STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
*Stat., Optim. Inf. Comput.*, Vol. **13**, March 2025, pp 1339–1356.
Published online in International Academic Press (www.IAPress.org)

# Application of the Periodic Self-Exciting Threshold Autoregressive Model

N. Bezziche, M. Merzougui *

*LaPS Laboratory, Badji Mokhtar-Annaba University, 12, P.O.Box, 23000 Annaba, Algeria*

**Abstract** In this paper, we analyze Algerian temperature data using the periodic self-exciting threshold autoregressive (PSETAR) model. While the periodic SETAR model offers significant advantages in capturing seasonal and threshold-based behaviors, it remains underutilized in practical applications. The primary motivation of this work is to demonstrate the utility of the PSETAR model for modeling complex temperature dynamics in Algeria. We aim to fill the gap in the literature by showcasing how this model effectively captures seasonal variations and threshold behaviors in temperature data. The paper contributes by providing a thorough analysis of the PSETAR model, discussing its estimation via the least squares method, and testing its linearity using the likelihood ratio test. Additionally, we extend the concept of local asymptotic normality to include p regimes. This study not only deepens the understanding of Algeria's temperature dynamics but also highlights the advantages of using the PSETAR model to address complex real-world time series problems.

**Keywords** Periodic Self-Exciting Threshold Autoregressive models, LS estimation, LR test, LAN property, Algeria's temperature.

**AMS 2010 subject classifications** 62F12, 62M10

**DOI:** 10.19139/soic-2310-5070-2254

## 1. Introduction

This study aims to demonstrate the validity and effectiveness of the Periodic Self-Exciting Threshold Autoregressive model. Our focus is on capturing the nonlinear dynamics and periodic regimes in climatic time series data, specifically the average temperature in Algeria. While numerous studies have applied time series models to various datasets, we emphasize the adaptability and robustness of the PSETAR model in scenarios characterized by distinct regimes, such as cold and warm periods. Despite Algeria's diverse climate, ranging from coastal to Saharan zones, our analysis centers on the national average for simplicity. This approach enables us to highlight the model's ability to identify broader periodic structures rather than focusing on regional variations. Climatic data has been extensively studied by researchers, including [13, 15, 17, 4], and others. Since temperature data often exhibit distinct regimes, we employ the Self-Exciting Threshold Autoregressive (SETAR) model [19, 8, 12] to address the nonlinearities inherent in such time series. This model facilitates segmentation through threshold selection, allowing the analysis of dynamic relationships within individual regimes. Moreover, because monthly temperature data typically follows seasonal patterns, we extend this framework by applying the PSETAR model, specifically designed to capture periodic nonlinearity. Periodic models are crucial for understanding seasonal behaviors, as shown in various theoretical and empirical studies [9, 1, 2, 10, 11]. The PSETAR model was initially introduced by [14], while adaptive tests for periodicity within SETAR frameworks were developed by [6]. Before applying the model, we examine essential statistical properties, including parameter estimation using the least squares method and linearity testing via the likelihood ratio test. This study also investigates the locally asymptotically normal

*Correspondence to: Merzougui Mouna (Email: mouna.merzougui@univ-annaba.dz). Department of Mathematics, Badji Mokhtar-Annaba University, 12, P.O.Box, 23000 Annaba, Algeria.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

(LAN) property, which has been extended to multiple regimes, as demonstrated by [5] in their adaptive estimation of the two-regime PSETAR model. The LAN property is a cornerstone of time series analysis [3, 16], offering a framework for simplifying estimation and hypothesis testing in complex, nonlinear models. While the discussion of the LAN property in this paper primarily aims to demonstrate its validity through simulations, its inclusion reinforces the theoretical foundation of this study. The remainder of this paper is organized as follows: Section 2 discusses the model specification, parameter estimation using the least squares method, and linearity testing. Section 3 delves into the LAN property in greater detail. Finally, Section 4 presents the application of the PSETAR model to Algeria's temperature series, highlighting key findings and implications.

## 2. Periodic Self Exciting Threshold Autoregressive Model

### *2.1. Model Estimation*

The stochastic process $\{X_t; t \in \mathbb{Z}\}$ is characterized by a PSETAR$(p, 1, \ldots, 1)$ representation with period $S$ ($S \geq 2$), defined as:

$$X_t = \sum_{i=1}^{p} \phi_{t,i} X_{t-1} \mathbb{I}(c_{t,i-1} < X_{t-1} \leq c_{t,i}) + \varepsilon_t, \quad t \in \mathbb{Z}. \tag{1}$$

Here, $\{\varepsilon_t; t \in \mathbb{Z}\}$ is an independent and identically distributed sequence with mean 0 and variance $\sigma_t^2$, where the probability density function is unknown. The autoregressive parameters $\phi_{t,1}, \ldots, \phi_{t,p}$ and the innovation variance $\sigma_t^2$ vary periodically with a period of $S$. The threshold parameters $c_{t,i}$ are defined such that $c_{t,0} = -\infty$, $c_{t,p} = +\infty$, $\forall t$, while the remaining thresholds are periodic with period $S$. For convenience, it is assumed that these thresholds are known, given the relatively subjective nature of their estimation.

The model can be reformulated using indicator functions:

$$X_t = \phi_{t,1} X_{t-1}^1 + \ldots + \phi_{t,p} X_{t-1}^p + \varepsilon_t, \tag{2}$$

where

$$X_{t-1}^i = X_{t-1} \mathbb{I}(c_{t,i-1} < X_{t-1} \leq c_{t,i}).$$

To clarify, let $t = S\tau + s$. Then we have:

$$X_{S\tau+s} = \phi_{s,1} X_{S\tau+s-1}^1 + \ldots + \phi_{s,p} X_{S\tau+s-1}^p + \varepsilon_{S\tau+s}, \quad s = 1, \ldots, S, \ \tau \in \mathbb{Z}. \tag{3}$$

Let $\underline{\phi} = (\underline{\phi}_1', \underline{\phi}_2', \ldots, \underline{\phi}_S')' \in \mathbb{R}^{pS}$, where $\underline{\phi}_s = (\phi_{s,1}, \ldots, \phi_{s,p})'$ represents the parameter vector.

A realization of the periodic SETAR(2,1,1) model is illustrated in Figure 1, where $\underline{\phi} = (0.7, -1.4; -0.5, 0.5)'$ and the threshold $c = 0$. The density is observed to be non-normal, as indicated by the Shapiro-Wilk normality test, with a $p$-value of 0.009199. The presence of nonlinearity is further confirmed by the Keenan test, yielding a $p$-value of $6.171129 \times 10^{-8}$.

We consider observations $\{X_1, \ldots, X_N\}$ from Equation (3), where $N = mS$ represents $m$ years and $S$ seasons. With known thresholds, the model is represented as a periodically linear regression, allowing us to derive the Least Squares (LS) estimators. Under necessary assumptions, we have:

$$\hat{\underline{\phi}}_s = \begin{bmatrix} \sum_{\tau=0}^{m-1} (X_{S\tau+s-1}^1)^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sum_{\tau=0}^{m-1} (X_{S\tau+s-1}^p)^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{\tau=0}^{m-1} X_{S\tau+s} X_{S\tau+s-1}^1 \\ \vdots \\ \sum_{\tau=0}^{m-1} X_{S\tau+s} X_{S\tau+s-1}^p \end{bmatrix}, \ s = 1, \ldots, S. \tag{4}$$

$$\hat{\sigma}_s^2 = \frac{1}{m} \sum_{\tau=0}^{m-1} \left( X_{S\tau+s} - \phi_{s,1} X_{S\tau+s-1}^1 - \ldots - \phi_{s,p} X_{S\tau+s-1}^p \right)^2, \ s = 1, \ldots, S.$$

The asymptotic distribution of the estimator is given by:

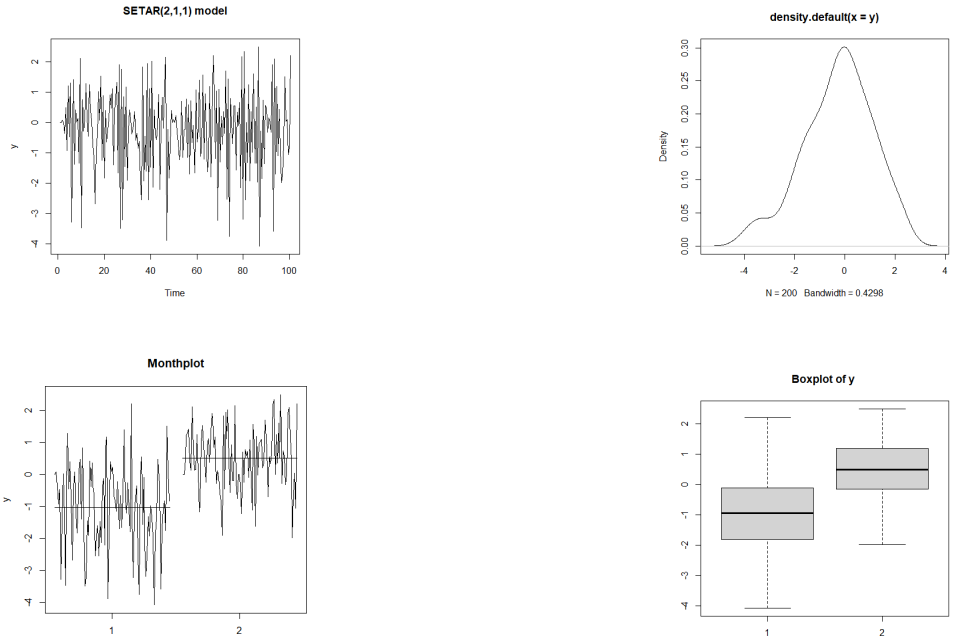


Figure 1. Periodic SETAR(2,1,1) model.

