Article

Unraveling Time Series Dynamics: Evaluating Partial Autocorrelation Function Distribution and Its Implications

Hossein Hassani [1,*], Leila Marvian [2], Masoud Yarmohammadi [2] and Mohammad Reza Yeganegi [3]

1   The Research Institute of Energy Management and Planning (RIEMP), University of Tehran, Tehran 19395-4697, Iran
2   Department of Statistics, Payame Noor University, Tehran 19395-4697, Iran; leila.marveian@imamreza.ac.ir (L.M.); masyar@pnu.ac.ir (M.Y.)
3   The International Institute for Applied Systems Analysis (IIASA), 2361 Laxenburg, Austria; yeganegi@iiasa.ac.at
*   Correspondence: hassani.stat@gmail.com

Abstract: The objective of this paper is to assess the distribution of the Partial Autocorrelation Function (PACF), both theoretically and empirically, emphasizing its crucial role in modeling and forecasting time series data. Additionally, it evaluates the deviation of the sum of sample PACF from normality: identifying the lag at which departure occurs. Our investigation reveals that the sum of the sample PACF, and consequently its components, diverges from the expected normal distribution beyond a certain lag. This observation challenges conventional assumptions in time series modeling and forecasting, indicating a necessity for reassessment of existing methodologies. Through our analysis, we illustrate the practical implications of our findings using real-world scenarios, highlighting their significance in unraveling complex data patterns. This study delves into 185 years of monthly Bank of England Rate data, utilizing this extensive dataset to conduct an empirical analysis. Furthermore, our research paves the way for future exploration, offering insights into the complexities and potential revisions in time series analysis, modeling, and forecasting.

Keywords: autocorrelation function (ACF); partial autocorrelation function (PACF); time series analysis; normality; PACF distribution; Hassani's $-\frac{1}{2}$ theorem; Bank of England rate

1. Introduction

Time series data, characterized by its sequential nature, underpins a vast array of applications, from economics [1] and finance [2] to climate modeling [3] and signal processing [4]. Thus, understanding and extracting meaningful insights from the data hinges on our ability to grapple with its inherent temporal dependencies. The Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) emerge as central pillars in this endeavor [5–7].

The power of ACF and PACF do not rest solely in their application; their theoretical and empirical distribution form the cornerstone of statistical inference in time series analysis. Therefore, understanding the expected distribution of ACF and PACF values under varying conditions serves as a compass guiding analysts through hypothesis testing, confidence interval estimation, model diagnostics, and forecasting [8].

For the time series analysis, a fundamental principle has long been established: the theoretical sum of ACF values is invariably zero for any white noise process characterized by a mean of zero and an arbitrary value of variance, denoted as σ^2. However, a compelling and intriguing discovery emerges when we delve into empirical observations: the sum of sample ACF consistently converges to a constant, specifically $-\frac{1}{2}$ (Hassani's $-\frac{1}{2}$ Theorem) [6,7]. This intriguing constancy persists regardless of diverse time series attributes, such as length, mean, variance, dependencies, or any other unique characteristics they may possess. Numerous research studies have underscored the significance of Hassani's $-1/2$

Theorem for practical applications and its integration into time series analysis and model development (refer, for example, to [9–12]).

The implications of this remarkable consistency are profound, particularly for the fields of time series model building and analysis [13–15]. For a selected recent work highlighting the importance of considering the sample ACF in analysis, especially in the context of Hassani's $-1/2$ Theorem, see, for instance, [16–18].

However, the sum of sample PACF has received relatively little attention in comparison to its counterpart, ACF. Therefore, this paper will primarily focus on exploring the PACF. It is essential to note that there exists a robust relationship between ACF and PACF for time series analysis and modeling.

This paper goes beyond theoretical conjectures, illustrating our findings with a rich array of practical examples. We employ both simulated data and real-world time series to validate and emphasize the practical relevance of the results obtained in our study. Through this interdisciplinary journey, we aim to contribute valuable insights that enhance the understanding and application of PACF in the field of time series analysis.

In Section 2, we provide a foundational understanding of PACF, its distribution, and the sum of sample ACF as the backbone of this research. Within this section, the definitions and mathematical derivations of these concepts are elucidated, offering readers a view into the mechanics of the PACF. The distribution of sample PACF subsection advances the discourse by focusing on the distribution aspects of the sample PACF. This subsection paints a detailed portrait of the PACF's behavior, showcasing its inherent properties and tendencies.

In Section 3, the paper shifts its focus to assess the behavior of the sample PACF and the normality of the sum of sample PACF. Through simulation studies, this section juxtaposes theoretical PACF with sample PACF, highlighting potential disparities and deviations. Building on prior findings, the section offers a fresh perspective on the distribution of sample PACF, underscored by empirical evidence. Additionally, this section evaluates the application of observations from the simulations using real data sets, specifically the 185 years of monthly Bank of England Rate data.

Section 4 summarizes the main themes and findings of the paper, bringing the discussion to a close. By synthesizing the concepts and reiterating the key discoveries, this section provides a cohesive wrap-up of the paper's central arguments and insights. Additionally, it offers several ideas for further research based on the findings of this paper.

2. Partial Autocorrelation Function

In order to establish the definition of the PACF, it is necessary to initially examine the notion of a linear regression, wherein a random variable Y_t is predicted using its own previous values up to a certain lag h [5].

The partial autocorrelation coefficient, denoted as ϕ_{hh}, represents the correlation between the time series variable Y_t and its lagged value Y_{t-h}, while controlling for the influence of all other preceding lags. The aforementioned number is derived by the resolution of the Yule–Walker equations pertaining to an autoregressive process with a lag of h. By utilizing empirical data, it is possible to derive an estimation for the PACF, conventionally represented as $\hat{\phi}_{hh}$. The formula for estimating the PACF, $\hat{\phi}_{hh}$, for a given realization of length T, $Y_T = (y_1, \cdots, y_T)$, is expressed as the estimation of coefficients in the linear regression of Y_T on $Y_{t-1}, Y_{t-2}, \cdots, Y_{t-h}$.

