaneously. This allows the model to capture the dynamics and interdependence among multiple time series, such as stock prices, exchange rates, and interest rates.

In the FIGARCHv model, each series has its own IGARCH component and is modeled jointly with other series through a covariance matrix of the conditional variances. This allows the model to capture the spillover effects of volatility shocks across different time series.

The FIGARCHv model has been applied in various fields for the modeling and forecasting of multiple time series data with long memory in volatility. For example, it has been used to model stock prices and exchange rates simultaneously, where the volatility of one series affects the volatility of the other series. The model is particularly useful in finance for risk management, portfolio optimization, and hedging strategies. However, the estimation of the model can be computationally intensive, and the interpretation of the multiple parameters can be challenging. The FIGARCH model is given by

$$[1 - \alpha(L) - \beta(L)](1 - L)^{d}\varepsilon_{t}^{2} = \alpha_{0} + [1 - \beta(L)]\nu_{t}, \quad 0 < d < 1$$
(18)

where

$$(1-L)^{d} = F(-d, 1, 1, L) = \sum_{k=0}^{\infty} \Gamma(k-d)\Gamma(k+1)^{-1}\Gamma(-d)^{-1}L^{k}$$
(19)

 LARCH(∞) and LARCH+(∞) are two types of time series models that extend the ARCH and GARCH models to allow for long memory in the conditional variance equation.

LARCH( $\infty$ ) is an extension of the ARCH model that includes an infinite number of lagged squared residuals in the variance equation. This allows the model to capture long memory in the volatility of time series data. LARCH+( $\infty$ ) is an extension of the GARCH model that includes both an infinite number of lagged squared residuals and an infinite number of lagged conditional variances in the variance equation. This allows the model to capture long memory and the persistence of shocks in financial markets.

Both LARCH( $\infty$ ) and LARCH+( $\infty$ ) have been applied in various fields for the modeling and forecasting of time series data with long memory in volatility. The models are particularly useful in finance for risk management, portfolio optimization, and option pricing. However, the estimation of these models can be computationally intensive, and the interpretation of the infinite number of parameters can be challenging. The LARCH model can be described as

$$r_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \left(a + \sum_{j=1}^\infty b_j r_{t-j}\right)^2 \quad t \in \mathbb{Z}$$
 (20)

where  $\{\varepsilon_t, t \in Z\}$  are iid random variables with zero mean and unit variance.

Stochastic volatility (SV) models are a type of time series model that allow the volatility of financial or economic time series data to vary over time in a random or stochastic manner. These models are based on the idea that the volatility itself is a random process that follows a certain distribution.

In an SV model, the conditional variance of the series is modeled as a function of its past values, as well as a random process that represents the stochastic component of the volatility. This allows the model to capture the time-varying nature of the volatility in the data.

SV models have been widely used in finance and economics for the modeling and forecasting of time series data with changing volatility. For example, they have been used to model stock prices, exchange rates, and interest rates, and are particularly useful for pricing options and other financial derivatives. The models are also used

for risk management and portfolio optimization, as they allow for the more accurate estimation of risk measures such as the value at risk (VaR) and expected shortfall (ES). However, the estimation of SV models can be computationally intensive, and the interpretation of the random component of the volatility can be challenging. The SV model is defined as follows:

$$y_t = \exp(\frac{h_t}{2})\varepsilon_t, \quad \varepsilon_t \sim NID(0,1)$$
 (21)

$$h_{t+1} = \alpha + \beta h_t + \eta_{t+1}, \quad \eta_{t+1} \sim NID(0, \sigma_{\eta}^2)$$
 (22)

$$h_1 \sim N\left(\frac{\alpha}{1-\beta}, \frac{\sigma_\eta^2}{1-\beta^2}\right) \tag{23}$$

where  $\sigma_t^2 = \exp(h_t)$  is the volatility of  $y_t$ . The log volatility  $h_t$  is specified by the AR(1) process with Gaussian innovation noise.

• Autoregressive Fractionally Integrated Moving Average (ARFIMA) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) are two widely used time series models in finance and economics.

ARFIMA models are used to model time series data that exhibit long memory or fractional integration, meaning that the autocorrelation of the series declines very slowly. These models extend the ARIMA models by incorporating fractional differencing, which allows them to capture the long-memory effect.

