

Research on the AR(P) Change Point Model with Unequal Variances before and after the Sequence

Jibin Yang, Fang Yu*, Guodong Li, Shuo Li

ABSTRACT—This study investigates the change point problem in time series analysis, a prominent research topic in statistics. We propose the CP-AR(P) model, an autoregressive framework integrating change points, primarily addressing scenarios where sequence variance shifts after a structural break. Notably, a constant variance is highlighted as a special case of this framework. The methodology employs the likelihood ratio test for model formulation, with change point estimates derived via the conditional maximum likelihood method. Additionally, a symmetric matrix is introduced to provide a universal expression for parameter estimation. To further validate the proposed model's efficacy, extensive simulations using the CP-AR(2) model, in particular considering almost all potential points in the sequence as change points (reserving only marginal samples for modeling), demonstrate its efficacy, revealing that performance depends on variance disparity and sample size. Comparative analyses show a significant reduction in anomaly prediction error when considering change points, enhancing confidence interval plausibility. Thus, we recommend testing for change points before applying ARMA models to improve their fitting performance. This study presents a CP-AR(P) methodological framework and novel time series perspectives, emphasizing the role of change point identification in enhancing model performance. The framework applies to financial risk management, quality control, and environmental monitoring, among other fields.

Keywords—AR(P), Change point, Parameter estimation, Data simulations

I. INTRODUCTION

THE concepts of change point detection were initially proposed by Page and gained widespread recognition for their early application in quality control [1-2]. Subsequently, Lai [3] developed a comprehensive theory for sequential change point detection, particularly within quality-control contexts. The field of change point detection has seen extensive research and practical applications, with a focus on time series change points, especially in ARMA models with change points. The core of these studies centres on estimating change point sequences, with the maximum likelihood ratio test being a primary method in this process. Pioneering work

by Horváth [4] detailed the computation of the asymptotic distribution of the maximum likelihood ratio test for identifying parameter shifts in normal observations at an unknown point. Subsequently, Horváth [5] expanded the research to include detecting mean shifts in linear processes. Ciuperca and Maciak [6] proposed a novel quantile estimation method for change points in multiple linear regression models. Sheikhrabari and Aminnayeri [7] developed a maximum likelihood estimation technique to evaluate the change point of stationary ARMA (1,1) models transiting to a non-stationary process. Further contributions in this domain include those by Liu [8], Shiohama and Taniguchi [9], and Wang [10], among others, who discussed the maximum likelihood method in the context of ARMA models with change points. For non-normal change point sequences, Gombay [11] examined change point detection in exponential distribution families. For the non-parametric method, Keriven et al. [12] investigated the issue of finding change points in multidimensional time series. A heuristic method was proposed based on ratings to estimate change points in non-parametric processes [13]. Comprehensive insights into the asymptotic behaviour of the maximum likelihood ratio in change point issues are provided by Yao and Davis [14] and Gombay and Horváth [15]. Moreover, the practical applications of change point detection are vast and varied. For instance, Manzano Sanchez et al. [16] extracted characteristics from the power consumption of smartphones through change point analysis. This method has also proven useful in identifying significant variations in such consumption patterns [17]. Inoue and Yamada [18] detected the time point of software failure by the change point detection method, while Awe and Adepoju [19] employed this method in environmental detection. For the application of the change point problem in other fields, please refer to references [20-22]. This brief outline of relevant work provides only a cursory insight into the extensive literature on change point problems. Interested readers are encouraged to consult comprehensive reviews, such as those by Aue and Horváth [23], Jandhyala et al. [24] or Atashgar [25]. Additional contributions include Lee [26], who expounds on five prevalent types of change point problems, among other key works.

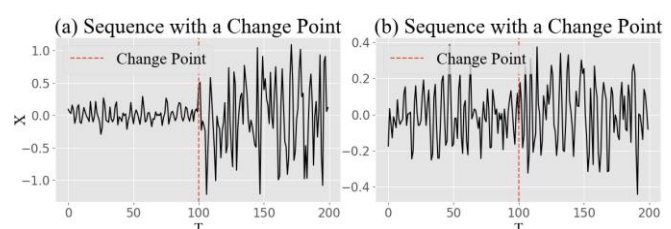


Fig. 1. Time series with change point ($T = 200$, $k=100$)

Manuscript received April 1, 2025; revised August 14, 2025.

This work was supported by the Natural Science Foundation of Xinjiang (No.2021D01C003) and Changji University Financial Big Data Key Laboratory Platform Construction Project ([2024] No. 46), and the Fund for Basic Scientific Research of Xinjiang Universities (No. XJEDU2022P116).

Jibin Yang is a lecturer at Changji University, Changji, 831100 China (email: penggar07416@163.com).

Fang Yu is a professor of Changji University, Changji, 831100 China (email: yicaizunb3741@163.com).

Guodong Li is a professor of Guilin University, Guilin, 541004 China (email: huailouka57@163.com).

Shuo Li is a professor of Changji University, Changji, 831100 China (email: nieoh8171@163.com).

First, the concept of change points and the necessity of detecting them will be introduced through illustrations. The sequences with a change point in Fig. 1 are all simulated using autoregressive models, referred to as CP-AR(P) models, which are the primary focus of this study. A change point here is defined as a moment where model parameters shift. While the approximate location of a change point can sometimes be identified directly from a time series plot, it is often challenging to confirm its existence. For instance, Fig. 1 presents two time series with a change point at position 100. In Fig. 1(a), the change point is relatively obvious: the sequence exhibits a noticeable increase in fluctuation midway through, marking the change point. In contrast, the change point in Fig. 1(b) is hard to detect, as there is no significant variation in the sequence's fluctuation. In such cases, modeling that overlooks the change point can compromise the model's performance.

Given the complexity of change point detection (e.g., non-significant variance changes and hidden structural shifts), research in this area has drawn substantial attention, and this paper aims to advance the field by addressing key gaps. Many researchers have discussed the invariance of variance before and after sequence change points. This paper further addresses the position of model change points, the estimation of model parameters, and the variance before and after a change point in cases of variance mutation. Notably, a key innovation lies in its simulation design. Unlike most existing simulation studies that typically consider only one or a few selected points, commonly the midpoint, as change points to demonstrate the effectiveness of the proposed methods, our study takes almost all points in the sequence into account as potential change points, reserving only a small portion of samples at the beginning and end of the sequence for modeling. This study aims to develop a method for detecting change points that can accurately identify change points in time series data and is effective for both structural changes in variance and other types of structural breaks. Therefore, we propose the CP-AR (P) model, designed to analyze time series data by integrating an autoregressive framework with a change point detection mechanism. In terms of methodology, the change point problem is transformed into a hypothesis test problem, and the likelihood ratio test is used to estimate the change point k . Additionally, a symmetric matrix is introduced to provide estimation expressions for variance and other parameters. This method, with a solid theoretical foundation, also exhibits broad practical applicability. To further validate its effectiveness and reliability, we conducted extensive simulation experiments. Specifically, using the CP-AR(2) model, we executed multiple simulations at varying positions within the sequence's change points to investigate how variance levels and sample sizes influence outcomes. We achieved notable results by comparing model errors with and without considering change points, confirming the method's superiority. The key finding is that the model's change point detection method accurately identifies change points in time series data, even in the presence of variance mutations. This insight is crucial for maintaining the model's accuracy and adaptability, particularly for extended time series data. Finally, the study's significance is multifaceted. Firstly, the CP-AR (P) model's change point detection approach offers a new theoretical

perspective and tool for time series analysis, enriching the research in this domain. Secondly, it holds broad practical application prospects, particularly in areas such as quality control, environmental monitoring, and consumer behavior detection, by ensuring model accuracy, adaptability, and efficiency in data processing. Thus, the findings of this study not only hold theoretical value but also have significant practical implications.