The estimation of these coefficients can be achieved by several ways, one of which is the Durbin–Levinson algorithm [19]. This algorithm is a recursive technique employed to solve the Yule–Walker equations specifically for autoregressive (AR) models. It is important to acknowledge that in order to obtain a precise estimation of the PACF, it is necessary to estimate the ACF for the given data. The utilization of PACF estimates, generated from sample PACF, is commonly employed to ascertain the appropriate number of lags to include in an autoregressive (AR) model [5,20].

2.1. Distribution of PACF

Let us now start this subsection by evaluating the distribution of PACF. Understanding the distribution of PACF is vital for any time series modeling and forecasting, as it provides crucial insights into the underlying dependencies within the data. By examining the distribution, we can better assess the behavior of the time series, identify patterns, and improve the accuracy of our models. This foundational knowledge is essential for developing robust forecasting techniques and ensuring the reliability of predictions. Let us first concentrate on the $AR(p)$ process.

Remark 1. *For a causal $AR(p)$ process, asymptotically $(n \rightarrow \infty)$, $\sqrt{(n)}\hat{\phi}_{hh} \rightarrow N(0, 1)$ for $h > p$ [19].*

The above remark implies that, for large sample sizes, the estimated partial autocorrelation coefficients beyond the order p follow a standard normal distribution, providing a basis for statistical inference in time series analysis. Consequently, the asymptotic normality of the sample PACF suggests that the sum of the sample PACF would also be asymptotically normal. This result is derived by applying the delta method to the sample PACF at different lags (see [19] for details on the delta method). Please note that the $AR(p)$ process, when $p = 0$, is a white noise series. Thus, this remark also applies to a white noise series. Furthermore, under the null hypothesis for statistical tests, we assume that we have a noise time series. Thus, the normality assumption plays a vital role.

The above remark indicates that, for an $AR(p)$ process, $\sum \hat{\phi}_{hh}$ for $h > p$ follow a normal distribution. Our aim here is to study this sum for various lags, even for $h < p$. While there is not enough research on $\sum \hat{\phi}_{hh}$, recent studies on the sum of sample ACF reveal several important observations for time series analysis and forecasting [6,7,21–25].

As mentioned in Remark 1, the sample PACF of any white noise process (as $AR(0)$ process) is asymptotically normal in all lags. In testing the null hypothesis $H_0 : \phi_{hh} = 0, h = 1, \cdots, T - 1$, this characteristic of sample PACF of white noise comes in handy, since under this null hypothesis, the time series will be a white noise process.

As mentioned above, in practical applications, the accurate values of the PACF, ϕ_{hh}, are often unknown and are approximated using their sample equivalents, $\hat{\phi}_{hh}$. Thus, in practical terms, our main interest lies with the sample data at hand. It is critical to underscore that the theoretical notions, inclusive of all peripheral theory and approximations, should ideally be observable in the application of these methods to real-world data. This practical application is pivotal, as it allows us to validate the theoretical models and ensure their relevance and utility in empirical analysis.

2.2. Sum of the Sample ACF

Let us now examine the sum of the Sample ACF (SACF). We will explore various properties of the SACF, largely building upon previous research, as exemplified in [6,7,21–25].

Hassani $-1/2$ Theorem [6]

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

$$S_{ACF} = \sum_{h=1}^{T-1} \hat{\rho}_h = \frac{-1}{2}. \tag{1}$$

The S_{ACF} has the following properties:

1. It remains independent of the time series length, denoted as T; specifically, S_{ACF} equals $\frac{-1}{2}$ for $T \geq 2$.
   This property is intriguing, because it implies that the total level of autocorrelation in a stationary time series, quantified by the sum of the ACF values, remains unaffected by the time series' length.

2.   The value of S_{ACF} remains constant at $\frac{-1}{2}$ for any stationary time series. Consequently, for instance, the S_{ACF} for any order (p, q) of an $ARMA(p, q)$ process is equivalent to that of a Gaussian white noise process, with both being equal to $\frac{-1}{2}$.

The second property of the theorem asserts that, for any stationary time series, the value of S_{ACF} is consistently $\frac{-1}{2}$. In practical terms, this signifies that the sum of sample ACF at each lag remains constant, irrespective of the time series' length. For instance, whether it is the sample ACF of an $ARMA(p, q)$ process of any order (p, q) or a Gaussian white noise process, both exhibit an unchanging value of $\frac{-1}{2}$ for S_{ACF}.

This result holds significant implications for the process of constructing ARMA models and conducting forecasting. Consequently, improper order detection may occur, leading to potential modeling inaccuracies.

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

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

This equation demonstrates 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 essence, it reveals that the sample ACF values are not independent of one another; instead, they are systematically interconnected.

4.   There is at least one negative $\hat{\rho}(h)$ for any stationary time series, even for $AR(p)$ with a positive ACF [6].

This property asserts that, in the context of any stationary time series, there is guaranteed to be at least one negative sample ACF value, even when dealing with autoregressive (AR) models characterized by positive ACF values.

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, [24,25]).

In conclusion, while the theoretical sum of ACF is always zero for any white noise process with a mean of zero and variance σ^2 (for example), it is important to note that the sum of sample ACF consistently remains at $\frac{-1}{2}$. This holds true regardless of the time series length, mean, variance values, dependencies, or any other time series characteristics.

Although the patterns of the sum of the sample ACF vary across models, the sum always converges to $\frac{-1}{2}$ as the sample size expands. This indicates that the sum of the sample ACF may not be a reliable metric for discerning long memory, especially when compared to the theoretical definition rooted in the ACF [25]. This insight is crucial, highlighting that exclusive reliance on the sample ACF for identifying long-memory processes might lead to erroneous conclusions.

3. Assessing Normality of Sum of Sample PACF

Let us now consider the behavior of the sample PACF, examining it both through simulation studies and by using real data sets. By leveraging simulations, we can explore the theoretical properties and potential anomalies of the sample PACF in a controlled environment. Additionally, applying our findings to real-world data, such as the 185 years of monthly Bank of England Rate, allows us to validate our theoretical insights and assess their practical implications. This dual approach not only enhances our understanding of the sample PACF but also strengthens the robustness of our conclusions in the context of time series modeling and forecasting.