$$\sqrt{m}(\hat{\underline{\phi}}_s - \phi_s) \xrightarrow{d} \mathcal{N}(\mathbf{0}_p, \sigma_s^2 \Gamma_s^{-1}), \quad \text{for } s = 1, \ldots, S.$$

where

$$\Gamma_s = \begin{bmatrix} E(X^1_{S\tau+s-1})^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & E(X^p_{S\tau+s-1})^2 \end{bmatrix}. \tag{5}$$

We conducted a small simulation study to verify the effectiveness of the estimators. Observations were generated from PSETAR$_2$(2,1,1) models with parameters $\underline{\phi} = (0.9, -0.3; -0.5, 0.8)'$ and variance $\sigma_t^2 = (1, 0.8)'$, using sample sizes ranging from 100 to 300. The threshold was set to 0. Table 1 presents the LS estimators and their variances. Additionally, we generated a PSETAR$_2$(3,1,1) model with parameters $\underline{\phi} = (0.9, -1, 0.2; -0.5, 1.6, -0.1)'$ and the same periodic variance. Sample sizes of $n = 300$ and $n = 600$ were used, with thresholds set at $-1$ and 1. The results are summarized in Table 2. Consistent means across different sample sizes suggest stability in the estimation process, while decreasing variances with larger sample sizes indicate increased precision and reliability in the parameter estimates.

Table 1. LS estimators of PSETAR$_2$(2,1,1) model.

| **n** | $\phi_{1,1}$ | $\phi_{1,2}$ | $\phi_{2,1}$ | $\phi_{2,2}$ | |
|---|---|---|---|---|---|
| 100 | Mean | 0.8857 | -0.3017 | -0.5045 | 0.7916 |
| | Var | 0.1408 | 0.0210 | 0.0192 | 0.0685 |
| 200 | Mean | 0.9093 | -0.2990 | -0.5015 | 0.7963 |
| | Var | 0.0629 | 0.0101 | 0.0085 | 0.0308 |
| 300 | Mean | 0.8972 | -0.2990 | -0.5011 | 0.7958 |
| | Var | 0.0390 | 0.0063 | 0.0059 | 0.0183 |

Table 2. LS estimators of $\text{PSETAR}_2(3,1,1)$ model.

| **n** | | $\phi_{1,1}$ | $\phi_{1,2}$ | $\phi_{1,3}$ | $\phi_{2,1}$ | $\phi_{2,2}$ | $\phi_{2,3}$ |
|---|---|---|---|---|---|---|---|
| 300 | Mean | 0.8964 | -0.9943 | 0.1959 | -0.5025 | 1.6063 | -0.1028 |
| | Var | 0.0134 | 0.0395 | 0.0084 | 0.0059 | 0.0316 | 0.0111 |
| 600 | Mean | 0.9059 | -1.0049 | 0.1975 | -0.5020 | 1.6031 | -0.0993 |
| | Var | 0.0075 | 0.0178 | 0.0043 | 0.0028 | 0.0169 | 0.0055 |

### *2.2. Test*

In the specific scenario where $\phi_{s,1} = \cdots = \phi_{s,p} = \phi_s, \forall s = 1, \ldots, S$, the model (3) simplifies to the periodic AR(1) model. In this section, we employ the likelihood ratio test (LR) to examine the linearity. The hypotheses under consideration are:

$$H_0 : X_t \sim \text{PAR}_S(1) \quad \text{vs} \quad H_1 : X_t \sim \text{PSETAR}_S(p, 1, \ldots, 1).$$

The LR test rejects $H_0$ at the asymptotic level $\alpha$ if

$$LR_m = m \sum_{s=1}^{S} \log\left(\frac{\tilde{\sigma}_s^2}{\hat{\sigma}_s^2}\right) > \chi^2_{(p-1)S}(1-\alpha),$$

where $\tilde{\sigma}_s^2$ and $\tilde{\phi}_s$ are the estimators under $H_0$, given by

$$\tilde{\sigma}_s^2 = \frac{1}{m} \sum_{\tau=0}^{m-1} \left(X_{S\tau+s} - \phi_s X_{S\tau+s-1}\right)^2,$$

and

$$\tilde{\phi}_s = \frac{\sum_{\tau=0}^{m-1} X_{S\tau+s-1} X_{S\tau+s}}{\sum_{\tau=0}^{m-1} X^2_{S\tau+s-1}}.$$

This LR test provides a statistical method to assess the linearity of the model, comparing the fit of the PAR model to the more general PSETAR model. The decision to reject or accept $H_0$ is based on the comparison of the LR statistic to the critical value $\chi^2_{(p-1)S}(1-\alpha)$.

## 3. Local Asymptotique Normality (LAN)

### *3.1. Notation and Assumptions*

Let $H_f^{(n)}(\underline{\phi})$ be a sequence of null hypotheses, under which $X_t^{(n)}$ is a realization of the process (3) with the parameter $\underline{\phi}$. $H_f^{(n)}(\underline{\phi}^{(n)})$ represents the sequence of alternatives with the central parameter:

$$\underline{\phi}^{(n)} = \underline{\phi} + \frac{1}{\sqrt{n}} \underline{h}^{(n)}, \quad \underline{h}^{(n)} \in \mathbb{R}^{pS},$$

such that $\sup_n \underline{h}^{(n)\prime} \underline{h}^{(n)} < \infty$. Let $\underline{\tau}^{(n)} = (\underline{\tau}_1^{(n)}, \underline{\tau}_2^{(n)}, \ldots, \underline{\tau}_S^{(n)})'$, where $\tau_s^{(n)} = (h_{s,1}^{(n)}, h_{s,2}^{(n)})'$, for $s = 1, 2, \ldots, S$.

To establish the LAN property, the following assumptions are required on the model to ensure differentiability in quadratic mean:

Assumption (A.1). The model is periodically stationary. A sufficient condition is given by $\phi_{s,1} < 1$, $\phi_{s,p} < 1$ and $\phi_{s,1}\phi_{s,p} < 1$, for $s = 1, 2, \ldots, S$.

Assumption (A.2). The innovation density $f(.)$ is supposed to satisfy the following conditions:

(i) $f(x) > 0$, $\forall x \in \mathbb{R}$, $\int x f(x)\,dx = 0$ and $\sigma_t^2 = \mathbb{E}(\epsilon_t^2) < \infty$.
(ii) $f(.)$ is absolutely continuous with respect to the Lebesgue measure $\mu$.
(iii) The Fisher information $I(f) = \int (\phi_f(x))^2 f(x)\,dx$ is finite, where $\boldsymbol{\phi}_f = -\frac{f'}{f}$.

### *3.2. Local asymptotic normality*

For simplicity, let's assume that the size of the observed time series, $n$, is a multiple of $S$, i.e., $n = mS$, $m \in \mathbb{N}^*$, and let $t = s + S\tau$, $s = 1, \ldots, S$ and $\tau = 0, 1, \ldots, m-1$. Then we have:

$$Z_{s,\tau}^{(n)}(\underline{\phi}_s) = X_{S\tau+s}^{(n)} - \phi_{s,1} X_{S\tau+s-1}^{(n)1} - \cdots - \phi_{s,p} X_{S\tau+s-1}^{(n)p}, \quad s = 1, \ldots, S,\ \tau = 0, \ldots, m-1.$$

These calculated residuals under $H_f^{(n)}(\underline{\phi})$ coincide with $\epsilon_{S\tau+s}$. Additionally, we have:

$$Z_{s,\tau}^{(n)}(\underline{\phi}_s^{(n)}) = Z_{s,\tau}^{(n)}(\underline{\phi}_s) - \frac{1}{\sqrt{n}} h_{s,1}^{(n)} X_{S\tau+s-1}^{(n)1} - \cdots - \frac{1}{\sqrt{n}} h_{s,p}^{(n)} X_{S\tau+s-1}^{(n)p}.$$