II. CP-AR(P) MODEL

For a noncentralized stationary time series,

$$\mu = \frac{\varphi_0}{1 - \varphi_1 - L - \varphi_p}$$

shifts it by a constant to convert it into a centralized sequence. Assuming that b satisfies the p -order autoregressive model, and if a delay operator is introduced, then this centered model can be expressed as: $\Phi(B)x_t = \xi_t$, where $\Phi(B) = 1 - \varphi_1 B^1 - \varphi_2 B^2 - L - \varphi_p B^p$, usually, we assume $\xi_t \sim \text{i.i.d.} \mathcal{WN}(0, \delta^2)$.

For a practical problem, the sequence may be affected by certain factors, resulting in a sudden change at time k . Therefore, it is divided into two sequences, $\{x_t^{(1)}\}$ and then $\{x_t\} = \{x_t^{(1)}, x_t^{(2)}\}$, a new model is obtained as follows:

$$X_t = \begin{cases} \varphi_1^{(1)} X_{t-1} + \varphi_2^{(1)} X_{t-2} + L + \varphi_p^{(1)} X_{t-p} + \xi_t^{(1)} & t = p+1, L, k \\ \varphi_1^{(2)} X_{t-1} + \varphi_2^{(2)} X_{t-2} + L + \varphi_p^{(2)} X_{t-p} + \xi_t^{(2)} & t = k+1, L, T \end{cases} \quad (1)$$

This is the CP-AR(P) model, where T is the length of the sequence, and k is the location of change point. The sequence before the change point is described by the first equation in the model, and the second equation describes the later part of the sequence. Here, an abrupt parameter changes means parameter values shift while the model structure and parameter count remain unchanged. If a delay operator is introduced, the above equation can be expressed as:

$$\begin{cases} \Phi^{(1)}(B)x_t^{(1)} = \xi_t^{(1)} & t = 3, L, k \\ \Phi^{(2)}(B)x_t^{(2)} = \xi_t^{(2)} & t = k+1, L, T \end{cases} \quad (2)$$

III. SEARCHING FOR CHANGE POINTS

A. $\sigma^{(1)} \neq \sigma^{(2)}$

Horváth [4] and Gombay [11] introduced the likelihood ratio method to find the location of the change point. Here, a similar approach is used to model time series. Therefore, the CP-AR(P) model can also be considered as a hypothesis testing problem for the following:

$$\begin{aligned} H_0 : X_t &= \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + L + \varphi_p X_{t-p} + \xi_t \quad p+1 \leq t \leq T \\ H_1 : X_t &= \begin{cases} \varphi_1^{(1)} X_{t-1} + \varphi_2^{(1)} X_{t-2} + L + \varphi_p^{(1)} X_{t-p} + \xi_t^{(1)} & p+1 \leq t \leq k \\ \varphi_1^{(2)} X_{t-1} + \varphi_2^{(2)} X_{t-2} + L + \varphi_p^{(2)} X_{t-p} + \xi_t^{(2)} & k+1 \leq t \leq T \end{cases} \quad (3) \end{aligned}$$

When H_0 is true, given $\{X_t\}$, denote $\bar{\theta}_0 = (\varphi_1, \varphi_2, L, \varphi_p, \sigma')'$ as the parameter vector. Since $\{\xi_t\}$ is assumed to be a Gaussian process, X_t is also a Gaussian process. When $t \geq 2$, X_t is also a Gaussian process given in the previous sequence $\{X_{t-1}\}$. Therefore, when $T > P$, the conditional logarithmic likelihood function is:

$$\begin{aligned} & \log f_{X_T, X_{T-1}, \dots, X_{p+1}} | X_p, \dots, X_1 (x_T, x_{T-1}, \dots, x_{p+1} | x_p, \dots, x_1; \theta) \\ &= -\frac{T-p}{2} \log(2\pi) - \frac{T-p}{2} \log(\sigma^2) \\ & - \frac{1}{2\sigma^2} \sum_{t=p+1}^T (X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - L - \phi_p X_{t-p})^2 \end{aligned} \quad (4)$$

Differentiating $\phi_1, \phi_2, L, \phi_p$ will result in a complex expression. So, to estimate these parameters, a symmetric matrix is introduced as below:

$$\text{Let } Z = \begin{bmatrix} \sum X_{t-1}^2 & \sum X_{t-1}X_{t-2} & L & \sum X_{t-1}X_{t-p} \\ \sum X_{t-2}X_{t-1} & \sum X_{t-2}^2 & L & \sum X_{t-2}X_{t-p} \\ M & M & O & M \\ \sum X_{t-p}X_{t-1} & \sum X_{t-p}X_{t-2} & L & \sum X_{t-p}^2 \end{bmatrix} \quad (5)$$

$$\Phi = \begin{pmatrix} \phi_1 \\ \phi_2 \\ M \\ \phi_p \end{pmatrix}, B = \begin{pmatrix} \sum X_t X_{t-1} \\ \sum X_t X_{t-2} \\ M \\ \sum X_t X_{t-p} \end{pmatrix} \quad (6)$$

Then, $Z\Phi = B$ is derived, which is:

$$\begin{aligned} & \begin{bmatrix} \sum X_{t-1}^2 & \sum X_{t-1}X_{t-2} & L & \sum X_{t-1}X_{t-p} \\ \sum X_{t-2}X_{t-1} & \sum X_{t-2}^2 & L & \sum X_{t-2}X_{t-p} \\ M & M & O & M \\ \sum X_{t-p}X_{t-1} & \sum X_{t-p}X_{t-2} & L & \sum X_{t-p}^2 \end{bmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ M \\ \phi_p \end{pmatrix} \\ &= \begin{pmatrix} \sum X_t X_{t-1} \\ \sum X_t X_{t-2} \\ M \\ \sum X_t X_{t-p} \end{pmatrix} \end{aligned} \quad (7)$$

Subsequently, assuming Z is invertible, $\Phi = Z^{-1}B$ can be obtained. Solving this system of linear equations yields the maximum likelihood estimate $\hat{\Phi} = (\hat{\phi}_1, \hat{\phi}_2, L, \hat{\phi}_p)$ of Φ . The estimation of σ^2 obtained using the same method is as follows:

$$\hat{\sigma}^2 = \frac{1}{T-p} \left[\sum_{t=p+1}^T (X_t - \hat{\phi}_1 X_{t-1} - \hat{\phi}_2 X_{t-2} - L - \hat{\phi}_p X_{t-p})^2 \right] \quad (8)$$

Under the assumption that H_1 is true, $\theta_1 = (\phi_1, \phi_2, L, \phi_p, \sigma^{(1)}, \sigma^{(2)})'$, the logarithmic likelihood function of the change point problem is:

$$\begin{aligned} & \log f_{X_T, X_{T-1}, \dots, X_{p+1}} | X_p, \dots, X_1 (x_T, x_{T-1}, \dots, x_{p+1} | x_p, \dots, x_1; \theta) \\ &= -\frac{T-1}{2} \log(2\pi) - \left(\frac{k-p}{2} \right) \log[\sigma^{(1)}]^2 - \left(\frac{T-k-p}{2} \right) \log[\sigma^{(2)}]^2 \\ & - \frac{1}{2[\sigma^{(1)}]^2} \sum_{t=p+1}^k (X_t - \phi_1^{(1)} X_{t-1} - \phi_2^{(1)} X_{t-2} - L - \phi_p^{(1)} X_{t-p})^2 \\ & - \frac{1}{2[\sigma^{(2)}]^2} \sum_{t=k+1}^T (X_t - \phi_1^{(2)} X_{t-1} - \phi_2^{(2)} X_{t-2} - L - \phi_p^{(2)} X_{t-p})^2 \end{aligned} \quad (9)$$