GARCH models, on the other hand, are used to model time series data that exhibit heteroskedasticity or volatility clustering, meaning that the variance of the series changes over time. These models extend the ARCH models by incorporating autoregressive components in the conditional variance equation, allowing them to capture the persistence of shocks in the data.

Both the ARFIMA and GARCH models have various applications in finance and economics. ARFIMA models are particularly useful in the modeling and forecasting of financial and economic time series with long memory, such as stock prices, exchange rates, and interest rates. GARCH models are widely used in risk management and portfolio optimization, as they allow for the more accurate estimation of risk measures such as the value at risk (VaR) and expected shortfall (ES). They are also used in option pricing and volatility forecasting. However, the estimation of these models can be computationally intensive, and the interpretation of the parameters can be challenging.

An ARFIMA process  $\{y_t\}$  may be defined by

$$\phi(B)y_t = \theta(B)(1-B)^{-\alpha}\varepsilon_t \tag{24}$$

where  $\phi(B) = 1 + \phi_1 B + \cdots + \phi_p B^p$  and  $\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$  are the autoregressive and moving average operators, respectively.  $(1 - B)^{-\alpha}$  is a fractional differencing operator defined by the binomial expansion

$$(1-B)^{-\alpha} = \sum_{j=0}^{\infty} \eta_j B^j = \eta(B)$$
 (25)

where

$$\eta(j) = \frac{\Gamma(j+\alpha)}{\Gamma(j+1)\Gamma(\alpha)}$$
(26)

for  $\alpha < \frac{1}{2}$ ,  $\alpha \neq 0, -1, -2, \cdots$  and  $\{\varepsilon_t\}$  is a white noise sequence with finite variance.

 A CAR(1) model, also known as a Conditional Autoregressive Model of Order 1, is a time series model that describes the dependence between observations in a series over time.

In this model, each observation in the series is assumed to be a function of the previous observation and a random error term. The term "conditional" in CAR(1) refers to the fact that the current observation is conditional on the previous observation.

The application of CAR(1) models is widely practised in econometrics, finance, and engineering for the forecasting and analysis of time series data. It is particularly useful in modeling and forecasting stock prices, exchange rates, and interest rates. It can also be used in modeling natural phenomena, such as climate patterns or population growth.

The CAR model explains the observations with *p* fixed effects and *n* spatial random effects:

$$\boldsymbol{y} \mid \boldsymbol{\beta}, \boldsymbol{s}, \boldsymbol{\tau_0} \sim N(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{I}_n \boldsymbol{S}, \boldsymbol{\tau_0} \boldsymbol{I}_n)$$
(27)

$$\boldsymbol{s} \mid \tau_j \sim N(\boldsymbol{0}, Q), \quad j = 1, \cdots, F-1$$
 (28)

where  $\tau_0 I_n$  and  $Q = \sum_{j=1}^{F-1} \tau_j Q_j$  are precision matrices, observations y and random effects s are  $n \times 1$ , design matrix X is  $n \times p$ , and the fixed effect regression parameter vector  $\beta$  is  $p \times 1$ .

#### 3.2. Results

Table 1 presents a comparison of theoretical and empirical results for the definition of long memory based on the sum of the sample autocorrelation function. Table 1 lists different financial models for long-memory time series data. Each model has a theoretical expectation for the long-memory process, represented by either  $\sum_{h=-\infty}^{\infty} |\rho(h)| \to \infty$  as  $T \to \infty$  or  $f(\lambda) \to \infty$  as  $\lambda \to 0$ . The table also lists the empirical results for each model, represented by  $\sum_{h=-(T-1)}^{(T-1)} \hat{\rho}(h) = 0$  or  $\hat{f}(0) = 0$ .

It is evident from the table that while the theoretical results based on the ACF or spectral density are infinite, the empirical spectral density or sum of the empirical ACF is finite and zero. This significant discrepancy between the two makes the detection of long memory misleading. The results presented in this table are, for example, time series that have been widely used in the literature (Hassani et al. 2012).

Table 2 presents the comparison between the theoretical and empirical results of longmemory process detection. The theoretical results and the empirical results are presented in the two columns of the table. The first row shows that, according to the theoretical results, the sum of the absolute values of the autocorrelation coefficients (i.e.,  $\rho(h)$ ) approaches infinity as the number of observations (*T*) increases. However, the empirical results show that the sum of the absolute values of the estimated autocorrelation coefficients (i.e.,  $\hat{\rho}(h)$ ) has a finite upper limit as *T* approaches infinity.