3.1. Simulation

In order to investigate the distribution of the sample PACF, a simulation study is conducted. The simulation consists of 5000 sample paths for each of the time series lengths $T = 50, 100, 500,$ and 1000. Each sample path is generated from Gaussian white noise, and the sample PACF is calculated for all possible lags (lags 1 to $T - 1$ for the generated

sample path of length *T*). Once the sample PACFs are generated for simulated sample paths, the Shapiro–Wilk test [26] is employed to test the normality of the cumulative sum of sample PACF in each lag. The null hypothesis in the normality test is as follows:

$$H_{0_{(h)}} : \sum_{i=1}^{h} \hat{\phi}_{ii} \sim N(\mu_{0_{(h)}}, \sigma_{0_{(h)}}), \quad h = 1, \cdots, T - 1,$$

where $\mu_{0_{(h)}}$ and $\sigma_{0_{(h)}}$ are the mean and standard deviation of $\sum_{i=1}^{h} \hat{\phi}_{ii}$ in the simulated data. The *p*-values of each test and the first lag in which the normality assumption is rejected are recorded. In order to account for the stochastic nature of the simulation study, the above-described simulation is repeated 100 times, and the median, 2.5, and 97.5 percentiles for *p*-values and the first lag in which the normality assumption is rejected are calculated. Furthermore, the density of the sum of sample PACF is estimated using the kernel method, and the median and 2.5 and 97.5 percentiles of the fitted densities are calculated as well. Estimated percentiles for the first lags in which the sum of sample PACF's normality is rejected are presented in Table 1.

Table 1. First lag in which Shpiro–Wilk null hypothesis is rejected in simulated white noise sample paths.

Percentiles of the First, Lag in Which Normality Is Rejected	Time Series Length			
	T = 50	*T* = 100	*T* = 500	*T* = 1000
2.5%	1	1	1	3
Median	9	18	36	70
97.5%	14	27	131	261

Figure 1 also shows the distribution of the first lags where the sum of the sample PACF's distribution significantly departs from a normal distribution. As can be seen, after a certain number of lags, in all the simulations, the distribution of the sum of the sample PACF departs from a normal distribution.
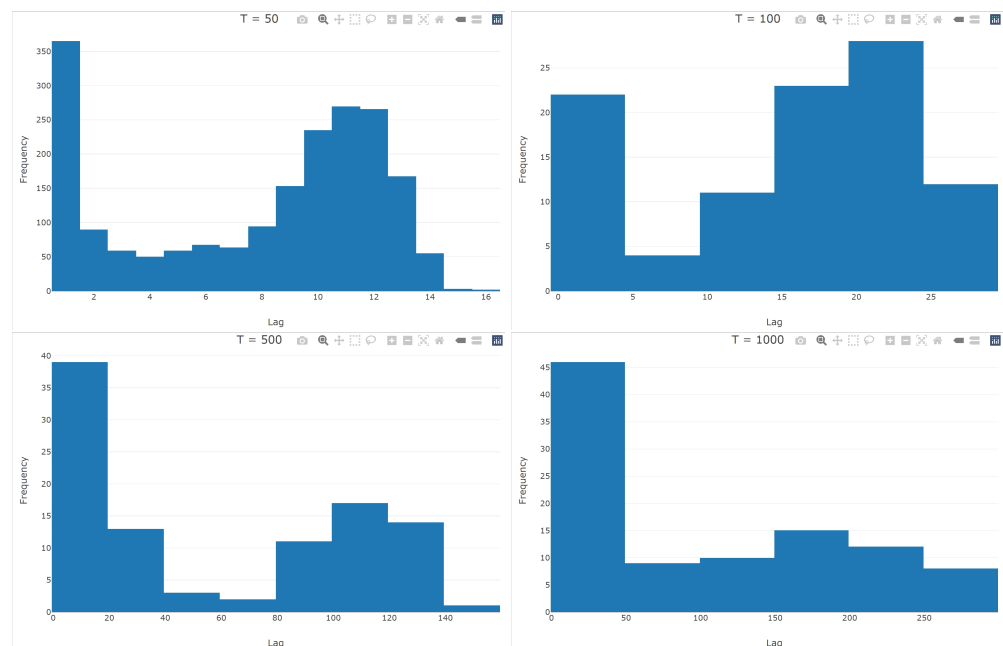

Figure 1. Histogram of first lag in which the normality of white noise's cumulative sum of sample PACF is rejected at $\alpha = 5\%$ significance level.

Figure 2 represents the trend of Shapiro–Wilk's *p*-value for testing the normality of cumulative sum of sample PACFs in different lags. The solid line shows the median of

p-values from different simulations and the bound shows the area between 2.5 and 97.5 percentiles of *p*-values from different simulations (representing 95% confidence interval).
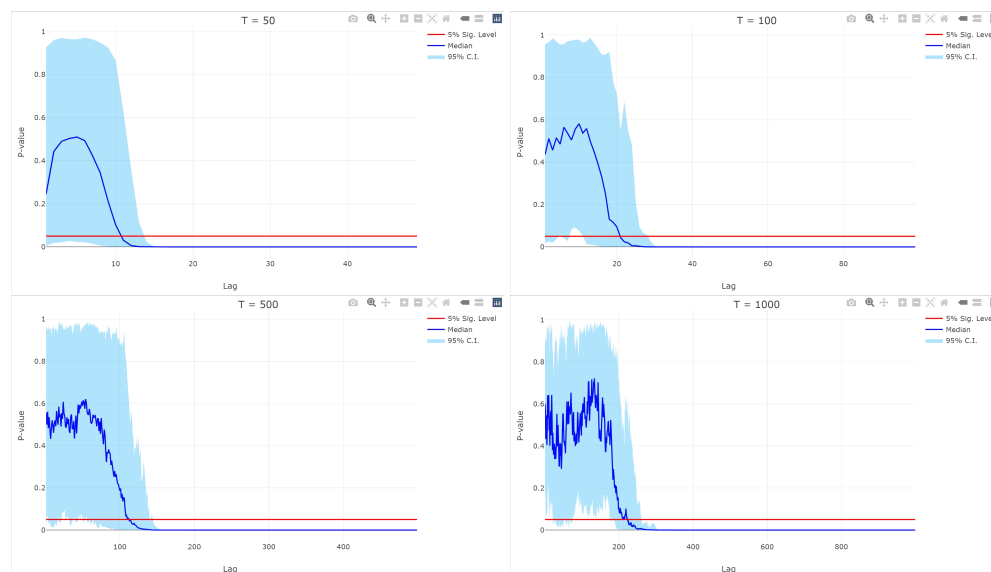

Figure 2. Trend of Shapiro–Wilk's *p*-values for testing the normality of cumulative sum of sample PACFs in different lags. Solid line shows the median of *p*-values from different simulations and the bound represents 95% confidence interval, calculated based on simulated white noise time series.