In this case, differentiating $\phi_1^{(1)}, \phi_2^{(1)}, L, \phi_p^{(1)}$ is only related to the previous sequence, while differentiating $\phi_1^{(2)}, \phi_2^{(2)}, L, \phi_p^{(2)}$ is only related to the subsequent sequence. Therefore, the sequence X is divided into two sequences U and V by the change point, so $X = \{U, V\}$, where $U = \{X_{p+1}, \dots, X_k\}$ is the sequence before the change point, and $V = \{X_{k+1}, \dots, X_T\}$ is the sequence after the change point. Then, similarly to the previous case, introduce the following symmetric matrix:

$$Z^{(1)} = \begin{bmatrix} \sum U_{t-1}^2 & \sum U_{t-1}U_{t-2} & L & \sum U_{t-1}U_{t-p} \\ \sum U_{t-2}U_{t-1} & \sum U_{t-2}^2 & L & \sum U_{t-2}U_{t-p} \\ M & M & O & M \\ \sum U_{t-p}U_{t-1} & \sum U_{t-p}U_{t-2} & L & \sum U_{t-p}^2 \end{bmatrix} \quad (10)$$

$$\Phi^{(1)} = \begin{pmatrix} \phi_1^{(1)} \\ \phi_2^{(1)} \\ M \\ \phi_p^{(1)} \end{pmatrix}, B^{(1)} = \begin{pmatrix} \sum U_t U_{t-1} \\ \sum U_t U_{t-2} \\ M \\ \sum U_t U_{t-p} \end{pmatrix} \quad (11)$$

Subsequently $\Phi^{(1)} = [Z^{(1)}]^{-1} B^{(1)}$, and solving this linear system of equations obtains the maximum likelihood estimate $\hat{\Phi}^{(1)} = (\hat{\phi}_1^{(1)}, \hat{\phi}_2^{(1)}, L, \hat{\phi}_p^{(1)})$. The maximum likelihood estimation of $[\hat{\sigma}^{(1)}]^2, [\hat{\sigma}^{(2)}]^2$ is:

$$[\hat{\sigma}^{(1)}]^2 = \frac{1}{k-p} \left[\sum_{t=p+1}^k (X_t - \hat{\phi}_1^{(1)} X_{t-1} - L - \hat{\phi}_p^{(1)} X_{t-p})^2 \right] \quad (12)$$

$$[\hat{\sigma}^{(2)}]^2 = \frac{1}{T-k-p} \left[\sum_{t=k+1}^T (X_t - \hat{\phi}_1^{(2)} X_{t-1} - L - \hat{\phi}_p^{(2)} X_{t-p})^2 \right] \quad (13)$$

The maximum likelihood ratio function is:

$$\begin{aligned} \lambda_k &= -\frac{\sum_{t=1}^T f(X_t; \hat{\phi}_1, \hat{\phi}_2, L, \hat{\phi}_p, \hat{\sigma}_T)}{\prod_{i=1}^k f(X_t; \hat{\phi}_1^{(1)}, \hat{\phi}_2^{(1)}, L, \hat{\phi}_p^{(1)}, \hat{\sigma}^{(1)}) \prod_{t=k+1}^T f(X_t; \hat{\phi}_1^{(2)}, \hat{\phi}_2^{(2)}, L, \hat{\phi}_p^{(2)}, \hat{\sigma}^{(2)})} \\ &= \frac{([\hat{\sigma}^{(1)}]^2)^{\frac{k-1}{2}} ([\hat{\sigma}^{(2)}]^2)^{\frac{T-k}{2}}}{(\hat{\sigma}_T^2)^{\frac{n-1}{2}}} \end{aligned} \quad (14)$$

Then:

$$\begin{aligned} \Lambda_k &= -2 \log(\lambda_k) = (T-1) \log(\hat{\sigma}_T^2) - (k-1) \log([\hat{\sigma}^{(1)}]^2) \\ & - (T-k) \log([\hat{\sigma}^{(2)}]^2) \end{aligned} \quad (15)$$

The maximum likelihood estimate of the change point k is:

$$\hat{k} = \arg \max_{p+1 \leq k \leq T-p-1} |\Lambda_k| \quad (16)$$

$$B, \sigma^{(1)} = \sigma^{(2)}$$

When the variance of the sequence before and after the change point is equal, the CP-AR (P) model can also be considered as a hypothesis testing problem for the following:

$$\begin{aligned}
 H_0 : X_t &= \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + L + \varphi_p X_{t-p} + \xi_t \quad p+1 \leq t \leq T \\
 H_1 : X_t &= \begin{cases} \varphi_1^{(1)} X_{t-1} + \varphi_2^{(1)} X_{t-2} + L + \varphi_p^{(1)} X_{t-p} + \xi_t & p+1 \leq t \leq k \\ \varphi_1^{(2)} X_{t-1} + \varphi_2^{(2)} X_{t-2} + L + \varphi_p^{(2)} X_{t-p} + \xi_t & k+1 \leq t \leq T \end{cases}
 \end{aligned} \quad (17)$$

When H_0 is true, the estimation results for each parameter can be found in (3), (4), and (8). Under the true assumption, $\theta_1 = (\varphi_1^{(1)}, \varphi_2^{(1)}, L, \varphi_p^{(1)}, \varphi_1^{(2)}, \varphi_2^{(2)}, L, \varphi_p^{(2)}, \sigma_k)^T$ the conditional logarithmic likelihood function is:

$$\begin{aligned}
 \log f''(X_T, X_{T-1}, \dots, X_{p+1} | X_p, \dots, X_1) &= \log f''(X_T, X_{T-1}, \dots, X_{p+1} | X_p, \dots, X_1; \theta) \\
 &= -\frac{T-1}{2} \log(2\pi) - \left(\frac{T-p}{2} \right) \log \sigma_k^2 - \frac{1}{2\sigma_k^2} \sum_{t=p+1}^T (X_t - \varphi_1^{(1)} X_{t-1} - \varphi_2^{(1)} X_{t-2} \\
 &\quad - L - \varphi_p^{(1)} X_{t-p})^2 - \frac{1}{2\sigma_k^2} \sum_{t=k+1}^T (X_t - \varphi_1^{(2)} X_{t-1} - \varphi_2^{(2)} X_{t-2} - L - \varphi_p^{(2)} X_{t-p})^2
 \end{aligned} \quad (18)$$

The maximum likelihood estimation of σ^2 is:

$$\begin{aligned}
 \hat{\sigma}_k^2 &= \frac{1}{T-p} \left[\sum_{t=p+1}^T (X_t - \hat{\varphi}_1^{(1)} X_{t-1} - \hat{\varphi}_2^{(1)} X_{t-2} - L - \hat{\varphi}_p^{(1)} X_{t-p})^2 \right. \\
 &\quad \left. + \sum_{t=k+1}^T (X_t - \hat{\varphi}_1^{(1)} X_{t-1} - \hat{\varphi}_2^{(1)} X_{t-2} - L - \hat{\varphi}_p^{(1)} X_{t-p})^2 \right]
 \end{aligned} \quad (19)$$

Estimations of $\hat{\Phi}^{(1)}$ and $\hat{\Phi}^{(2)}$ are in the previous section, so the maximum likelihood ratio function is:

$$\begin{aligned}
 \lambda_k &= -\frac{\prod_{t=1}^T f(X_t; \hat{\varphi}_1, \hat{\varphi}_2, L, \hat{\varphi}_p, \hat{\sigma})}{\prod_{t=1}^k f(X_t; \hat{\varphi}_1^{(1)}, \hat{\varphi}_2^{(1)}, L, \hat{\varphi}_p^{(1)}, \hat{\sigma}_k) \prod_{t=k+1}^T f(X_t; \hat{\varphi}_1^{(2)}, \hat{\varphi}_2^{(2)}, L, \hat{\varphi}_p^{(2)}, \hat{\sigma}_k)} \\
 &= (\hat{\sigma}_k^2 / \hat{\sigma}^2)^{\frac{T-k}{2}}
 \end{aligned} \quad (20)$$

Then, the maximum likelihood estimate of the change point k is:

$$\hat{k} = \arg \max_{p+1 \leq k \leq T-p-1} |\lambda_k| \quad (21)$$

IV. SIMULATION RESEARCH

To further verify and study the above conclusions, we take the 2nd-order model as an example for simulation. According to (3), the CP-AR(2) can also be transformed into a hypothesis testing problem, as shown below:

$$\begin{aligned}
 H_0 : X_t &= \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \xi_t \quad 3 \leq t \leq T \\
 H_1 : X_t &= \begin{cases} \varphi_1^{(1)} X_{t-1} + \varphi_2^{(1)} X_{t-2} + \xi_t^{(1)} & 3 \leq t \leq k \\ \varphi_1^{(2)} X_{t-1} + \varphi_2^{(2)} X_{t-2} + \xi_t^{(2)} & k+1 \leq t \leq T \end{cases}
 \end{aligned} \quad (22)$$