These are the calculated residuals under the alternatives.
The logarithm of the likelihood ratio is given by:
$\Lambda_f^{(n)}\left(\underline{\phi} + \frac{1}{\sqrt{n}}\underline{\tau}^{(n)}\right) = \sum_{s=1}^{S}\sum_{\tau=0}^{m-1}\left(\log f_{\sigma_s}\left(Z_{s,\tau}^{(n)}(\underline{\phi}_s) - \frac{1}{\sqrt{n}} h_{s,1}^{(n)} X_{S\tau+s-1}^{(n)1} - \cdots - \frac{1}{\sqrt{n}} h_{s,p}^{(n)} X_{S\tau+s-1}^{(n)p}\right) - \log f_{\sigma_s}\left(Z_{s,\tau}^{(n)}(\underline{\phi}_s)\right)\right) + o_p(1).$
The central sequence is given by:

$$\Delta^{(n)}(\underline{\phi}) = \begin{bmatrix} \Delta_1^{(n)}(\underline{\phi}_1) \\ \vdots \\ \Delta_S^{(n)}(\underline{\phi}_S) \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{n}}\sum_{\tau=0}^{m-1} \boldsymbol{\phi}_{\sigma_1}\left(Z_{1,\tau}^{(n)}(\underline{\phi}_1)\right) X_{S\tau}^{(n)1} \\ \vdots \\ \frac{1}{\sqrt{n}}\sum_{\tau=0}^{m-1} \boldsymbol{\phi}_{\sigma_1}\left(Z_{1,\tau}^{(n)}(\underline{\phi}_1)\right) X_{S\tau}^{(n)p} \\ \vdots \\ \frac{1}{\sqrt{n}}\sum_{\tau=0}^{m-1} \boldsymbol{\phi}_{\sigma_S}\left(Z_{S,\tau}^{(n)}(\underline{\phi}_S)\right) X_{S-1+S\tau}^{(n)1} \\ \vdots \\ \frac{1}{\sqrt{n}}\sum_{\tau=0}^{m-1} \phi_{\sigma_S}\left(Z_{S,\tau}^{(n)}(\underline{\phi}_S)\right) X_{S-1+S\tau}^{(n)p} \end{bmatrix}.$$

The $pS \times pS$ squared block diagonal information matrix is expressed as:

$$\Gamma(\underline{\phi}) = \begin{bmatrix} \frac{\Gamma_1(\underline{\phi})}{\sigma_1^2} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{\Gamma_S(\underline{\phi})}{\sigma_S^2} \end{bmatrix},$$

where $\Gamma_s(\underline{\phi}) = \Gamma_s$, given by formula (5).
The following proposition establishes the Local Asymptotic Normality property.

**Proposition 3.1.** Under the regularity conditions (A1)-(A2) and under $H_f^{(n)}(\underline{\phi})$, we have, as $n \to \infty$ :

$$1) \quad \Lambda_f^{(n)}\left(\underline{\phi} + \frac{1}{\sqrt{n}}\underline{\tau}^{(n)}\right) = \underline{\tau}^{(n)\prime}\Delta^{(n)}(\underline{\phi}) - \frac{I(f_1)}{2S}{}^{(n)\prime}\Gamma(\underline{\phi})\underline{\tau}^{(n)} + o_p(1),$$

$$2) \quad \mathcal{L}\left(\frac{\Delta^{(n)}(\underline{\phi})}{\mathbb{P}_{n,\underline{\phi}}}\right) \Rightarrow N\left(0, \frac{I(f_1)}{S}\Gamma(\underline{\phi})\right).$$

**Proof** The proof of this LAN result relies on Swensen's Conditions (1985) [18] and is a straightforward extension of the case $p = 2$ in [5]. The most crucial point is the differntiability in mean quadratic:

$$\int \left(f_{\sigma_s}^{\frac{1}{2}}\left(x - \frac{1}{\sqrt{n}} h_{s,1}^{(n)} X_{S\tau+s-1}^{(n)1} - \ldots - \frac{1}{\sqrt{n}} h_{s,p}^{(n)} X_{S\tau+s-1}^{(n)p}\right) - f_{\sigma_s}^{\frac{1}{2}}(x)\right.$$

$$-\frac{1}{2}\phi_{\sigma_s}(x)\left(\frac{1}{\sqrt{n}}h_{s,1}^{(n)}X_{S\tau+s-1}^{(n)1}-\ldots-\frac{1}{\sqrt{n}}h_{s,p}^{(n)}X_{S\tau+s-1}^{(n)p}\right)\Bigg)^2 dx=o(\|h_s\|^2).$$

On the other hand, for $s$ fixed, this reduces to SETAR$(p,1,\ldots,1)$ model considered by ([7].

### *3.3. Simulation results*

In our simulation studies, we have confirmed the asymptotic normality of the central sequence. Specifically, we simulated PSETAR(2,1,1) models with a periodicity of 2, varying the sample sizes at $n=100,500$, and 1000. The parameter values used for the models are $\phi=(0.9,-0.3;-0.5,0.8)'$ and $\sigma^2=(1,0.8)'$. The central sequence, denoted as $\Delta_k^{(n)}(\varphi)$ for $k=1$ to 1000, was computed for $k=1$ to 1000 under both logistic and Gaussian densities.

Figures 2, 3, and 4 offer a detailed comparison via kernel density estimates of the central sequence against its asymptotic counterpart, as stated in Proposition 3.1. Additionally, a Normal Q-Q plot under the Gaussian density is presented for further insight. The results are compelling, demonstrating a strong alignment between the central sequence and the theoretical normal distribution.

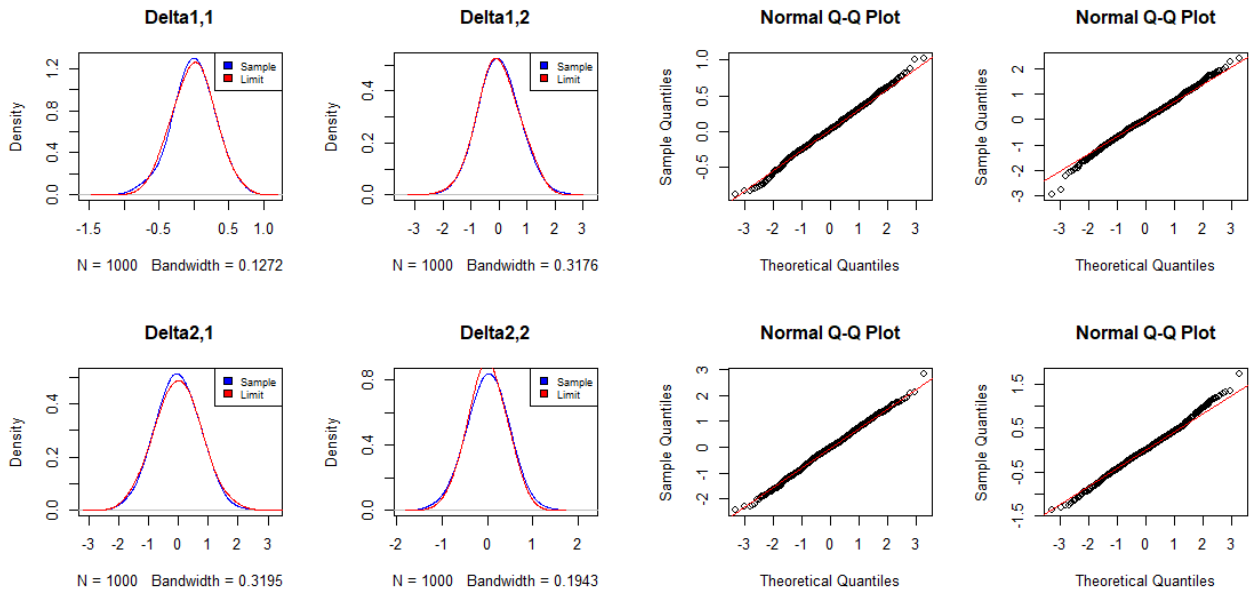


Figure 2. Theoretical and asymptotic law of $\Delta_k^{(n)}(\varphi)$ with Normal innovation and $n=100$.

On the other hand, we conducted simulations of the same model with a different set of parameters, specifically $\phi=(0.8,-0.4;-0.6,0.7)'$ under the logistic density.

Figures 5, 6, and 7 present the probability densities of the central sequence alongside the Gaussian density, with the mean and variance of $\Delta_k^{(n)}(\varphi)$. Additionally, Q-Q plots are included to visually assess the goodness of fit of the distributions.

The affirmation of normality is consistent across all cases, as corroborated by the Normal Q-Q plots. These plots systematically demonstrate that the central sequence adheres closely to a normal distribution. Regardless of the parameter variations and the law in Figures 2, to 7, the observed quantiles align closely with the expected quantiles under the assumption of normality.