To better visualize the departure of the cumulative sum of sample PACF from a normal distribution, the density of cumulative sum of sample PACF are estimated for the lags in Table 1. Please note that the use of the keyword "cumulative" here indicates that the sum is taken from lag 1 to lag *h*. Otherwise, the meaning remains the same without using the term "cumulative". This is just for a more precise representation of the results.

Since, in practical applications of the sample PACF (e.g., finding the order of *AR* component in a model), it is common to use the sample PACF up to the lag that is $T/4$ (were *T* is the length of observed time series), the density is also estimated for lag $T/4$ (for further discussion on this selection, see, for instance, [23]). The median and 95% confidence intervals for estimated densities are presented in Figures 3–6.

As evident in Figures 3–6, as the number of lags in the sum of the sample PACF increases, the distribution departs from normality. However, the rate of this departure depends on the length of the time series.

As discussed by Fuller [27], when the coefficients of a stationary AR model are close to the unit root boundaries, the asymptotic normality of the estimated parameters converges at a slower rate, necessitating a very large number of observations for accurate inference. Fuller's argument is particularly relevant to the sample PACF distribution, especially in the $AR(1)$ case, since the first lag sample PACF, the first lag sample ACF, and the coefficient estimate are equal.

As shown in Figures 3–6, the departure of the sample PACF distribution from normality continues at later lags, even in white noise. Figure 7 also illustrates the cumulative sum of sample PACFs and their density across different lags, for 1000 simulated sample paths from the following $AR(1)$ time series:

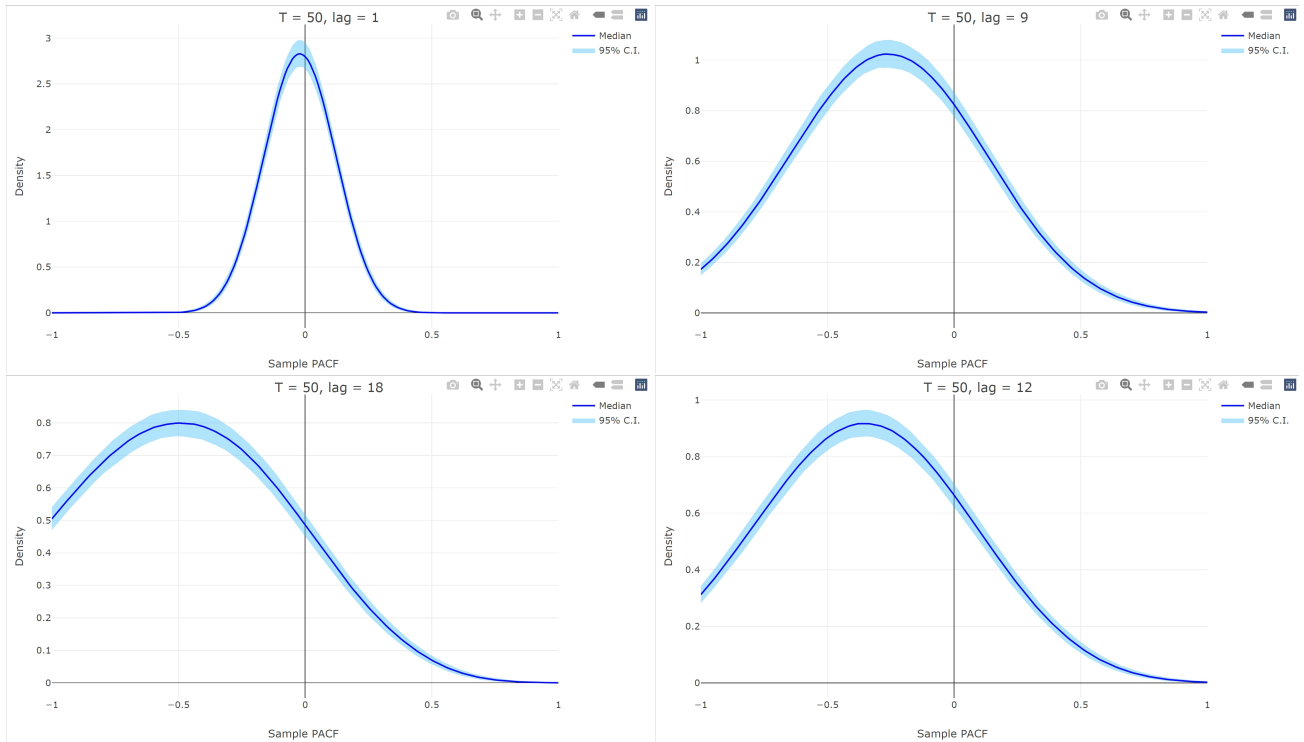$$y_t = 0.9y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0,1), \quad t = 1 \cdots, 1000.$$