When H_0 is true, $\theta_0 = (\varphi_1, \varphi_2, \sigma)^T$, according to the method in section 3, the estimation expression for each parameter is as follows:

$$\hat{\varphi}_1 = \frac{\sum_{t=3}^T X_{t-2} \sum_{t=3}^T X_{t-1} X_t - \sum_{t=3}^T X_{t-2} X_{t-1} \sum_{t=3}^T X_{t-2} X_t}{\sum_{t=3}^T X_{t-2}^2 \sum_{t=3}^T X_{t-1}^2 - \sum_{t=3}^T (X_{t-2} X_{t-1})^2} \quad (23)$$

$$\hat{\varphi}_2 = \frac{\sum_{t=3}^T X_{t-1}^2 \sum_{t=3}^T X_{t-2} X_t - \sum_{t=3}^T X_{t-2} X_{t-1} \sum_{t=3}^T X_{t-1} X_t}{\sum_{t=3}^T X_{t-2}^2 \sum_{t=3}^T X_{t-1}^2 - \sum_{t=3}^T (X_{t-2} X_{t-1})^2} \quad (24)$$

$$\hat{\sigma}^2 = \frac{1}{T-2} \left[\sum_{t=3}^T (X_t - \hat{\varphi}_1 X_{t-1} - \hat{\varphi}_2 X_{t-2})^2 \right] \quad (25)$$

When H_1 is true, $\theta_1 = (\varphi_1^{(1)}, \varphi_2^{(1)}, \varphi_1^{(2)}, \varphi_2^{(2)}, \sigma^{(1)}, \sigma^{(2)})^T$, according to the method in section 3, the estimation expression for each parameter is as follows:

$$\hat{\varphi}_1^{(1)} = \frac{\sum_{t=3}^k X_{t-2} \sum_{t=3}^k X_{t-1} X_t - \sum_{t=3}^k X_{t-2} X_{t-1} \sum_{t=3}^k X_{t-2} X_t}{\sum_{t=3}^k X_{t-2}^2 \sum_{t=3}^k X_{t-1}^2 - \sum_{t=3}^k (X_{t-2} X_{t-1})^2} \quad (26)$$

$$\hat{\varphi}_2^{(1)} = \frac{\sum_{t=3}^k X_{t-1}^2 \sum_{t=3}^k X_{t-2} X_t - \sum_{t=3}^k X_{t-2} X_{t-1} \sum_{t=3}^k X_{t-1} X_t}{\sum_{t=3}^k X_{t-2}^2 \sum_{t=3}^k X_{t-1}^2 - \sum_{t=3}^k (X_{t-2} X_{t-1})^2} \quad (27)$$

$$\hat{\varphi}_1^{(2)} = \frac{\sum_{t=k+1}^T X_{t-2} \sum_{t=k+1}^T X_{t-1} X_t - \sum_{t=k+1}^T X_{t-2} X_{t-1} \sum_{t=k+1}^T X_{t-2} X_t}{\sum_{t=k+1}^T X_{t-2}^2 \sum_{t=k+1}^T X_{t-1}^2 - \sum_{t=k+1}^T (X_{t-2} X_{t-1})^2} \quad (28)$$

$$\hat{\varphi}_2^{(2)} = \frac{\sum_{t=k+1}^T X_{t-1}^2 \sum_{t=k+1}^T X_{t-2} X_t - \sum_{t=k+1}^T X_{t-2} X_{t-1} \sum_{t=k+1}^T X_{t-1} X_t}{\sum_{t=k+1}^T X_{t-2}^2 \sum_{t=k+1}^T X_{t-1}^2 - \sum_{t=k+1}^T (X_{t-2} X_{t-1})^2} \quad (29)$$

$$[\hat{\sigma}^{(1)}]^2 = \frac{1}{k-2} \left[\sum_{t=3}^k (X_t - \hat{\varphi}_1^{(1)} X_{t-1} - \hat{\varphi}_2^{(1)} X_{t-2})^2 \right] \quad (30)$$

$$[\hat{\sigma}^{(2)}]^2 = \frac{1}{T-k-2} \left[\sum_{t=3}^{k+1} (X_t - \hat{\varphi}_1^{(2)} X_{t-1} - \hat{\varphi}_2^{(2)} X_{t-2})^2 \right] \quad (31)$$

Estimation details of $\hat{k}$ are provided in (14) to (16), which are not repeated here. For example, for the following CP-AR(2) model:

$$X_t = \begin{cases} 0.5X_{t-1} - 0.7X_{t-2} + \xi_t^{(1)} & 3 \leq t \leq k \\ 0.6X_{t-1} - 0.4X_{t-2} + \xi_t^{(2)} & k+1 \leq t \leq T \end{cases} \quad (32)$$

First, a sample sequence will be generated based on this model. Fig. 1 shows a sequence of $T=200$ and $k=100$. When the sample size is increased to 600, the estimated k is shown

in Fig. 2.

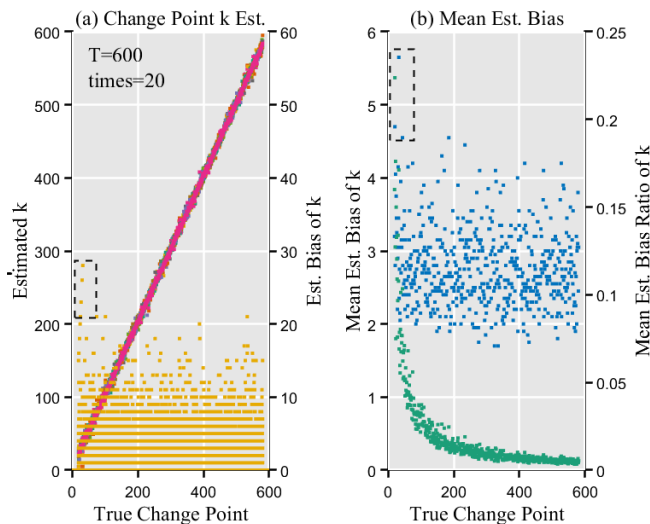


Fig. 2. The estimation of Change point k and bias. Note that $T = 600$, $\phi_1^{(1)} = 0.5$, $\phi_2^{(1)} = -0.7$, $\phi_1^{(2)} = 0.6$, $\phi_2^{(2)} = -0.4$, $D^{(1)} = 0.1$, $D^{(2)} = 0.5$ (where D represent variance), the number of simulations times=20 for each k . For Fig. 2(a), the abscissa is $t(k)$, the ordinate is $\hat{k}$, the secondary axis is the estimated deviation of k ; that is:

$$e_k = |k - \hat{k}|, \quad k \in (20, \dots, 580]$$

for Fig. 2(b), the ordinate axis is the average estimation deviation of k , that is:

$$\bar{e}_k = \frac{1}{n} \sum_{n=1}^n e_k, \quad k \in (20, \dots, 580], n = 600$$

and the secondary axis is the ratio of the average deviation to the estimated value of k , that is:

$$r_k = \bar{e}_k / k, \quad k \in (20, \dots, 580]$$

Figure 2 shows the results of estimating the inflection points in sequence X_t using the method introduced above, as well as the estimation bias. Like the Q-Q plot for determining normal distribution, the better the estimation, the more concentrated the points in the plot are on a slope-1 straight line. Here, the sample size T is 600. To ensure that the sample size of the preceding and succeeding sequence segments is at least 20, the actual position of the change point, k , ranges from 20 to 580. With 20 simulations performed for each of these 560 values of k (times = 20), a total of $20 \times 560 = 11,200$ simulations are conducted. In Fig. 2 (a), the primary axis (left) represents the predicted change point position, and nearly all estimated values fall near the straight line with a slope of 1 and pass through the origin. The secondary coordinate axis of Fig. 2(a) is the absolute deviation of k , and Fig. 2(b) is a supplement to Fig. 2(a), mainly showing the predicted deviation of k . The primary axis (left) represents the average absolute deviation of the change point. Most values range between 1 and 4, indicating small and uniformly distributed deviations. The secondary axis (right) shows the ratio of the average absolute deviation to $\hat{k}$, which tends to decrease as the sample size increases. This trend indicates improved prediction accuracy with larger sample sizes (as shown in the bottom part of Fig. 2(b)). For a more detailed breakdown of these results, selected data points are presented in Table I.