The second row presents a similar discrepancy between the theoretical and empirical results, with the theoretical results indicating that the sum of the autocorrelation coefficients approaches infinity as *T* approaches infinity, while the empirical results show that the sum of the estimated autocorrelation coefficients is equal to zero.

The third row compares the behavior of the theoretical and empirical spectral densities. The theoretical results show that the spectral density (i.e.,  $f(\lambda)$ ) approaches infinity as the frequency  $(\lambda)$  approaches zero, while the empirical results show that the estimated spectral density (i.e.,  $\hat{f}(\lambda)$ ) is equal to zero at zero frequency. Overall, the table shows that there are discrepancies between the theoretical and empirical results of long-memory process detection, indicating that the assumptions made in the theoretical analysis may not hold in practice.

Case	Models	Long-Memory Process	<b>Empirical Results</b>	References
1	GARMA(p,q)	$\sum_{h=-\infty}^{\infty} \rho(h) \to \infty$ as $T \to \infty$	$\sum_{h=-(T-1)}^{(T-1)} \rho(h) = 0$	(Bertelli and Caporin 2002)
2	IGARCH	$f(\lambda) \to \infty$ as $\lambda \to 0$	$\widehat{f}(0)=0$	(Rapach et al. 2008)
3	$ARCH(\infty)$	$f(\lambda) \to \infty$ as $\lambda \to 0$	$\widehat{f}(0)=0$	(Bertail et al. 2006)
4	$IARCH(\infty)$	$f(\lambda) \to \infty$ as $\lambda \to 0$	$\hat{f}(0) = 0$	(Teyssière and Kirman 2002)
5	FIGARCHv	$f(\lambda) \to \infty$ as $\lambda \to 0$	$\widehat{f}(0)=0$	(Zivot and Wang 2013)
6	$LARCH(\infty)$	$f(\lambda) \to \infty$ as $\lambda \to 0$	$\hat{f}(0) = 0$	(Beran et al. 2013)
7	$LARCH_{+}(\infty)$	$f(\lambda) \to \infty$ as $\lambda \to 0$	$\hat{f}(0) = 0$	(Dedecker et al. 2007)
8	SV models (Stochastic Volatility)	$\sum_{h=-\infty}^{\infty} \rho(h) \to \infty$ as $T \to \infty$	$\sum_{h=-(T-1)}^{(T-1)} \rho(h) = 0$	(Zheng et al. 2018)
9	ARFIMA – GARCH	$f(\lambda) \to \infty$ as $\lambda \to 0$	$\widehat{f}(0)=0$	(Barbieri 2017)
10	CAR(1)	$\sum_{h=-\infty}^{\infty} \rho(h) \to \infty$ as $T \to \infty$	$\sum_{h=-(T-1)}^{(T-1)} \rho(h) = 0$	(Beran et al. 2013)

Table 1. Some examples of long-memory time series—theoretical vs. empirical results.

Table 2. The theoretical and empirical results of a long-memory process.

Theoretical Results	Empirical Results
$\sum_{h=-\infty}^{\infty}  \rho(h)  \to \infty$ as $T \to \infty$	$2 \leq \lim_{T  o \infty} \sum_{h=-(T-1)}^{(T-1)} \mid \hat{ ho}(h) \mid < \infty$
$\sum_{h=-\infty}^{\infty} \rho(h) \to \infty$ as $T \to \infty$	$\sum_{h=-(T-1)}^{(T-1)} ho(h)=0$
$f(\lambda)  o \infty$ as $\lambda  o 0$	$\hat{f}(0)=0$

Table 3 provides a comparison of four popular time series models, namely ARFIMA, GARMA, IGARCH, and CAR(1), based on four important characteristics of time series data: long memory, stationarity, volatility clustering, and the autocorrelation function. ARFIMA and GARMA are both long-memory models, meaning that they can capture the long-range dependence present in time series data. However, neither of these models guarantees stationarity, which is a desirable property in many applications. On the other hand, IGARCH and CAR(1) are both stationary models, but they do not capture long-memory dependence. IGARCH is designed specifically to model volatility clustering, which is a common phenomenon in financial time series data. In contrast, CAR(1) assumes that autocorrelation decreases exponentially with the lag and does not account for volatility clustering.