Figure 3. Estimated densities of the cumulative sum of sample PACF lags in Table 1 and the lag $T/4$. The solid line represents the median of estimated densities and the bound represents the 95% confidence interval, calculated based on simulated white noise time series of length 50.
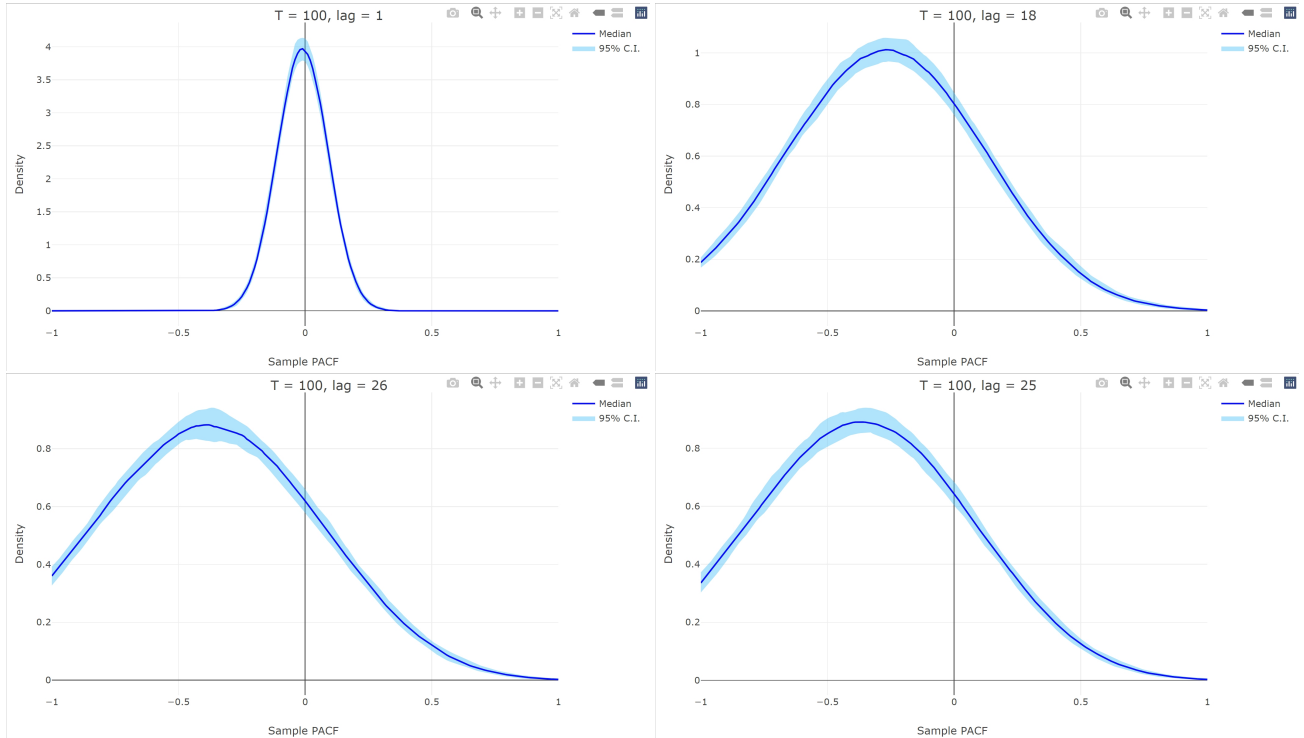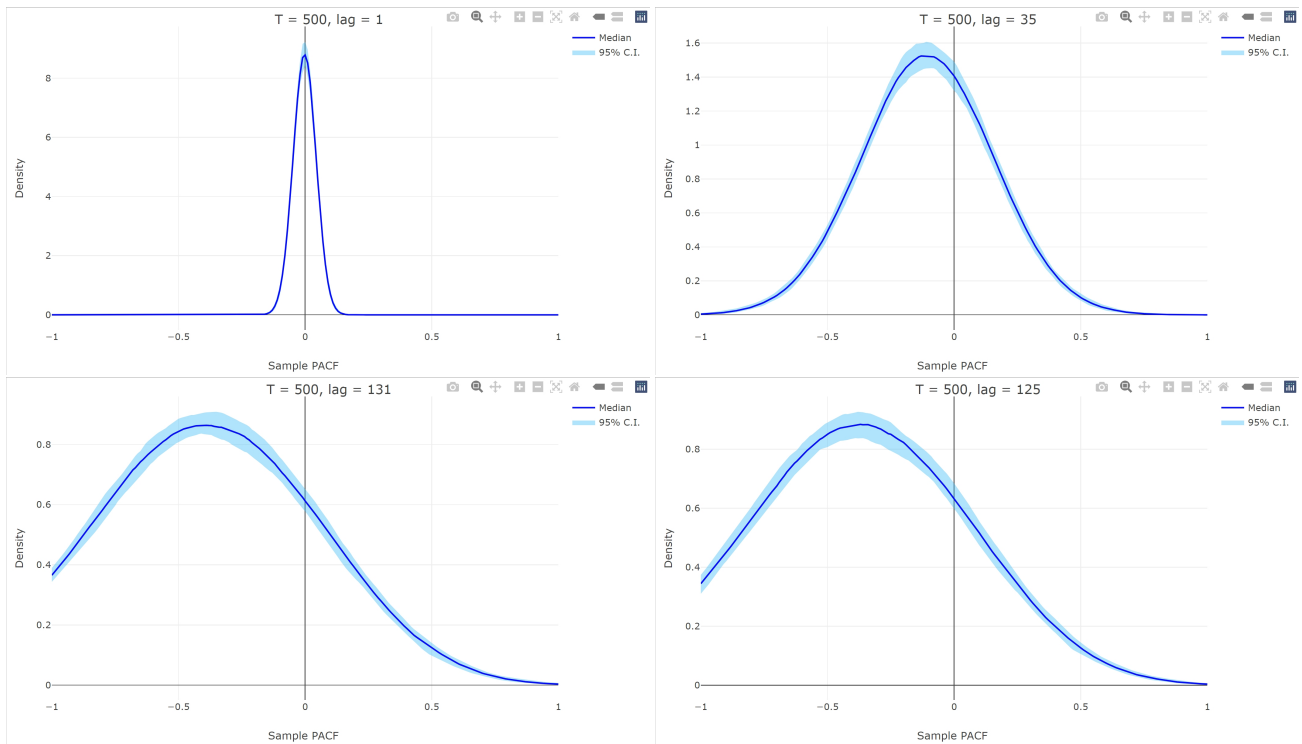