TABLE I
MEAN EST.BIAS AND BIAS RATIO OF K

k	21	30	50	100	200	300
r_k	12.73%	9.67%	5.10%	3.25%	1.25%	0.75%
$\bar{e}_k$	4.05	2.9	2.6	3.25	2.5	2.25

As shown in Table I, when the sample size reaches 30, the average absolute deviation is small and relatively stable, supporting its recommendation as the minimum sample size for modeling. A sample size of 50 or more is even better, as it further reduces deviations. Additionally, from Fig. 2, it is noted that there are many cases of large deviations near the starting and ending positions (see the dashed boxes in Fig. 2(a) and Fig. 2(b)). This is caused by an excessively forward change point position, which leads to an insufficient sample size. Although Table I shows that the average deviation is not much larger, the stability is obviously poor.

The above analysis indicates that a larger sample size corresponds to a longer sequence. Apart from the influence of sequence length (i.e., sample size) on the stability of estimation bias, are there other contributing factors? To address this, further research will explore the impact of different variance differences between pre- and post-change point sequences.

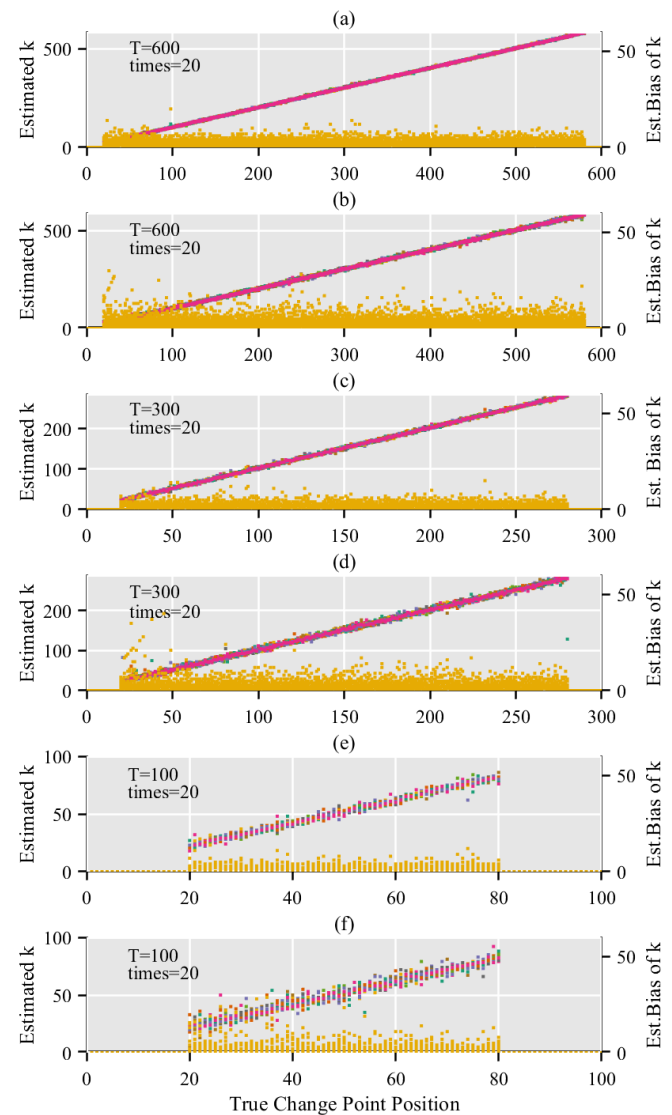


Fig. 3. Change point k estimation with different sample size and variance. Note that in (a), (c), and (d), $D^{(1)} = 0.1$, $D^{(2)} = 0.5$, and the sample size T is 600, 300, and 100, respectively. In (b), (d), and (f), $D^{(1)} = 0.1$, $D^{(2)} = 0.3$, and the sample size T is 600, 300, and 100, respectively. The number of simulation times for each position k is 20.

The sub-diagram in Fig. 3 is like Fig. 2. The horizontal axis represents the position of the change point, the vertical axis represents $\hat{k}$, and the secondary axis is the change point estimated deviation. It further illustrates scenarios with different samples and also displays simulation results with different variances. Based on variance differences, it can be divided into two groups: Fig. 3(a), Fig. 3(c), and Fig. 3(e), and Fig. 3(b), Fig. 3(d), and Fig. 3(f); based on different sample sizes, it can be divided into three groups: Fig. 3(a) and Fig. 3(b), Fig. 3(c) and Fig. 3(d), Fig. 3(e) and Fig. 3(f). For the sample grouping, as the sample size decreases from 600

to 100, the estimates of the change points are more widely scattered, especially when $T=100$, with a consequent decrease in prediction accuracy. This implies that more points deviate from the line with a slope of 1 in the graph. For the grouping based on variance, comparing Fig. 3(a) and Fig. 3(b), Fig. 3(c) and Fig. 3(d), Fig. 3(e) and Fig. 3(f) when the difference in the variance between the preceding and following sequences becomes smaller, the range of deviations increases and becomes more unstable, so the accuracy of change point estimation is affected by the difference in variance between the preceding and following sequences. When the variance difference is small enough, it can be assumed that it is constant during structural changes. In this case, the method described in Section 3 can be used for estimation.

TABLE II
PARAMETER ESTIMATION RESULTS ($T = 600, k = 300, [\sigma^{(1)}]^2 = 0.1, [\sigma^{(2)}]^2 = 0.3$)

Not considering C.P. ($\hat{\theta}_0$)					Considering C.P. ($\hat{\theta}_1$)				
$\hat{k}_1$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\sigma}^2$	$\hat{\phi}_1^{(1)}$	$\hat{\phi}_2^{(1)}$	$\hat{\phi}_1^{(2)}$	$\hat{\phi}_2^{(2)}$	$[\hat{\sigma}^{(1)}]^2$	$[\hat{\sigma}^{(2)}]^2$
296	0.5148	-0.3556	0.2144	0.4790	-0.6383	0.5111	-0.3039	0.1046	0.2790
303	0.5078	-0.3829	0.2338	0.5339	-0.6326	0.5004	-0.3435	0.1001	0.3142
306	0.5459	-0.4175	0.2256	0.5093	-0.6216	0.5466	-0.3900	0.0994	0.3043
302	0.4904	-0.3789	0.2145	0.4154	-0.6717	0.4925	-0.3289	0.0989	0.2835
304	0.4602	-0.3966	0.2261	0.4741	-0.6559	0.4539	-0.3474	0.1030	0.3017
302	0.4131	-0.3111	0.2132	0.5051	-0.6132	0.3921	-0.2438	0.1007	0.2809
303	0.5310	-0.3539	0.2264	0.4572	-0.6299	0.5308	-0.3187	0.0991	0.3025
301	0.4893	-0.3108	0.2357	0.4905	-0.6127	0.4824	-0.2687	0.1012	0.3149
300	0.4989	-0.3486	0.2251	0.4913	-0.6386	0.4932	-0.3041	0.0981	0.2996
307	0.5385	-0.4067	0.2274	0.4958	-0.6212	0.5391	-0.3814	0.0944	0.3092

TABLE III
PARAMETER ESTIMATION RESULTS ($T = 600, k = 300, [\sigma^{(1)}]^2 = 0.1, [\sigma^{(2)}]^2 = 0.15$)