Table 3 also shows that the autocorrelation functions of ARFIMA, GARMA, IGARCH, and CAR(1) all decrease over time, but with different rates. For ARFIMA and GARMA, the autocorrelation function decreases to zero, which is indicative of the long-memory dependence captured by these models. In contrast, the autocorrelation functions of IGARCH and CAR(1) decrease exponentially, reflecting the short-range dependence present in these models. Understanding the properties of different time series models can help researchers and practitioners to choose the most appropriate model for their specific application and improve the accuracy of their forecasts.

Let us now consider the differences between the theoretical and empirical results of these models. Figure 1 illustrates these differences. Figure 1 presents 1000 realizations from the ARFIMA, GARMA, IGARCH, and CAR(1) processes. These models shows the behavior of the empirical sum of the autocorrelation function (see package "Hassani.SACF" in R). Figure 1 shows the sum of the sample autocorrelation function (ACF) for various long-memory models, including ARFIMA, GARMA, IGARCH, and CAR(1). These models are widely used in practice to capture long-range dependence in time series data.

Model	Long Memory	Stationarity	Volatility Clustering	Autocorrelation Function
ARFIMA	Yes	No	No	Decreases to zero
GARMA	Yes	No	No	Decreases to zero
IGARCH	No	Yes	Yes	Decreases exponentially
CAR(1)	No	Yes	No	Decreases exponentially

Table 3. Comparison of ARFIMA, GARMA, IGARCH, and CAR(1).

While the patterns of the sum of the sample ACF are different for each model, they all ultimately converge to  $\frac{-1}{2}$  as the sample size increases. This means that the sum of the sample ACF cannot be used as an accurate measure of long memory in the way that the theoretical definition of long memory is based on the ACF. This finding is significant because it suggests that relying solely on the sample ACF to identify long-memory processes can be misleading.



Figure 1. SACF for "ARFIMA", "GARMA", "IGARCH", and "CAR(1)" processes.

# 4. Comparison of Parametric and Non-Parametric/Semi-Parametric Approaches for Long-Term Memory Time Series Detection

There are several approaches available for the detection of long-term memory in time series data. Broadly speaking, these approaches can be classified into two categories: parametric and non-parametric/semi-parametric.

Parametric approaches assume a specific functional form for the underlying process and estimate its parameters using a maximum likelihood method. These approaches are typically used when the underlying process can be well approximated by a known stochastic process, such as ARIMA, ARFIMA, or GARCH. Parametric approaches have the advantage of being computationally efficient and providing explicit statistical inference, such as hypothesis testing and confidence intervals. However, they may not be appropriate when the underlying process does not follow the assumed functional form, or when the data are contaminated with outliers or measurement errors.

Non-parametric/semi-parametric approaches, on the other hand, do not assume a specific functional form for the underlying process and estimate its properties using more flexible methods. These approaches include wavelet-based methods, periodogrambased methods, detrended fluctuation analysis (DFA), and local Whittle estimation. Nonparametric/semi-parametric approaches have the advantage of being more robust to deviations from the underlying assumptions and can capture more complex dependence structures. However, they may require more computational resources and provide less explicit statistical inference than parametric approaches. Table 4 summarizes the strengths and weaknesses of parametric and non-parametric/semi-parametric approaches for long-term memory time series detection.

**Table 4.** Strengths and weaknesses of parametric and non-parametric/semi-parametric approaches for long-term memory time series detection.

Methods	Parametric	Semi-Parametric	Non-Parametric
Basic assumption	Assumes specific probability dis- tribution of errors	Allows for flexible modeling of errors	Makes no assumptions about the probability distribution of errors
Strengths	Accurate modeling of errors and autocorrelation function	Combines flexibility and accuracy	Robust to deviations from distribu- tion assumptions
Weaknesses	Sensitive to distributional as- sumptions	May not fully capture complex de- pendencies	May be less accurate in small sam- ple sizes
Examples	ARIMA, ARCH/GARCH	Fractional ARIMA, Wavelet analy- sis	Hurst exponent, rescaled range analysis

It is worth noting that both parametric and non-parametric/semi-parametric approaches have their own strengths and weaknesses. The choice of approach should depend on the specific characteristics of the data and the research question at hand. Researchers and practitioners should carefully evaluate the assumptions and limitations of each approach and select the most appropriate one for their application.

#### 5. Discussion

The empirical results presented in Table 3 provide a comparison of four popular time series models: ARFIMA, GARMA, IGARCH, and CAR(1). These models were evaluated based on four important characteristics of time series data: long memory, stationarity, volatility clustering, and the autocorrelation function.