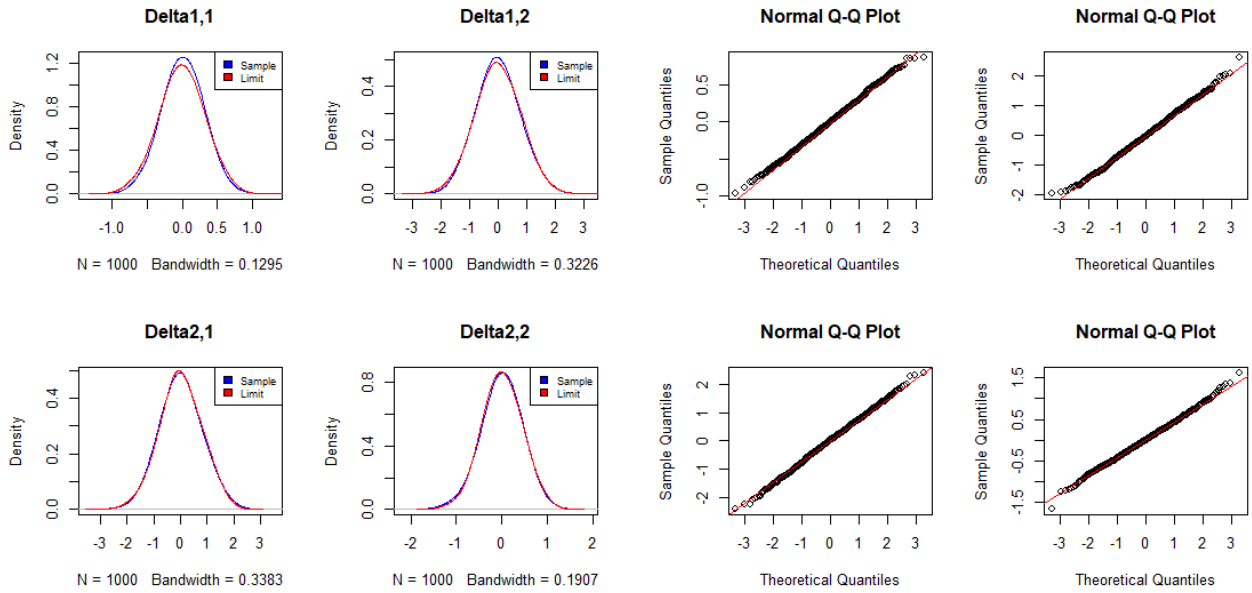


Figure 3. Theoretical and asymptotic law of $\Delta_k^{(n)}(\varphi)$ with Normal innovation and $n = 500$.

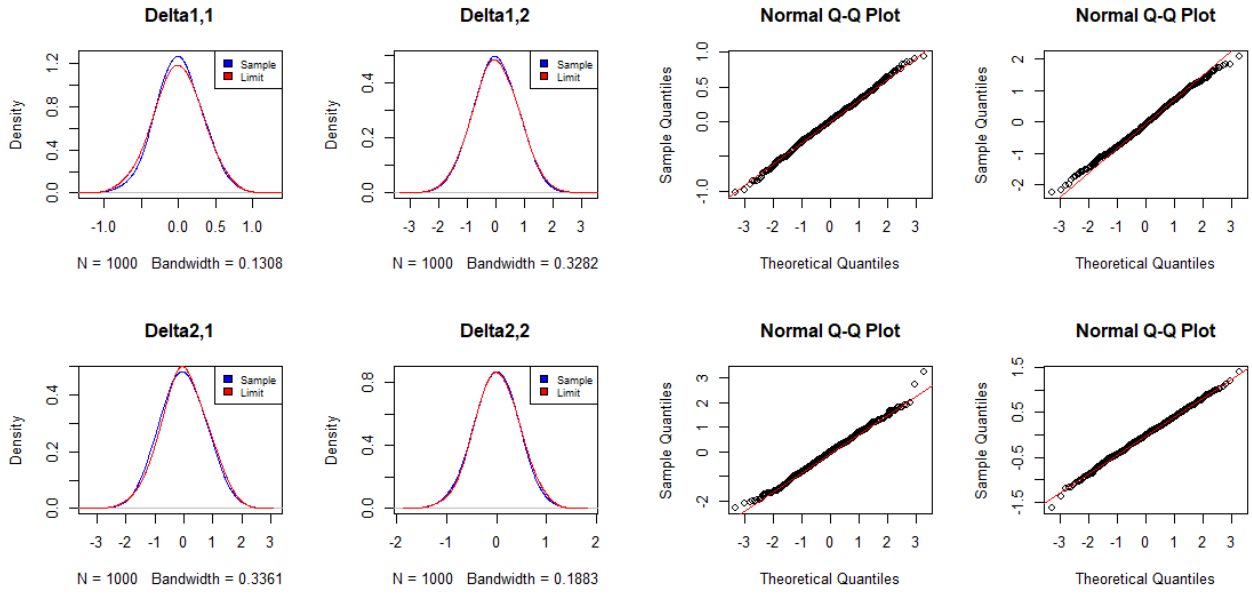


Figure 4. Theoretical and asymptotic law of $\Delta_k^{(n)}(\varphi)$ with Normal innovation and $n = 1000$.

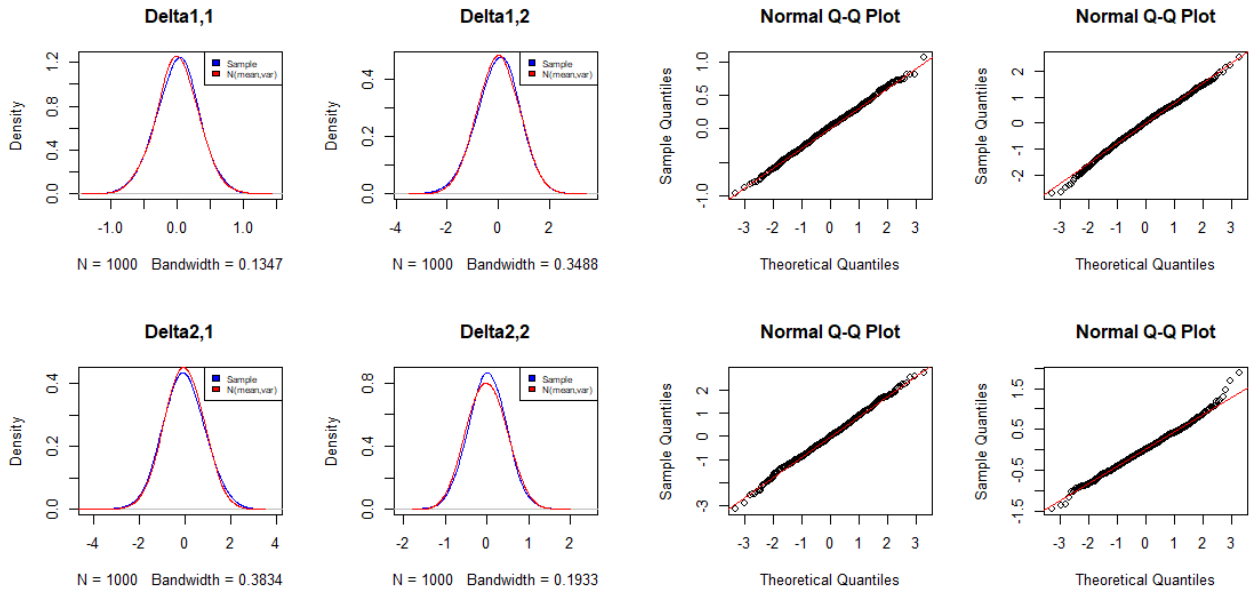


Figure 5. Theoretical and asymptotic law of $\Delta_k^{(n)}(\varphi)$ with logistic innovation and $n = 100$.

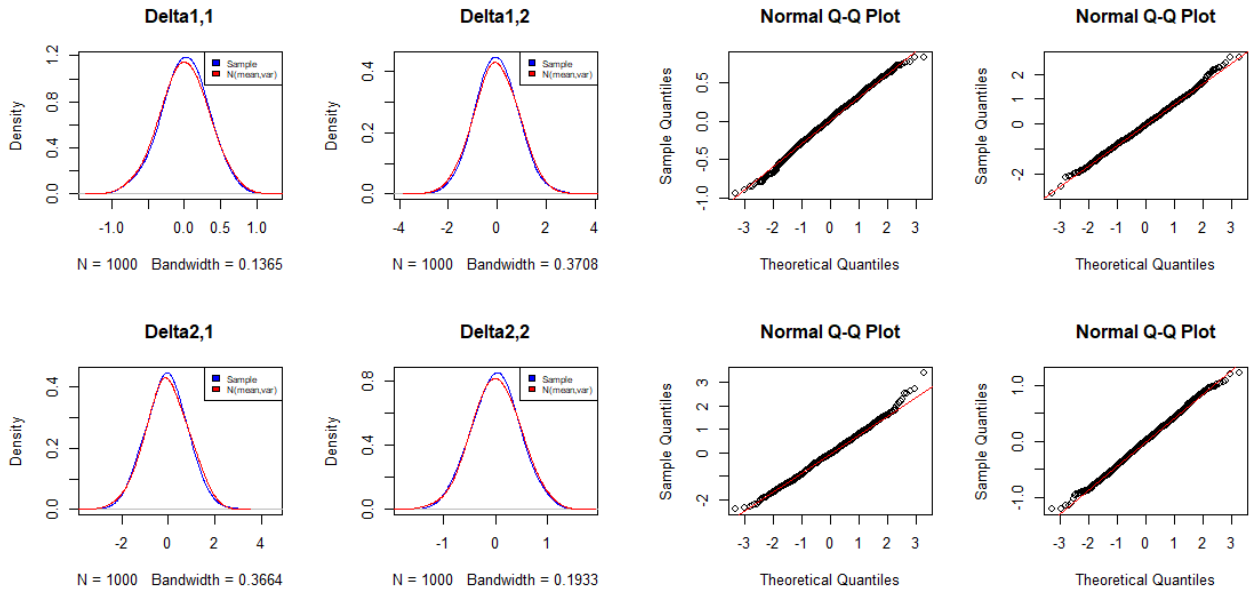


Figure 6. Theoretical and asymptotic law of $\Delta_k^{(n)}(\varphi)$ with logistic innovation and $n = 500$.