Not considering C.P. ($\hat{\theta}_0$)					Considering C.P. ($\hat{\theta}_1$)				
$\hat{k}_1$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\sigma}^2$	$\hat{\phi}_1^{(1)}$	$\hat{\phi}_2^{(1)}$	$\hat{\phi}_1^{(2)}$	$\hat{\phi}_2^{(2)}$	$[\hat{\sigma}^{(1)}]^2$	$[\hat{\sigma}^{(2)}]^2$
299	0.4212	-0.4443	0.1319	0.3564	-0.6271	0.4404	-0.3561	0.0978	0.1556
302	0.4290	-0.4222	0.1281	0.3480	-0.5177	0.4589	-0.3836	0.0992	0.1509
299	0.4078	-0.4775	0.1271	0.4137	-0.6116	0.4022	-0.4026	0.0973	0.1497
302	0.4370	-0.4429	0.1255	0.4426	-0.5895	0.4271	-0.3367	0.1033	0.1421
313	0.4284	-0.4219	0.1279	0.3216	-0.5804	0.4626	-0.3573	0.0936	0.1542
318	0.4274	-0.3840	0.1245	0.4110	-0.5100	0.4312	-0.3047	0.1021	0.1442
305	0.4161	-0.4645	0.1326	0.3699	-0.6914	0.4251	-0.3104	0.0987	0.1538
301	0.3996	-0.4290	0.1249	0.3638	-0.5741	0.4117	-0.3536	0.0961	0.1467
307	0.4198	-0.4325	0.1252	0.3048	-0.5357	0.4576	-0.3986	0.0918	0.1510
281	0.4177	-0.4318	0.1213	0.3562	-0.5326	0.4419	-0.3870	0.0980	0.1378

TABLE IV
PARAMETER ESTIMATION RESULTS ($T = 200, k = 100, [\sigma^{(1)}]^2 = 0.1, [\sigma^{(2)}]^2 = 0.3$)

Not considering C.P. ($\hat{\theta}_0$)					Considering C.P. ($\hat{\theta}_1$)				
$\hat{k}_1$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\sigma}^2$	$\hat{\phi}_1^{(1)}$	$\hat{\phi}_2^{(1)}$	$\hat{\phi}_1^{(2)}$	$\hat{\phi}_2^{(2)}$	$[\hat{\sigma}^{(1)}]^2$	$[\hat{\sigma}^{(2)}]^2$
97	0.4689	-0.2939	0.1910	0.3729	-0.5978	0.4730	-0.2468	0.1006	0.2447
107	0.4278	-0.3378	0.2036	0.4460	-0.3819	0.4267	-0.3334	0.0885	0.2847
104	0.4601	-0.3720	0.1998	0.3139	-0.5495	0.4781	-0.3473	0.1028	0.2669
102	0.5311	-0.3775	0.2306	0.3667	-0.5451	0.5410	-0.3673	0.0926	0.3155
106	0.4710	-0.3828	0.2224	0.3833	-0.3977	0.4808	-0.3832	0.0990	0.3084
96	0.3804	-0.5105	0.1987	0.4100	-0.5384	0.3778	-0.5085	0.0927	0.2621

103	0.4052	-0.3017	0.2148	0.3448	-0.5693	0.4092	-0.2683	0.0935	0.2924
102	0.5022	-0.3756	0.2507	0.2859	-0.5667	0.5182	-0.3618	0.1084	0.3392
102	0.3706	-0.3375	0.2134	0.3109	-0.6466	0.3749	-0.2883	0.0959	0.2869
103	0.4016	-0.3751	0.2164	0.5311	-0.5716	0.3800	-0.3391	0.0998	0.2928

TABLE V
PARAMETER ESTIMATION RESULTS $\left(T = 200, k = 100, [\sigma^{(1)}]^2 = 0.1, [\sigma^{(2)}]^2 = 0.15\right)$

Not considering C.P. $\left(\vec{\theta}_0\right)$				Considering C.P. $\left(\vec{\theta}_1\right)$					
$\hat{k}_1$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\sigma}^2$	$\hat{\phi}_1^{(1)}$	$\hat{\phi}_2^{(1)}$	$\hat{\phi}_1^{(2)}$	$\hat{\phi}_2^{(2)}$	$[\hat{\sigma}^{(1)}]^2$	$[\hat{\sigma}^{(2)}]^2$
105	0.4985	-0.4570	0.2266	0.4150	-0.5257	0.5114	-0.4504	0.1144	0.3071
96	0.4034	-0.3254	0.2134	0.3032	-0.5643	0.4120	-0.2981	0.0950	0.2810
107	0.4818	-0.3312	0.2128	0.4615	-0.4523	0.4840	-0.3196	0.0877	0.2989
102	0.5151	-0.3737	0.2272	0.4462	-0.4204	0.5213	-0.3711	0.0939	0.3116
100	0.4466	-0.3565	0.2305	0.4830	-0.6235	0.4405	-0.3187	0.0978	0.3101
101	0.4000	-0.3816	0.2282	0.4409	-0.5977	0.3936	-0.3449	0.1069	0.3056
100	0.5570	-0.4363	0.1930	0.4604	-0.5617	0.5697	-0.4194	0.1037	0.2521
106	0.5603	-0.4113	0.2201	0.4338	-0.4791	0.5745	-0.4080	0.1070	0.3007
99	0.4440	-0.3466	0.2248	0.4688	-0.7069	0.4326	-0.2759	0.0957	0.2973
109	0.4483	-0.2914	0.2133	0.4279	-0.4332	0.4504	-0.2727	0.0998	0.2978

TABLE VI
PARAMETER ESTIMATION RESULTS $\left(T = 100, k = 50, [\sigma^{(1)}]^2 = 0.1, [\sigma^{(2)}]^2 = 0.3\right)$

Not considering C.P. $\left(\vec{\theta}_0\right)$				Considering C.P. $\left(\vec{\theta}_1\right)$					
$\hat{k}_1$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\sigma}^2$	$\hat{\phi}_1^{(1)}$	$\hat{\phi}_2^{(1)}$	$\hat{\phi}_1^{(2)}$	$\hat{\phi}_2^{(2)}$	$[\hat{\sigma}^{(1)}]^2$	$[\hat{\sigma}^{(2)}]^2$
56	0.3457	-0.1805	0.2423	0.3157	-0.6206	0.3456	-0.1153	0.0928	0.3488
52	0.4251	-0.3236	0.2095	0.3779	-0.4321	0.4324	-0.3141	0.0926	0.2896
52	0.4297	-0.3555	0.1779	0.5192	-0.4955	0.4156	-0.3291	0.0882	0.2418
52	0.4804	-0.3386	0.2058	0.4081	-0.4875	0.4913	-0.3201	0.1014	0.2798
52	0.4873	-0.2753	0.2110	0.4323	-0.2622	0.4936	-0.2782	0.0881	0.2937
49	0.4991	-0.4363	0.2255	0.4713	-0.3112	0.5031	-0.4487	0.0922	0.3055
44	0.5585	-0.5267	0.2313	0.1289	-0.3764	0.5706	-0.5341	0.0733	0.3027
68	0.5410	-0.5601	0.2163	0.4268	-0.4308	0.5831	-0.6119	0.1356	0.3319
53	0.3316	-0.2706	0.2279	0.4554	-0.6095	0.3151	-0.2169	0.0924	0.3174
45	0.4766	-0.4067	0.2184	0.4820	-0.6828	0.4731	-0.3649	0.0993	0.2804

TABLE VII
PARAMETER ESTIMATION RESULTS $\left(T = 100, k = 50, [\sigma^{(1)}]^2 = 0.1, [\sigma^{(2)}]^2 = 0.15\right)$