ARFIMA and GARMA were found to be long-memory models capable of capturing long-range dependence in time series data. However, neither of these models guarantees stationarity, which is a desirable property in many applications. On the other hand, IGARCH and CAR(1) were identified as stationary models, but they do not capture long-memory dependence. IGARCH is specifically designed to model volatility clustering, a common phenomenon observed in financial time series data. In contrast, CAR(1) assumes a decreasing exponential autocorrelation function and does not account for volatility clustering. The theoretical and empirical results of these models were compared, as illustrated in Figure 1. It was observed that the patterns of the sum of the sample ACF differed for each model. However, all the models ultimately converged to approximately -1/2 as the sample size increased. This implies that relying solely on the sum of the sample ACF as a measure of long memory can be misleading. It deviates from the theoretical definition of long memory, which is based on the ACF.

This finding has important implications as it highlights the limitation of using the sample ACF alone to identify long-memory processes. Researchers and practitioners should exercise caution when interpreting results solely based on the sum of the sample ACF, as it may not accurately capture the presence of long-range dependence.

The results obtained from the empirical analysis raise questions about the generalizability of these findings to other approaches and methods that rely on the sample autocorrelation function (ACF) for the detection of long-memory processes. The answer to this question is affirmative, as demonstrated by Hassani's -1/2 theorem. This theorem suggests that methods relying solely on the sample ACF may fail to accurately detect long-memory processes, indicating a limitation in their theoretical foundation.

It is important to note that many existing approaches and methods to identify longmemory processes are based on asymptotic behaviors and various assumptions. However, the empirical results presented in this study demonstrate that these assumptions may not hold in real-world scenarios, as evident from the examples provided. This suggests that relying solely on such approaches can lead to inaccurate conclusions and may not fully capture the true nature of long memory in time series data.

To address these limitations, a data-driven approach rather than a model-based approach could be a potential solution. By adopting a data-driven approach, researchers can explore the inherent patterns and structures in the data themselves, rather than relying solely on predefined models or assumptions. This approach acknowledges the complex and diverse nature of real-world data and allows for a more comprehensive investigation of long-memory processes.

Further investigation is warranted to explore and develop data-driven methods to detect long memory in time series data. Such investigations could involve the development of innovative techniques that leverage machine learning algorithms, advanced statistical methodologies, or nonparametric approaches. By incorporating the richness and complexity of the data into the analysis, these data-driven methods have the potential to provide the more accurate and robust identification of long-memory processes.

#### 6. Concluding Remarks

The paper discusses the issue of detecting long-range dependence in time series data and the discrepancies between the theoretical definition and empirical identification of long-memory processes. The theoretical definition of long memory is based on the autocorrelation function (ACF), which measures the correlation between observations at different time lags.

However, the paper highlights that the commonly used empirical measure of long memory, the sample autocorrelation sum, is a predetermined constant for any stationary time series, regardless of sample size. This means that it cannot identify long-memory processes in the same way as the theoretical definition.

The implications of this are significant, as it suggests that theoretical results based on the ACF may be misleading if the empirical identification of long memory is incorrect. The paper presents an analysis of various long-memory models to demonstrate this point.

The main conclusion of the paper is that alternative approaches to identifying longmemory processes are necessary, given the limitations of the sample autocorrelation sum. Researchers and practitioners who use long-memory models should consider alternative methods of detecting long-range dependence.

For further research, the authors plan to investigate alternative measures and approaches for the detection of long-range dependence in time series data. Specifically, we aim to explore the use of alternative approaches such as wavelet-based methods and fractional integration techniques, which have shown promise in identifying long-memory processes. By using these alternative approaches, we hope to develop more reliable and accurate methods for the detection of long-range dependence in time series data, which could have important implications for model building and forecasting in various fields.

**Author Contributions:** Conceptualization, H.H., M.Y. and L.M.M.; Methodology, H.H., M.Y. and L.M.M.; Writing—review and editing, H.H., M.Y. and L.M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data that support the findings of this study are available on request.

Conflicts of Interest: The authors declare no conflict of interest.