Figure 4. Estimated densities of the cumulative sum of sample PACF lags in Table 1 and the lag $T/4$. The solid line represents the median of estimated densities and the bound represents the 95% confidence interval, calculated based on simulated white noise time series of length 100.
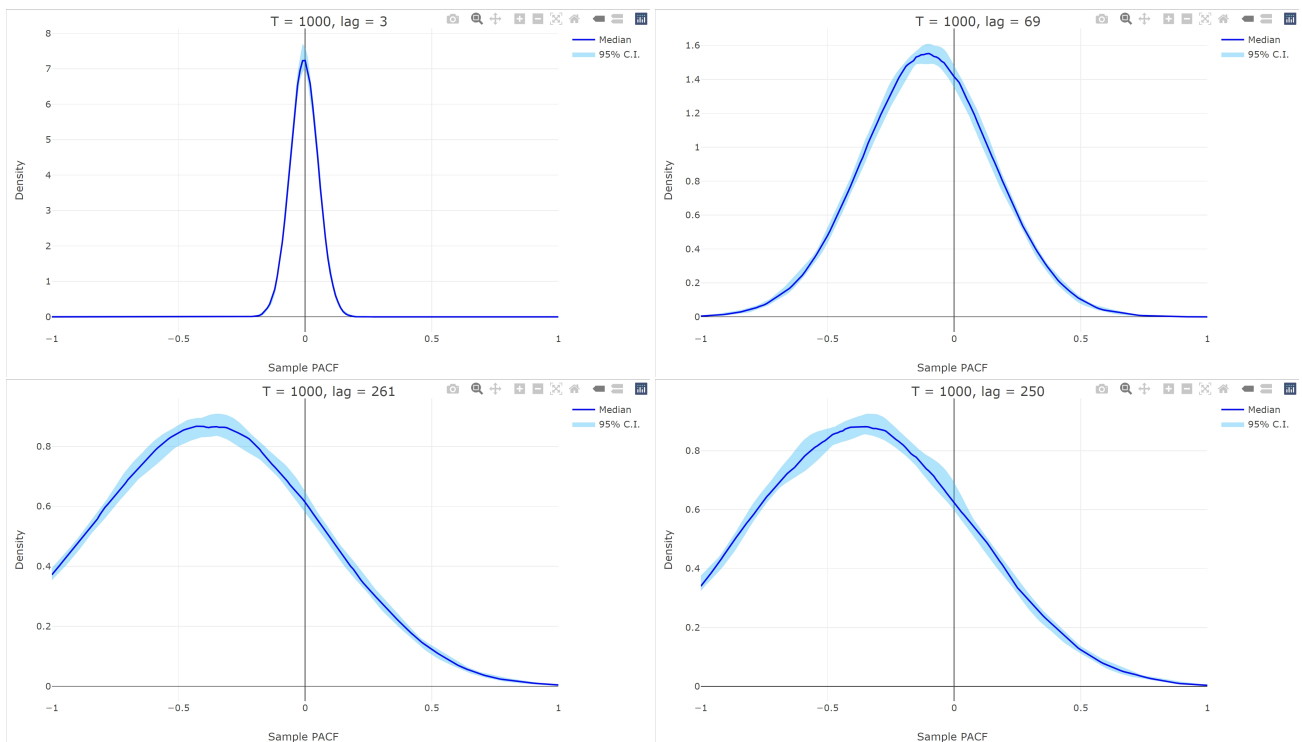
Figure 5. Estimated densities of the cumulative sum of sample PACF lags in Table 1 and the lag $T/4$. The solid line represents the median of estimated densities and the bound represents the 95% confidence interval, calculated based on simulated white noise time series of length 500.
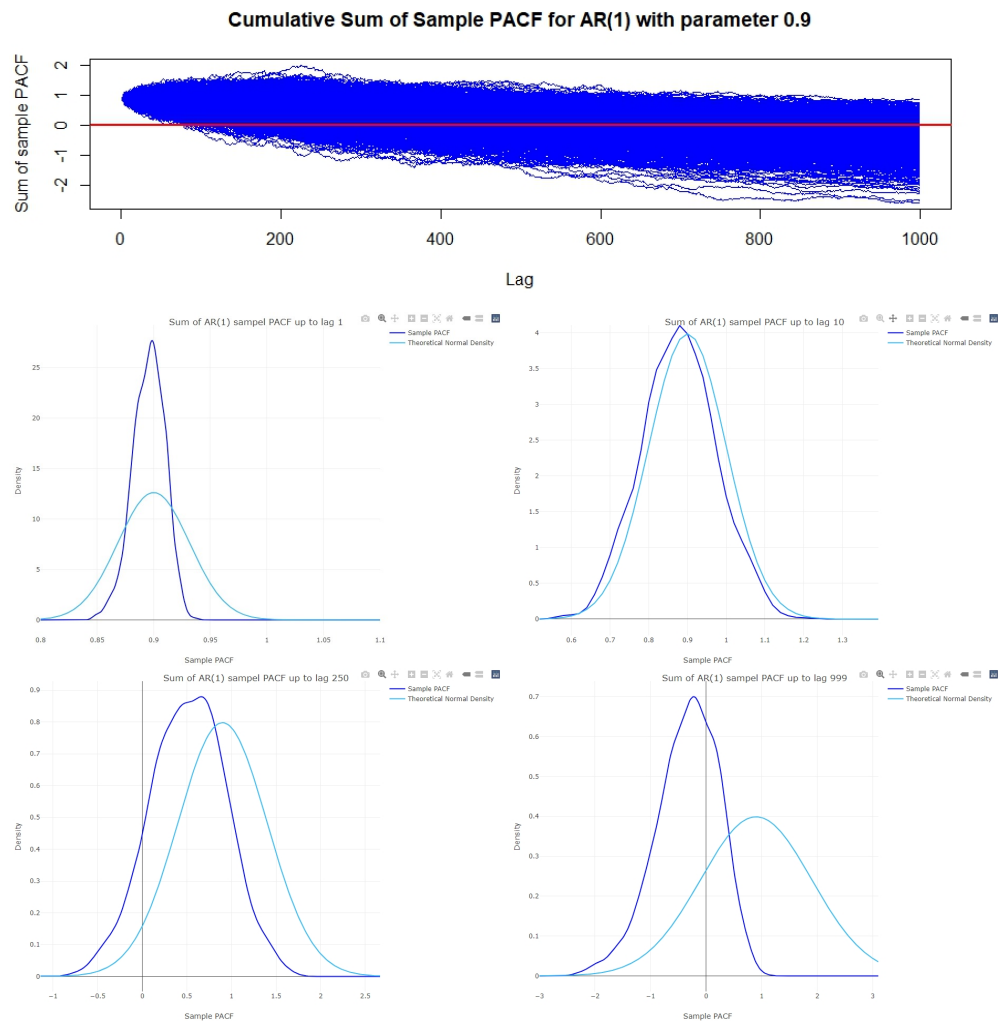

Figure 6. Estimated densities of the cumulative sum of sample PACF lags in Table 1 and the lag $T/4$. The solid line represents the median of estimated densities and the bound represents the 95% confidence interval, calculated based on simulated white noise time series of length 1000.