## 4. Application

### *4.1. Data Description*

The temperature data used in this analysis was obtained from the World Bank Climate Knowledge Portal, covering the entire 59-year period from January 1901 to December 1960. The data is recorded monthly and represents the

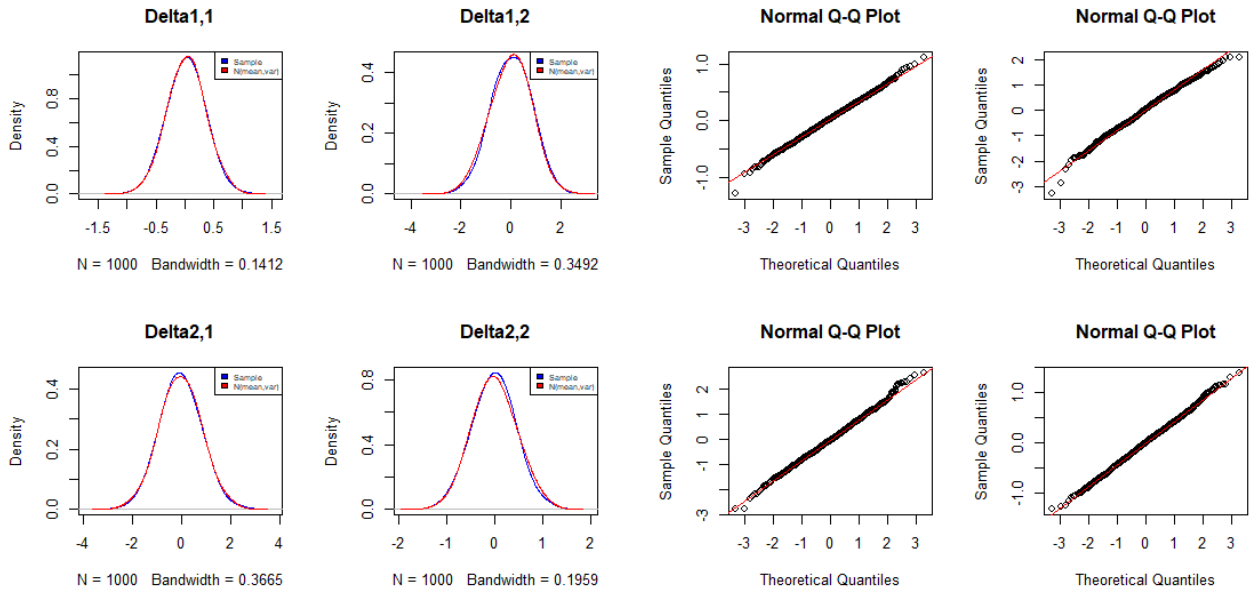


Figure 7. Theoretical and asymptotic law of $\Delta_k^{(n)}(\varphi)$ with logistic innovation and $n = 1000$.

mean temperature observed in Algeria, measured in degrees Celsius. In general, temperature readings are collected from various weather stations across Algeria, providing a representative overview of the country's climate over time. Figure 8 shows the plot of this time series.

Table 3 provides a detailed summary of the descriptive statistics for the full temperature series and for each individual month. The mean temperature exhibits a clear seasonal trend, with lower values in the winter months (January, February, December) and higher values in the summer months (June, July, August). The standard deviation is generally low, indicating limited variation within each month, with July and August being the most stable months.

Skewness and kurtosis values reveal additional patterns. For instance, positive skewness in months like January, February, and May indicates that these months occasionally experience higher-than-average temperatures. Conversely, negative skewness in September suggests a tendency toward slightly lower temperatures during this month. Kurtosis values close to zero, as seen in most months, suggest that the temperature distribution is fairly normal. However, months like August and April show more negative kurtosis, indicating a flatter distribution, meaning fewer extreme temperatures are recorded during these months.

These statistics provide insights into the seasonal dynamics of Algeria's temperature, highlighting predictable cyclical patterns, with more stability during summer and greater variability in other seasons. This understanding can be valuable for both climate study and regional policy planning.

The cyclical behavior observed in the scatter plot in Figure 9 indicates nonlinearity in the data, and the empty space in the middle suggests the presence of a limit cycle. The p-value of Keenan's test for the log-transformed temperature data is 0.0006, confirming the nonlinearity of this time series. To further validate the quality of the dataset, we performed homogeneity tests to ensure that the series does not contain significant structural breaks. The results of Pettitt's test (p-value = 0.5022) and the Buishand Range Test (p-value = 0.989) indicate that the data are homogeneous over the analyzed period. These findings confirm the dataset's suitability for modeling without requiring additional pre-processing for structural changes.

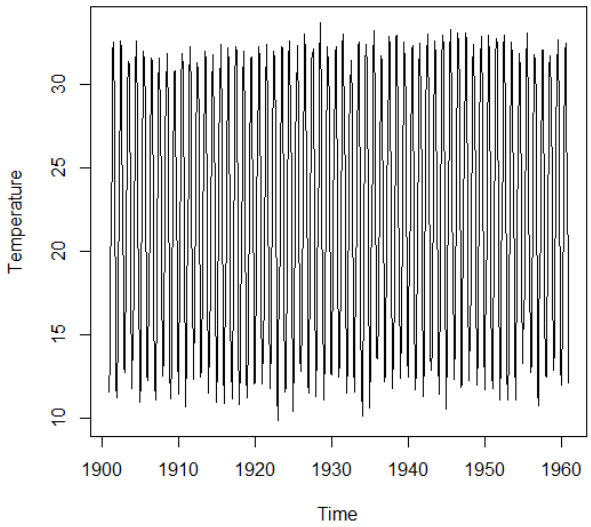


Figure 8. Algeria's temperature.

Table 3. Monthly Patterns in the Algerian Temperature Series

| Month | Mean | SD | Min | Max | Skewness | Kurtosis |
|---|---|---|---|---|---|---|
| Annual | 22.4240 | 7.2115 | 9.9069 | 33.6628 | -0.0521 | -1.4155 |
| Jan | 11.8054 | 0.8673 | 9.9069 | 14.3062 | 0.2889 | -0.1126 |
| Feb | 14.3950 | 1.0643 | 12.2803 | 16.6114 | 0.3312 | -0.5690 |
| Mar | 17.9453 | 0.9017 | 15.6129 | 20.1139 | -0.1044 | -0.0641 |
| Apr | 21.9764 | 0.8037 | 20.3271 | 23.4943 | -0.2333 | -0.8546 |
| May | 26.2224 | 0.7697 | 24.7761 | 28.3895 | 0.4362 | -0.2369 |
| Jun | 30.6230 | 0.8359 | 29.0544 | 32.8770 | 0.3042 | -0.5954 |
| Jul | 32.2668 | 0.6051 | 30.7174 | 33.6628 | -0.1507 | -0.4413 |
| Aug | 31.7176 | 0.5302 | 30.5925 | 32.6099 | -0.2634 | -1.0383 |
| Sep | 28.5553 | 0.6139 | 26.6743 | 29.6487 | -0.7878 | 0.7806 |
| Oct | 23.2166 | 0.7716 | 21.2238 | 24.8586 | -0.2723 | -0.1069 |
| Nov | 17.4501 | 0.8222 | 14.9068 | 19.3124 | -0.3028 | 0.3538 |
| Dec | 12.9144 | 0.7985 | 10.8395 | 14.6173 | -0.1078 | -0.5655 |

### *4.2. SETAR modelization*

As pointed out by a referee, we benchmarked the performance of the PSETAR model by comparing it with the classic SETAR model. Using the `tsDyn` package in R, we fitted the data to the following SETAR model:

$$X_t = \begin{cases} \underset{(0.1487)}{0.6424} + \underset{(0.0330)}{0.9758} X_{t-1} - \underset{(0.0620)}{0.1445} X_{t-2} + \epsilon_t, & \text{if } X_{t-1} \leq 2.748, \\ \underset{(0.0427)}{0.9038} + \underset{(0.0160)}{1.7332} X_{t-1} - \underset{(0.0202)}{1.0277} X_{t-2} + \epsilon_t, & \text{if } X_{t-1} > 2.748. \end{cases}$$

Here, $X_t$ represents the log-transformed data, and the values in parentheses denote the standard errors. The threshold value of 2.748 corresponds to a temperature of 15.6113°C, defining two regimes:

- **Low regime:** Colder temperatures ($X_{t-1} \leq 2.748$) with 23.68% of the observations.

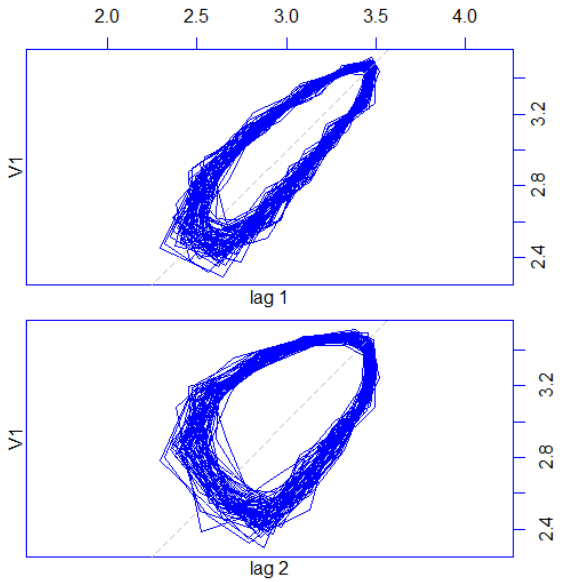


Figure 9. Lag plot of Algeria's temperature.