#### References

- Barbieri, Riccardo, Enzo Pasquale Scilingo, and Gaetano Valenza. 2017. *Complexity and Nonlinearity in Cardiovascular Signals*. New York: Springer. [CrossRef]
- Beran, Jan. 1994. *Statistics for Long-Memory Processes*. London: Chapman & Hall, p. 315. Available online: https://www.routledge.com/ Statistics-for-Long-Memory-Processes/Beran/p/book/9780412049019 (accessed on 30 March 2023).
- Beran, Jan, Yuanhua Feng, Sucharita Ghosh, and Rafal Kulik. 2013. *Long-Memory Processes: Probabilistic, statistical, and Economic Issues*. Berlin and Heidelberg: Springer Science & Business Media.

- Bertelli, Stefano, and Massimiliano Caporin. 2002. A note on calculating auto-covariances of long-memory processes. *Journal of Time Series Analysis* 23: 503–8. [CrossRef]
- Bertail, Patrice, Paul. Doukhan, and Philippe Soulier. 2006. Dependence in Probability and Statistics. In *Springer Lecture Notes on Statistics* 187. New York: Springer. [CrossRef]
- Das, K.J., and Sharad Nath Bhattacharya. 2021. Detection of Long Memory in the Indian Stock Market using Fractional Integration Analysis. *Business Studies* 1–2: 28. Available online: https://www.caluniv.ac.in/dj/BS-Journal/vol-31-32/long\_memory.pdf (accessed on 7 June 2023).
- Dedecker, Jérôme, Paul Doukhan, Gabriel Lang, León R. José Rafael, Sana Louhichi, and Clémentine Prieur. 2007. Weak dependence: In *Weak Dependence: With Examples and Applications*. Lecture Notes in Statistics. New York: Springer, p. 190. [CrossRef]
- Ding, Zhuanxin, Clive W. J. Granger, and Robert F. Engle. 1993. A long memory property of stock market returns and a new model. *Journal of Empirical Finance* 1: 83–106. [CrossRef]
- Doukhan, Paul, George Oppenheim, and Murad Taqqu. 2003. Theory and Applications of Long-Range Dependence. Boston: Birkhäuser.
- Hassani, Hossein. 2009. Sum of the sample autocorrelation function. *Random Operators and Stochastic Equations* 17: 125–30. [CrossRef] Hassani, Hossein. 2010. A note on the sum of the sample autocorrelation function. *Physica A* 389: 1601–6. [CrossRef]
- Hassani, Hossein, Leonenko Nikolai, and Kerry Patterson. 2012. The sample autocorrelation function and the detection of long-memory processes. *Physica A: Statistical Mechanics and Its Applications* 391: 6367–79. [CrossRef]
- Hassani, Hossein, and Emmanuel Sirimal Silva. 2015. A Kolmogorov-Smirnov Based Test for Comparing the Predictive Accuracy of Two Sets of Forecasts. *Econometrics* 3: 590–609. [CrossRef]
- Hassani, Hossein, Mohammad Reza Yeganegi, Sedigheh Zamani Mehreyan, and Abdolreza Sayyareh. 2021. On the Sample Autocorrelation Functions Absolute Summability. *Fluctuation and Noise Letters* 21: 2250004. [CrossRef]
- Hosking, J. R. M. 1981. Fractional differencing. Biometrika 8: 165–76. [CrossRef]
- Hurst, Harold Edwin. 1951. Long term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers* 116: 770–99. [CrossRef]
- Rapach, David E., Mark E. Wohar, and Hamid Beladi. 2008. *Forecasting in the Presence of Structural Breaks and Model Uncertainty*. Amsterdam: Elsevier Science Ltd., vol. 3, 700p. Available online: https://books.emeraldinsight.com/page/detail/Forecasting-in-the-Presence-of-Structural-Breaks-and-Model-Uncertainty/?k=9780444529428 (accessed on 30 March 2023).
- Tsay, Ruey. S. 2010. Analysis of Financial Time Series. New York: John Wiley & Sons.
- Teyssière, Gilles, and Alan Kirman. 2002. Microeconomic models for long memory in the volatility of financial time series. *Physica A* 370: 26–31.
- Wei, William. W. S. 2006. Time Series Analysis Univariate and Multivariate Methods, 2nd ed. New York: Addison Wesley, vol. 29, pp. 33–59.
- Zivot, Eric, and Jiahui Wang. 2013. Modeling Financial Time Series with S-PLUS®, 2nd ed. Berlin and Heidelberg: Springer.
- Zheng, Min, Liu Ruipeng, and Li Youwei. 2018. Long memory in financial markets: A heterogeneous agent model perspective International Review of Financial Analysis 58: 38–51. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.