Not considering C.P. $\left(\vec{\theta}_0\right)$				Considering C.P. $\left(\vec{\theta}_1\right)$					
$\hat{k}_1$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\sigma}^2$	$\hat{\phi}_1^{(1)}$	$\hat{\phi}_2^{(1)}$	$\hat{\phi}_1^{(2)}$	$\hat{\phi}_2^{(2)}$	$[\hat{\sigma}^{(1)}]^2$	$[\hat{\sigma}^{(2)}]^2$
52	0.5178	-0.4386	0.1222	0.1831	-0.4862	0.5948	-0.4562	0.0858	0.1474
34	0.3203	-0.4528	0.1229	-0.1096	-0.3763	0.3616	-0.4708	0.0659	0.1429
65	0.3720	-0.3953	0.1192	0.4107	-0.5749	0.2442	0.1462	0.1160	0.1044
58	0.2448	-0.4582	0.1180	0.4314	-0.5961	0.1052	-0.3685	0.0898	0.1448
43	0.5574	-0.2976	0.1323	0.6171	-0.4517	0.5505	-0.2622	0.0843	0.1595
47	0.4694	-0.4282	0.1258	0.2618	-0.5946	0.5073	-0.4043	0.0773	0.1557
48	0.3256	-0.4246	0.1223	0.2429	-0.5099	0.3476	-0.4070	0.0828	0.1511
61	0.5787	-0.6075	0.1056	0.4885	-0.7802	0.6406	-0.5093	0.0787	0.1308
50	0.4079	-0.2570	0.1332	0.4356	-0.5149	0.3806	-0.1118	0.1032	0.1543
63	0.5059	-0.4880	0.1283	0.4591	-0.6417	0.5285	-0.3889	0.0954	0.1688

Tables II to VII list the estimated results under different conditions, including the change point, the variance of sequences with and without considering the change point, and the estimated results of other parameters in the model, and each Table provides 10 estimates. In Table II, the estimated value of the change point fluctuates slightly around 300, with most errors within 4 and a maximum error of 7,

indicating good estimation performance. In Table III, the variance of the two sequences is further reduced, and the estimation error of k also slightly increases, with a maximum absolute error of 18, but the overall deviation is not significant. The sample size in Table IV and Table V is 200, and the estimation results are still good. Further investigation revealed that when the sample size was reduced to 100 (as

shown in Tables V and VI), the prediction error increased. However, the situation was more pronounced in Table VII. As the variance differences in Table VII narrowed, its prediction outcomes became unstable. This instability was directly reflected in an increase in the deviation of the change point, with a maximum deviation reaching 16 and a mean deviation reaching 7.7, significantly exceeding the previous range of prediction errors. This finding indicates that the model's predictive performance may be severely compromised by reduced sample size and narrowed variance differences. Therefore, in future research, we should handle such data more cautiously and consider adopting more sophisticated or robust prediction models to improve the accuracy and stability of predictions.

The above findings have been validated through extensive simulations on numerous sequences. To further investigate the predictive performance of the CP-AR (2) model, a sequence is randomly selected for analysis with $T=200$ as an example.

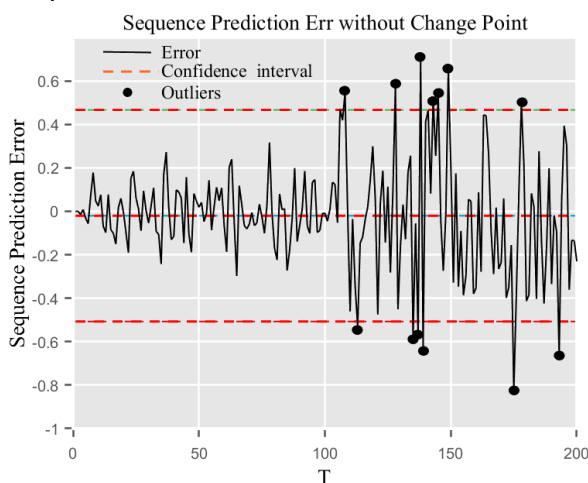


Fig. 4. Sequence Prediction Error. Without Change Point ($k = 100$, $\hat{k} = 105$). Note that the two dotted lines above and below are the 95% confidence interval, and the solid dots mark the points that exceed the interval.

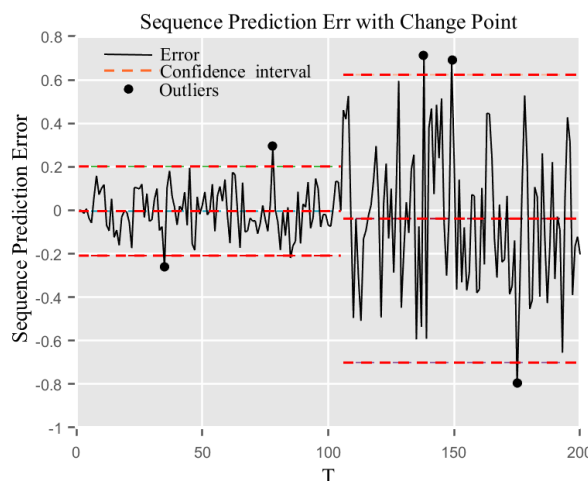


Fig. 5. Sequence Prediction Error. with Change Point ($k = 100$, $\hat{k} = 105$). Note that the two dotted lines above and below are the 95% confidence interval, and the solid dots mark the points that exceed the interval. Since modeling takes into account change points, the confidence intervals of errors before and after these points are different.

Fig. 4 and Fig. 5 demonstrate the prediction error situation through a sequence. This sequence contains a change point at $k=100$. The predicted change point position is 102. When modeling and predicting, Fig. 4 shows the error when change

points are not considered, while Fig. 5 shows the error situation after considering the change point modeling. For Fig. 4, when the change point is not considered, 13 points (marked by solid dots) lie outside twice the standard deviation from the mean. All outliers are in the sequence after the change point, while the earlier sequence contains no outliers, even with significant tolerance. In fact, after multiple simulation experiments, this situation remains the same, and all the outliers are concentrated in the latter half of the sequence. This situation is clearly unreasonable. Unlike Fig. 4, Fig. 5 considers the change point, and since the model parameters before and after the change point differ, the confidence intervals also vary accordingly. At this time, there are 5 points above twice the standard deviation of the mean and uniformly distributed, with 2 in the sequence before the change point and 3 in the later sequence. The number of outliers is significantly less than without considering the change point. To further assess the overall error profile, Table VII provides the analysis results of RMSE.

TABLE VIII
RMSE COMPARISON (TIMES=20)

T	k	$[\sigma^{(1)}]^2 / [\sigma^{(2)}]^2$	Not considering C.P. RMSE	Considering C.P. RMSE	RMSE comparison
600	300	0.1/0.3	0.1886	0.1691	11.50%
600	300	0.1/0.15	0.2061	0.1822	13.15%
200	100	0.1/0.3	0.2143	0.1912	12.11%
200	100	0.1/0.15	0.2237	0.1978	13.11%
100	50	0.1/0.3	0.2344	0.2101	11.57%
100	50	0.1/0.15	0.2524	0.2246	12.36%

Table VIII compares the RMSE values obtained from modeling with and without change points for stationary autoregressive sequences. The former involves identifying change points and deriving model expressions before and after using the method proposed in this paper, whereas the latter involves direct modeling. The first column of the table indicates the sequence length, the second column shows the location of sequence change points, and the third column displays the variance of the sequences before and after the change points. For each of the six scenarios presented in the table, data simulation and modeling were conducted 20 times based on specific conditions, and the RMSE values were calculated as the average of these 20 results. In all six scenarios, the RMSE values obtained from modeling with change points were consistently lower than those obtained without change points. This indicates that considering change points can significantly enhance models' fitting effectiveness and prediction accuracy when dealing with stationary autoregressive time series with change points. Furthermore, as the sequence length increases from 100 to 200 and then to 600, the RMSE values with and without considering change points generally decrease to some extent, suggesting that longer sequence data helps the model learn patterns and reduce errors. However, the reduction in RMSE was more pronounced when change points were considered, highlighting its advantage. Additionally, even if parameters such as the variance before and after the change points vary under the same sequence length, modeling with change points demonstrates greater stability and superiority.