- **High regime:** Hotter temperatures ($X_{t-1} > 2.748$) with 76.32% of the observations.

This indicates that the higher regime occurs significantly more frequently than the lower regime. Figure 10 shows the regime-switching plot, while Figure 11 provides a filtered time series derived from the model. The root mean square error (RMSE) for this model is 0.0707, and the Akaike Information Criterion (AIC) is -3796.


Figure 10. The regime switching plot.

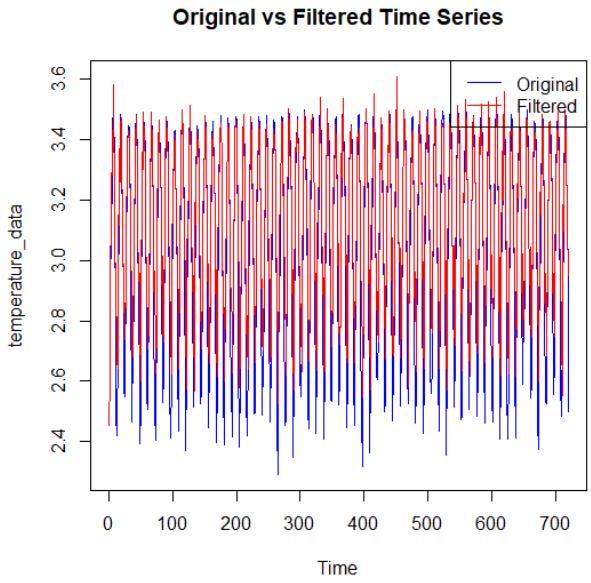


Figure 11. Simulation from the SETAR(2,2,2) model.

### *4.3. Periodic SETAR modelization*

To fit the temperature data using a periodic SETAR model, we initially selected the thresholds for each period by employing a grid search approach. The optimal thresholds were determined by minimizing the Residual Sum of Squares (RSS), as shown in Figure 12. However, the residual analysis revealed significant correlations in two specific months, suggesting that the chosen thresholds might not fully capture the seasonal dynamics of the data.

As a refinement, we opted to set the thresholds based on the seasonal mean, calculated as the average of monthly temperatures over the years. Table 4 presents a comparison between the thresholds estimated using the grid search method and those derived from the monthly means. This approach ensures that the model incorporates seasonally relevant information while addressing potential shortcomings identified during the residual analysis. Consequently, the mean was subtracted from $X_t$ to seasonally center the data.

Starting with the LR test from Section 2, we found $LR_m = 44.6118 > \chi^2_{12}(0.95) = 21.0260$. As a result, we reject the linear PAR model in favor of the PSETAR model. The parameter estimates, their $t$-statistics, and variances are presented in Table 5.

Table 4. Comparison of Threshold Values Estimated by Grid Search and Monthly Means for the PSETAR Model.

| $s$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Threshold | 2.4292 | 2.6357 | 2.9161 | 3.1095 | 3.2645 | 3.4547 | 3.4581 | 3.4558 | 3.3685 | 3.0991 | 2.9036 | 2.5725 |
| Mean | 2.4659 | 2.6642 | 2.8860 | 3.0893 | 3.2661 | 3.4213 | 3.4738 | 3.4567 | 3.3516 | 3.1443 | 2.8582 | 2.5564 |

The AIC value is -4919.129, and the global RMSE is approximately 0.0398. Considering the significance of the parameters, the PSETAR$_{12}$(2,1,1) model is represented as follows:

$$X_{12\tau+1} - 2.4659 = \begin{cases} 0.2649(X_{12\tau} - 2.4659) + \varepsilon_{12\tau+1} & \text{if } X_{12\tau} \leq 2.4659 \\ 0.3356(X_{12\tau} - 2.4659) + \varepsilon_{12\tau+1} & \text{if } X_{12\tau} > 2.4659 \end{cases}$$

$$X_{12\tau+2} - 2.6642 = \begin{cases} 0.4474(X_{12\tau+1} - 2.6642) + \varepsilon_{12\tau+2} & \text{if } X_{12\tau+1} \leq 2.6642 \\ 0.5005(X_{12\tau+1} - 2.6642) + \varepsilon_{12\tau+2} & \text{if } X_{12\tau+1} > 2.6642 \end{cases}$$

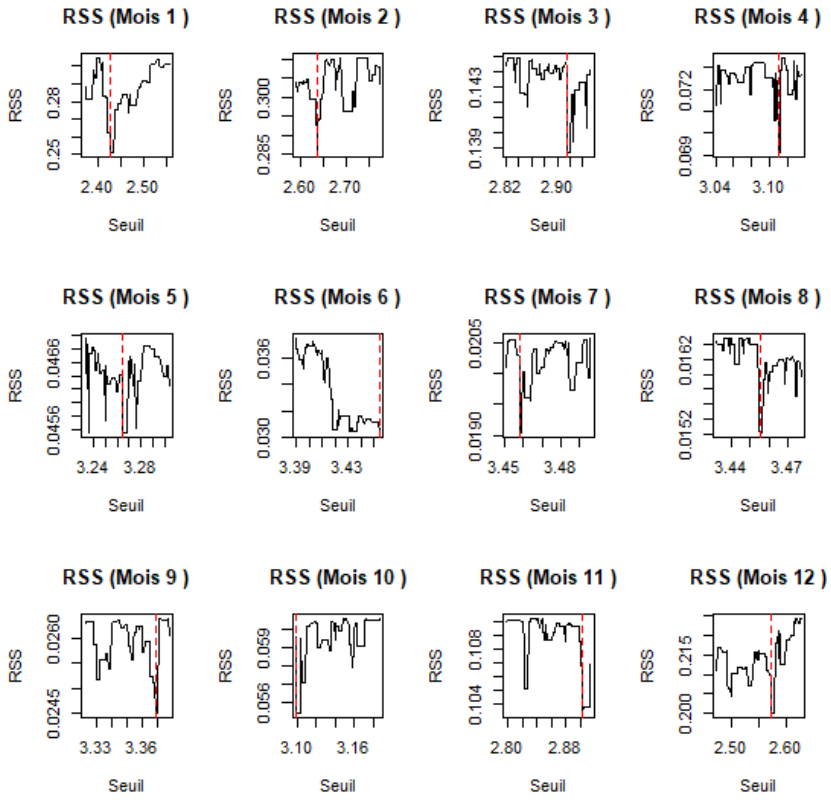


Figure 12. Optimal Threshold Selection Using RSS.

Table 5. Estimates in PSETAR$_{12}$(2,1,1).

| $s$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\phi_{s,1}$ | 0.2649 | 0.4474 | 0.3015 | 0.3491 | 0.3687 | 0.5311 | 0.4307 | 0.4619 | 0.4911 | 0.0560 | 0.8078 | 0.6567 |
| t-stat | 4.1287 | 4.6937 | 4.5204 | 3.4061 | 2.5206 | 1.7523 | 1.7054 | 1.2548 | 1.1439 | 1.5762 | 2.2604 | 3.2566 |
| $\phi_{s,2}$ | 0.3356 | 0.5005 | 0.3773 | 0.2092 | 0.4139 | 0.3677 | 0.4869 | 0.4171 | 0.4244 | 0.5748 | 0.6577 | 0.4719 |
| t-stat | 2.8289 | 3.4828 | 3.5925 | 2.2839 | 1.5858 | 1.4377 | 1.2790 | 0.8572 | 0.7448 | 0.8871 | 1.4669 | 2.1172 |
| $\sigma_s^2$ | 0.0049 | 0.0040 | 0.0018 | 0.0010 | 0.0006 | 0.0005 | 0.0001 | 0.0002 | 0.0004 | 0.0010 | 0.0016 | 0.0029 |

$$X_{12\tau+3} - 2.8860 = \begin{cases} 0.3015(X_{12\tau+2} - 2.8860) + \varepsilon_{12\tau+3} & \text{if } X_{12\tau+2} \leq 2.8860 \\ 0.3773(X_{12\tau+2} - 2.8860) + \varepsilon_{12\tau+3} & \text{if } X_{12\tau+2} > 2.8860 \end{cases}$$

$$X_{12\tau+4} - 3.0893 = \begin{cases} 0.3491(X_{12\tau+3} - 3.0893) + \varepsilon_{12\tau+4} & \text{if } X_{12\tau+3} \leq 3.0893 \\ 0.2092(X_{12\tau+3} - 3.0893) + \varepsilon_{12\tau+4} & \text{if } X_{12\tau+3} > 3.0893 \end{cases}$$

$$X_{12\tau+5} - 3.2661 = \begin{cases} 0.3687(X_{12\tau+4} - 3.2661) + \varepsilon_{12\tau+5} & \text{if } X_{12\tau+4} \leq 3.2661 \\ \varepsilon_{12\tau+5} & \text{if } X_{12\tau+4} > 3.2661 \end{cases}$$