Figure 7. Cumulative sample PACF for sample paths simulated from $AR(1)$ model with 0.9 coefficient and their estimated density.

As can be seen, the distribution is skewed from the first lag. Furthermore, it is evident that the cumulative sum of sample PACF follows a stochastic decreasing pattern.

3.2. Monthly Bank of England Rate for 185 Years

In our investigation of time series analysis, particularly focusing on PACF, it was imperative to apply our theoretical insights to a real-world context. For this purpose, we selected a comprehensive and historically significant dataset: 185 years of monthly Bank of England Rate (Figure 8, top). This dataset not only offered a vast span of data, but also presented a unique opportunity to examine economic trends over a long period, providing a practical ground for testing our theoretical propositions. Full description and comprehensive analysis of the data are presented in [28].

Our approach involved decomposing the time series data to isolate the random noise component, thereby enabling a clearer analysis of the PACF. This was essential in understanding how theoretical models of PACF hold up against actual, observed data. The residual series, after confirming that they are white noise, is then simulated 1000 times in order to generate the distribution of the sample PACF. We applied the Shapiro–Wilk test to evaluate the normality of the sum of sample PACF distribution. This analysis was pivotal in confirming or challenging our earlier simulations and theoretical deductions regarding the behavior of sample PACF in real-world data.

The analysis of this dataset highlighted the deviations of empirical PACF from its theoretical expectations, especially beyond certain lags (see Figure 8, bottom). Thus, this real-world application of PACF not only validated our previous findings, but also provided a deeper understanding of its practical implications. It underscored the necessity for analysts and economists to consider these deviations in their models and forecasts. Moreover, it opened up avenues for future research, particularly in refining PACF models to better account for the complexities and idiosyncrasies of real-world economic data.
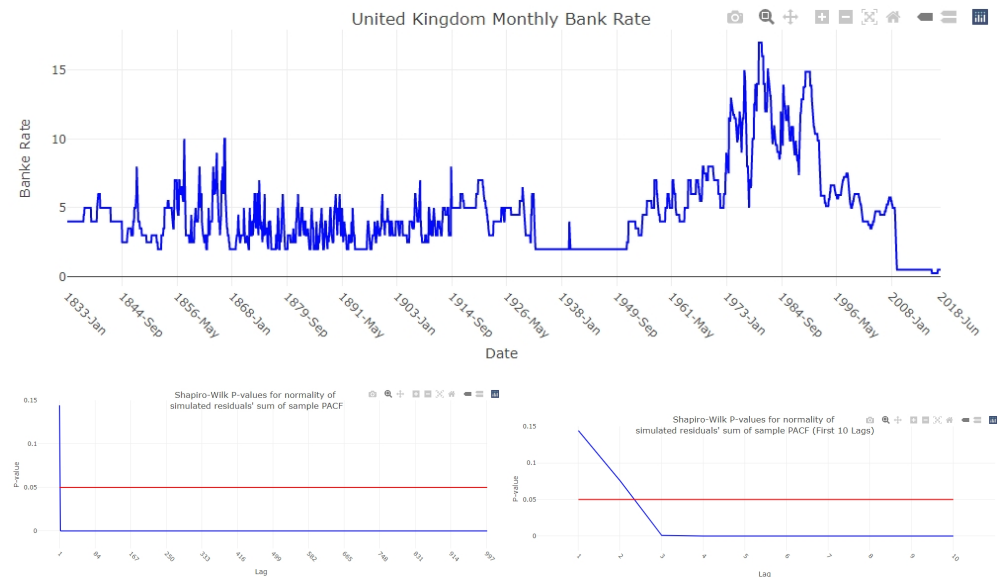

Figure 8. The **top** plot shows 185 years of monthly Bank of England rate data, while the **bottom** plot displays the *p*-value of the Shapiro–Wilk test for the simulated residual series after modeling the extracted signal.

4. Conclusions

Our analysis has demonstrated that while the PACF is a fundamental tool in uncovering the intricacies of time series data, its behavior in empirical scenarios can deviate significantly from theoretical expectations. Notably, we observed that the sum of the sample PACF and, consequently, its components, diverges from the expected normal distribution beyond a certain lag.

This deviation becomes more pronounced at larger lags, thus posing challenges to the underlying assumptions of normality in time series modeling.

The implications of these findings are profound. They compel us to reconsider and possibly revise standard practices in time series analysis, especially in the context of model selection and forecasting. The recognition that sample PACF can deviate from theoretical predictions underscores the need for a more nuanced approach in analyzing time series data, one that accounts for potential deviations and adapts to the specific characteristics of the dataset.

Furthermore, our study has highlighted the importance of a comprehensive approach that integrates both theoretical understanding and empirical analysis. By applying our findings to a real-world dataset of such historical significance, we have managed to bridge the gap between theory and practice, providing insights that are not only academically intriguing but also practically relevant.

In conclusion, our journey through the complex world of time series analysis, guided by the lens of the PACF, has revealed new challenges and opportunities. It underscores the need for continuous exploration and adaptation in the field, encouraging future research to delve deeper into these findings and further refine our understanding of time series analysis.