$$X_{12\tau+6} - 3.4213 = \begin{cases} 0.5311(X_{12\tau+5} - 3.4213) + \varepsilon_{12\tau+6} & \text{if } X_{12\tau+5} \leq 3.4213 \\ \varepsilon_{12\tau+6} & \text{if } X_{12\tau+5} > 3.4213 \end{cases}$$

$$X_{12\tau+7} - 3.4738 = \begin{cases} 0.4307(X_{12\tau+6} - 3.4738) + \varepsilon_{12\tau+7} & \text{if } X_{12\tau+6} \leq 3.4738 \\ \varepsilon_{12\tau+7} & \text{if } X_{12\tau+6} > 3.4738 \end{cases}$$

$$X_{12\tau+8} - 3.4567 = \varepsilon_{12\tau+8}$$

$$X_{12\tau+9} - 3.3516 = \varepsilon_{12\tau+9}$$

$$X_{12\tau+10} - 3.1443 = \varepsilon_{12\tau+10}$$

$$X_{12\tau+11} - 2.8582 = \begin{cases} 0.8078(X_{12\tau+10} - 2.8582) + \varepsilon_{12\tau+11} & \text{if } X_{12\tau+10} \leq 2.8582 \\ \varepsilon_{12\tau+11} & \text{if } X_{12\tau+10} > 2.8582 \end{cases}$$

$$X_{12\tau+12} - 2.5564 = \begin{cases} 0.6567(X_{12\tau+11} - 2.5564) + \varepsilon_{12\tau+12} & \text{if } X_{12\tau+11} \leq 2.5564 \\ 0.4719(X_{12\tau+11} - 2.5564) + \varepsilon_{12\tau+12} & \text{if } X_{12\tau+11} > 2.5564 \end{cases}$$

**Explanation of the Model**

In this model, for each month, the log-transformed temperature $X_t$ is compared to a threshold. If the temperature is below the threshold, the model applies one set of autoregressive (AR) coefficients; if it is above the threshold, a different set is used. This structure captures the nonlinearity inherent in temperature dynamics across the year.

For example:

- In January: If the temperature exceeds the threshold (2.4659, approximately 11.77°C), a stronger autoregressive effect is observed, meaning higher persistence from past values.
- In February: The equation suggests a higher level of stability above the threshold, meaning temperatures revert more slowly to the mean.
- From June to July: The autoregressive process is strong when temperatures are below the threshold. When the temperature exceeds the threshold, no autoregressive effect is present, suggesting temperatures revert quickly without significant influence from past values.
- From August to October: There is no autoregressive process, and the temperature is purely driven by error terms, indicating more random behavior with no significant past dependence.
- In November: A strong autoregressive effect is seen below the threshold, but above the threshold, the model suggests random behavior.

The correlograms of the residuals, shown in Figure 13, validate the periodic SETAR model by demonstrating the absence of significant autocorrelation. This is further confirmed by the Box-Ljung tests (BL), where all p-values exceed 0.05, as presented in Table 6. The same table also confirms the normality of the residuals (Shapiro-Wilk test (SW)), with the exception of September, where normality is rejected. Additionally, the residuals' homoscedasticity is confirmed with Breusch-Pagan test (BP). These findings support the conclusion that the PSETAR model provides an appropriate fit for Algeria's temperature series.

| $s$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| BL | 0.9611 | 0.7137 | 0.9987 | 0.7939 | 0.9762 | 0.6142 | 0.3927 | 0.8915 | 0.9507 | 0.4667 | 0.0624 | 0.6330 |
| SW | 0.6251 | 0.7780 | 0.0349 | 0.5261 | 0.9963 | 0.4940 | 0.8315 | 0.5710 | 0.0084 | 0.4367 | 0.6549 | 0.8979 |
| BP | 0.7074 | 0.6469 | 0.7226 | 0.1765 | 0.0816 | 0.6457 | 0.2302 | 0.8776 | 0.2244 | 0.2327 | 0.3757 | 0.2646 |

Table 6. p-values of Autocorrelation, Normality, and Heteroscedasticity Tests on Residuals.

Figure 14 shows the temperature time series with distinct regimes. The red points represent periods when the temperature exceeds the threshold (indicating warmer periods), while the blue points correspond to periods

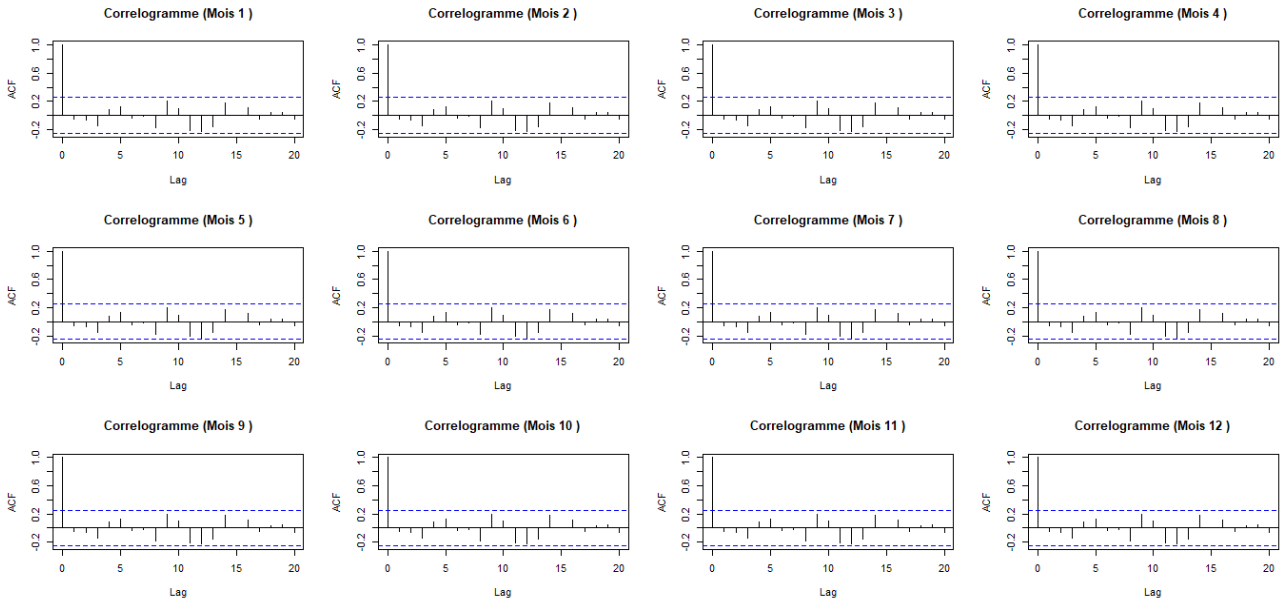


Figure 13. Correlograms of PSETAR residuals.

when the temperature is below the threshold (indicating cooler periods). In Figure 15, the threshold lines mark the monthly thresholds around which the temperature fluctuates. This figure illustrates the monthly temperature cycles, with regime-specific thresholds throughout the year, and highlights the transitions between warmer and cooler periods. It provides a clear view of how the temperature regimes evolve over the months.

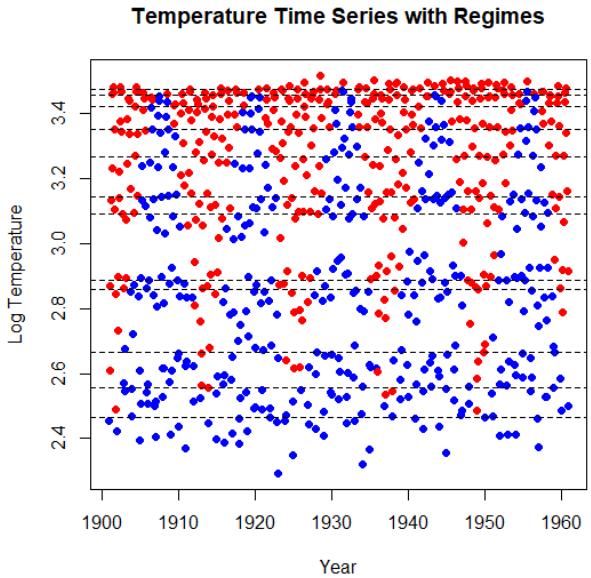


Figure 14. Temperature Time Series with Regime-Specific Thresholds: Hot vs. Cold Periods

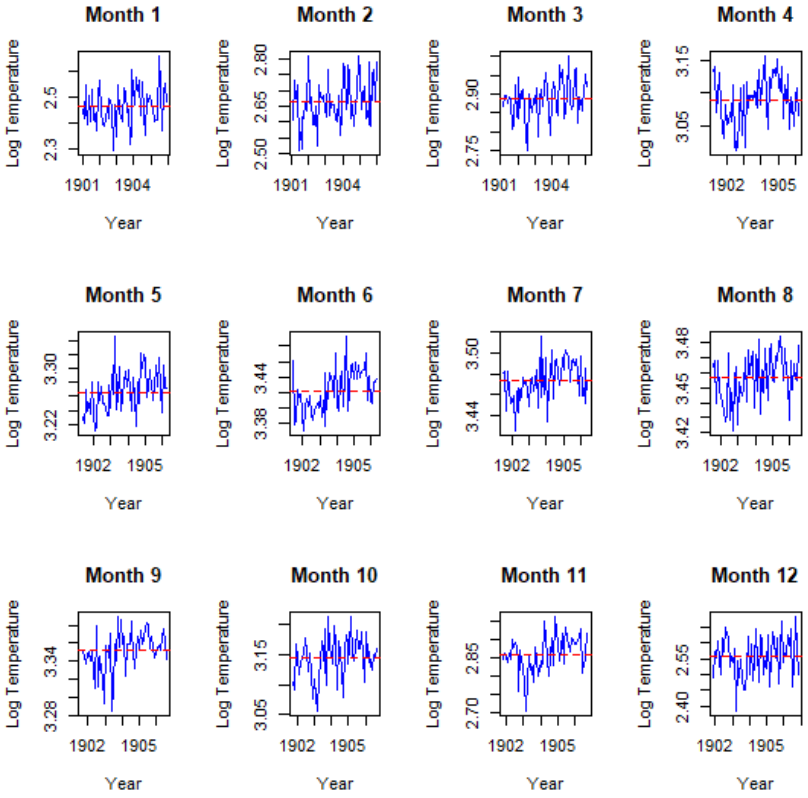


Figure 15. Monthly Temperature Cycles with Regime-Based Thresholds Across the Year.

The AIC and RMSE values for the PSETAR$_{12}$(2,1,1) model, when compared to the SETAR(2,2,2) model, indicate that the periodic model offers a better fit to the data, as it more effectively explains the variability around the fitted values. The filtered time series from the model is shown in Figure 16, while the predicted time series is presented in Figure 17.

## 5. Conclusion

In this study, we employed the periodic Self-Exciting Threshold Autoregressive model to analyze Algeria's temperature data. By incorporating periodicity, the PSETAR model demonstrated its advantage over the classical SETAR framework in capturing both seasonal variations and nonlinear dynamics. This allowed for a clearer distinction between temperature regimes during cooler and warmer months, with heightened sensitivity to past temperatures observed in warmer periods. Our analysis revealed strong nonlinear dynamics, particularly in months like June, July, and November, where the autoregressive behavior shifted significantly depending on whether the temperature exceeded the threshold. These findings align with real-world seasonal patterns, highlighting persistent trends in summer, characterized by gradual changes, and greater variability during transitional seasons such as spring and autumn.

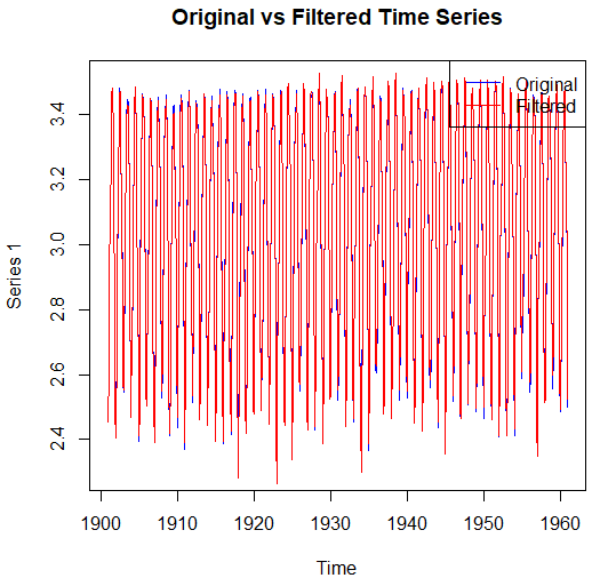


Figure 16. Simulation from PSETAR.

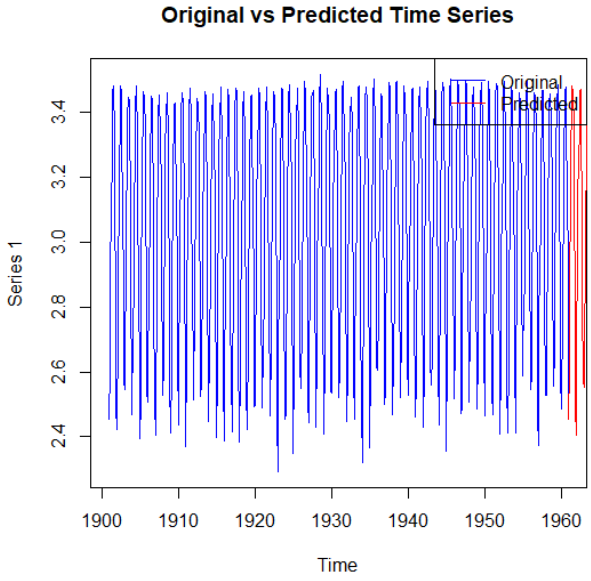


Figure 17. Predicted time series from PSETAR

## Acknowledgement

We sincerely thank the referees for their valuable comments and suggestions, which have greatly contributed to clarifying and improving the quality of this work.

## REFERENCES


1. A. Aknouche, *Periodic autoregressive stochastic volatility.*, Statistical Inference for Stochastic Processes, 20, 139-177, 2017.
2. A. Aknouche, B. Almohaimeed, and S. Dimitrakopoulos, *Periodic autoregressive conditional duration.*, Journal of Time Series Analysis, 43, 5-29, 2022.
3. A. Amimour, and K. Belaide, *Local asymptotic normality for a periodically time varying long memory parameter.* , Communications in Statistics - Theory and Methods, 51(9): 2936-2952. doi: 1080/03610926.2020.1784435, 2020.
4. P. L. Anderson, F. Sabzikar, and M. M. Meerschaert, *Parsimonious time series modeling for high frequency climate data.*, Journal of Time Series Analysis, 42. DOI:10.1111/jtsa.12579. 2020.
5. M. Bentarzi, and L. Djeddou, *Adaptive Estimation of Periodic First-Order Threshold Autoregressive Model.*, Communications in Statistics - Simulation and Computation, 43(7), 1611-1630, 2014.
6. M. Bentarzi, and M. Merzougui, *Adaptive Test for Periodicity in Self-Exciting Threshold Autoregressive Models.*, Communications in Statistics- Simulation and Computation. 38, 1592-1609, 2009.
7. Beong-Soo, *On adaptive estimation in the non-linear threshold AR(1) model.*, Perdue university. Technical report 90-45, 1991.
8. K. S. Chan, J. D. Petruccelli, H. Tong, and S. W. Woolford, *A Multiple-Threshold AR(1) Model.*, Journal of Applied Probability, Vol. 22, No. 2. pp. 267-279, 1985.
9. P. Franses, and R. Paap, *Periodic time series models.*, Oxford University Press, 2004.
10. A. Ghezal, M. Balegh, and I. Zemmouri, *Markov-switching threshold stochastic volatility models with regime changes.*, AIMS Mathematics, 9(2), 3895-3910, 2024.
11. A. Ghezal, and O. Alzeley, *Probabilistic properties and estimation methods for periodic threshold autoregressive stochastic volatility.* , AIMS Mathematics, 9(5), 11805-11832, 2024.
12. F. Hamdi, and A. Khalfi, *Predictive density criterion for SETAR models. Communications in Statistics - Simulation and Computation, 51(2), 443–459. https://doi.org/10.1080/03610918.2019.1653915*, Communications in Statistics - Simulation and Computation, 51(2), 443–459, https://doi.org/10.1080/03610918.2019.1653915, 2019.
13. A. Lama, K.N. Singh, H. Singh, et al. *Forecasting monthly rainfall of Sub-Himalayan region of India using parametric and non-parametric modelling approaches.* , Model. Earth Syst. Environ. https://doi.org/10.1007/s40808-021-01124-5, 2021.
14. P. A. W. Lewis, and B. K. Ray, *Nonlinear modelling of periodic threshold autoregressions using TSMARS. Journal of Time Series Analysis, 23(4), 459-471.* , Journal of Time Series Analysis, 23(4), 459-471, 2002.
15. S. E. Lim, M. Misiran, H. Sapiri, and Z. Md Yusof, *Forecasting the Monthly Temperature and Rainfall in Chuping Perlis.* , EDUCATUM Journal of Science, Mathematics and Technology, 11(2), 20-33. https://doi.org/10.37134/ejsmt.vol11.2.3.2024, 2024.
16. S. Regui, A. Akharif, and A. Mellouk, *Locally optimal tests against periodic linear regression in short panels.* , Communications in Statistics-Simulation and Computation, 1-15. DOI: 10.1080/03610918.2024.2314662, 2024.
17. U. Rupassara, D. Udokop, and F. Ozordi, *Time Series Analysis in Forecasting Monthly Average Rainfall and Temperature (Case Study, Minot ND, USA).* , International Journal of Data Science and Analysis. Vol. 8, No. 3, pp. 82-93. doi: 10.11648/j.ijdsa.20220803.12, 2022.
18. A. R. Swensen, *The Asymptotic Distribution of the likelihood ratio for autoregressive time series with a regression trend.* , Journal of Multivariate Analysis, 16, 54-70, 1985.
19. H. Tong, *On threshold models. In C. H. Chen (Ed.), In Pattern Recognition and Signal Processing.* , Amsterdam: Sijhoff and Noordhoff, 1978.