References

1. Tsay, R.S. *Analysis of Financial Time Series*, 3rd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2010.
2. Kirman, A.; Teyssière, G. Microeconomic Models for Long Memory in the Volatility of Financial Time Series. *Stud. Nonlinear Dyn. Econom.* **2002**, *5*. [CrossRef]
3. Zheng, X.; Basher, R.E. Structural Time Series Models and Trend Detection in Global and Regional Temperature Series. *J. Clim.* **1999**, *12*, 2347–2358. [CrossRef]
4. Sanei, S.; Hassani, H. *Singular Spectrum Analysis of Biomedical Signals*; CRC Press: Boca Raton, FL, USA, 2015.
5. Wei, W.W. *Time Series Analysis: Univariate and Multivariate Methods*, 2nd ed.; Pearson Modern Classics for Advanced Statistics Series; Pearson Addison Wesley: London, UK, 2006.
6. Hassani, H. Sum of the sample autocorrelation function. *Random Oper. Stoch. Equ.* **2009**, *17*, 125–130. [CrossRef]
7. Hassani, H. A note on the sum of the sample autocorrelation function. *Phys. Stat. Mech. Appl.* **2010**, *389*, 1601–1606. [CrossRef]
8. Brockwell, P.J.; Davis, R.A. *Introduction to Time Series and Forecasting*, 3rd ed.; Springer: Cham, Switzerland, 2016.
9. Liu, S.; Xie, Y.; Fang, H.; Du, H.; Xu, P. Trend Test for Hydrological and Climatic Time Series Considering the Interaction of Trend and Autocorrelations. *Water* **2022**, *14*, 3006. [CrossRef]
10. Dimitriadis, P.; Koutsoyiannis, D. Climacogram versus autocovariance and power spectrum in stochastic modelling for Markovian and Hurst–Kolmogorov processes. *Stoch. Environ. Res. Risk Assess.* **2015**, *29*, 1649–1669. [CrossRef]
11. Phojanamongkolkij, N.; Kato, S.; Wielicki, B.A.; Taylor, P.C.; Mlynczak, M.G. A Comparison of Climate Signal Trend Detection Uncertainty Analysis Methods. *J. Clim.* **2014**, *27*, 3363–3376. [CrossRef]
12. Xie, Y.; Liu, S.; Fang, H.; Wang, J. Global autocorrelation test based on the Monte Carlo method and impacts of eliminating nonstationary components on the global autocorrelation test. *Stoch. Environ. Res. Risk Assess.* **2020**, *34*, 1645–1658. [CrossRef]
13. Belmahdi, B.; Louzazni, M.; Bouardi, A.E. One month-ahead forecasting of mean daily global solar radiation using time series models. *Optik* **2020**, *219*, 165207. [CrossRef]
14. Yang, Y.; Qin, S.; Liao, S. Ultra-chaos of a mobile robot: A higher disorder than normal-chaos. *Chaos Solitons Fractals* **2023**, *167*, 113037. [CrossRef]
15. Gostischa, J.; Massolo, A.; Constantine, R. Multi-Species Feeding Association Dynamics Driven by a Large Generalist Predator. *Frontiers Mar. Sci.* **2021**, *8*, 739894. [CrossRef]
16. Orlando, G.; Bufalo, M. Empirical Evidences on the Interconnectedness between Sampling and Asset Returns' Distributions. *Risks* **2021**, *9*, 88. [CrossRef]
17. Bai, M.; Zhou, Z.; Chen, Y.; Liu, J.; Yu, D. Accurate four-hour-ahead probabilistic forecast of photovoltaic power generation based on multiple meteorological variables-aided intelligent optimization of numeric weather prediction data. *Earth Sci. Inf.* **2023**, *16*, 2741–2766. [CrossRef]
18. Wang, X.; Yang, J.; Yang, F.; Wang, Y.; Liu, F. Multilevel Residual Prophet Network Time Series Model for Prediction of Irregularities on High-Speed Railway Track. *J. Transp. Eng. Part A Syst.* **2023**, *149*, 04023012. [CrossRef]
19. Shumway, R.H.; Stoffer, D.S. *Time Series Analysis and Its Applications: With R Examples*; Springer International Publishing: Berlin/Heidelberg, Germany, 2017. [CrossRef]
20. Box, G.E.P.; Jenkins, G.M.; Reinsel, G.C. *Time Series Analysis: Forecasting and Control*; Wiley Series in Probability and Statistics; John Wiley & Sons: Hoboken, NJ, USA, 2008. [CrossRef]
21. Hassani, H.; Leonenko, N.; Patterson, K. The sample autocorrelation function and the detection of long-memory processes. *Phys. A Stat. Mech. Appl.* **2012**, *391*, 6367–6379. [CrossRef]
22. Hassani, H.; Yeganegi, M.R. Sum of squared ACF and the Ljung–Box statistics. *Phys. A Stat. Mech. Appl.* **2019**, *520*, 81–86. [CrossRef]
23. Hassani, H.; Yeganegi, M.R. Selecting optimal lag order in Ljung–Box test. *Phys. A Stat. Mech. Appl.* **2020**, *541*, 123700. [CrossRef]
24. Hassani, H.; Yeganegi, M.R.; Mehreyan, S.Z.; Sayyareh, A. On the Sample Autocorrelation Function's Absolute Summability. *Fluct. Noise Lett.* **2022**, *21*, 2250004. [CrossRef]
25. Hassani, H.; Yarmohammadi, M.; Mashhad, L.M. Uncovering Hidden Insights with Long-Memory Process Detection: An In-Depth Overview. *Risks* **2023**, *11*, 113. [CrossRef]
26. Shapiro, S.S.; Wilk, M.B. An analysis of variance test for normality (complete samples). *Biometrika* **1965**, *52*, 591–611. [CrossRef]

27. Fuller, W.A. *Introduction to Statistical Time Series*; Wiley Series in Probability and Statistics; John Wiley & Sons: Hoboken, NJ, USA, 1995. [CrossRef]
28. Hassani, H.; Yeganegi, M.R.; Cuñado, J.; Gupta, R. Forecasting interest rate volatility of the United Kingdom: Evidence from over 150 years of data. *J. Appl. Stat.* **2020**, *47*, 1128–1143. [CrossRef]