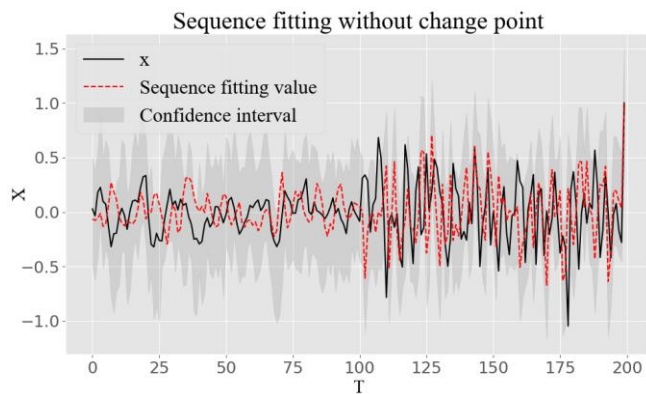


Fig. 6. Sequence fitting without a change point. Note that the black line is the sequence. The red dashed line is a fitting value. The gray area is the confidence interval.

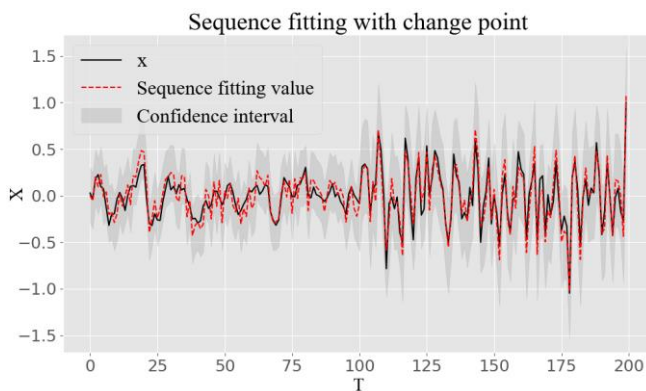


Fig. 7. Sequence fitting with a change point. Note that the black line is the sequence. The red dashed line is the fitting value. The gray area is the confidence interval.

Fig. 6 shows the prediction and fitting of the sequence without considering the change point. Fig. 7 shows the prediction and fitting of the sequence while considering the change point. The solid line indicates the actual value, the dashed red line represents the fitted trend of the data series, and the shaded area indicates the 95% confidence interval of the sequence fitting values. When considering change points, the fit is better compared to the scenario without change points. The confidence interval in Fig. 6 shows slight fluctuation, even though the sequence variance before and after the change point has changed significantly. The confidence interval in Fig. 7 is smaller before the change point but suddenly becomes larger after the change point, mainly due to the increase in sequence fluctuation after the change point. The pattern in the confidence interval when considering the change point is much more reasonable, and the model is more reliable.

V. CONCLUSION

This paper primarily investigates the change-point problem in autoregressive (P) models, defining these issues as CP-AR (P) models. While existing studies have explored cases where sequence variance remains unchanged at change points, it further examines scenarios with varying variance, providing estimates for the change point k , model parameters, and variances before and after the change point. In the derivation process, symmetric matrices are introduced, thereby deriving the estimation expression for model coefficients. Then, variance estimates are obtained, and the

estimation expression for the location of change points is provided in conjunction with the likelihood ratio. In Chapter 4, a large number of data simulations are conducted based on the CP-AR(2) model. Unlike most data simulations for change-point problems, this study simulates almost every point in the sequence as a change point, with each point undergoing 20 simulations. To ensure a sufficient sample size, only some sequence points at the beginning and end are excluded. This design more fully verifies the reliability of the model. Also, it leads to the following important conclusions: Firstly, it is found that the estimation bias rate of change-point locations is closely related to the sequence length (i.e., sample size). The larger the sample, the smaller the bias rate. It is recommended to retain at least 30 sample points before and after the change point to ensure estimation accuracy. For details, please refer to Fig. 2, Fig. 3, and Table I. Additionally, it has been observed that the greater the variance difference between the sequences before and after the change point, the higher the accuracy of the change-point estimation. Some simulation results are detailed in Tables II to VII. To further verify the model's effectiveness, a randomly selected sequence is compared regarding the error and fitting performance with and without considering change points. Refer to Figs. 4 to 7 and Table VIII. It is found that the model considering change points is indeed superior. However, change points in the sequence are often not easily detectable (Figs. 1 and 2). When there are change points in the sequence, direct modeling can lead to poor fitting performance. Therefore, it is recommended to first perform change-point detection. The specific process is shown in Fig. 8.

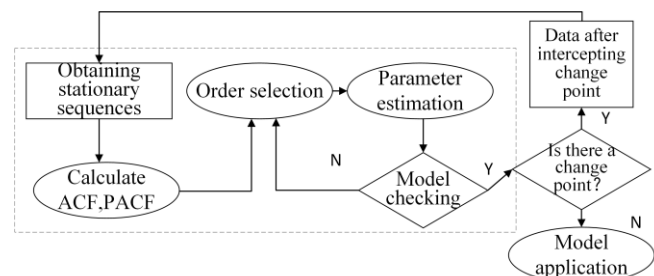


Fig. 8. ARIMA modeling flowchart considering change point.

- Step 1:** Preprocess the time series data to make it stable.
- Step 2:** Calculate ACF and PACF and determine model order.
- Step 3:** Find the change point and verify the estimated change point. The SIC method can be used here.
- Step 4:** If the change point k is not close to the sequence's end or beginning, use the sequence following the change point for remodeling.

Fig. 8 applies not only to the modeling of CP-AR(P) models but also to ARMA models with change points, which can be further discussed later. From a practical application perspective, detecting changes in the time series before modeling is recommended to improve the reliability of model predictions.

Although this study has conducted some theoretical research and drawn a series of valuable conclusions through simulation verification and practical analysis, there are still some limitations that warrant attention. First, the method exhibits higher computational complexity for higher-order models: In simulation, the second-order model is primarily

exemplified, featuring a relatively complex estimation expression. The detailed process can be referred to in formulas (5) to (16), from which the second-order parameter estimation expressions can be derived, as shown in (23) to (29). When estimating Φ , (5) to (7) yield $\Phi=Z^{-1}B$, necessitating matrix operations. For higher-order scenarios, computational complexity increases exponentially. Although high-order models are seldom employed in practical applications, this issue requires further investigation. Second, estimation instability occurs at sequence boundaries (Fig. 2 and 3), primarily when fewer than 30 samples are available for modeling in segments around change points. Insufficient observations fail to constrain the parameter space, increasing estimate variance by 20 to 50% compared to central change points. Furthermore, Table V and Table VI reveal that the small sample size can cause inaccuracy. Therefore, the method is more suitable for long series. Additionally, the model assumes a single change point and a Gaussian distribution. Preliminary tests with non-Gaussian distributions (e.g., heavy-tailed) showed reduced accuracy, though detailed results are omitted here. Future work should address multi-change-point scenarios and non-parametric distributions.

REFERENCES

- [1] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1, pp. 100–115, Jun. 1954.
- [2] E. S. Page, "A test for a change in a parameter occurring at an unknown point," *Biometrika*, vol. 42, no. 3–4, pp. 523–527, Dec. 1955.
- [3] T. L. Lai, "Sequential changepoint detection in quality control and dynamical systems," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 57, no. 4, pp. 613–644, Nov. 1995.
- [4] L. Horváth, "The maximum likelihood method for testing changes in the parameters of normal observations," *The Annals of Statistics*, vol. 21, no. 2, pp. 671–680, Jun. 1993.
- [5] L. Horváth, "Detection of changes in linear sequences," *Ann. Inst. Stat. Math.*, vol. 49, pp. 271–283, Jun. 1997.
- [6] G. Ciuperca and M. Maciak, "Change-point detection in a linear model by adaptive fused quantile method," *Scand. J. Stat.*, vol. 47, no. 2, pp. 425–463, Jun. 2020.
- [7] R. Sheikhrabari and M. Aminnayeri, "Maximum likelihood estimation of the change point in stationary state of auto regressive moving average (ARMA) models, using SVD-based smoothing," *Commun. Stat. Theory Methods*, vol. 51, no. 22, pp. 7801–7818, Feb. 2021.
- [8] C. Liu, "The maximum likelihood estimation of the change point of ARMA models," *Math. Appl.*, vol. 5, pp. 111–113, Jun. 1992.
- [9] T. Shiohama and M. Taniguchi, "Asymptotic estimation theory of change-point problems for time series regression models and its applications," *Lecture Notes Monogr. Ser.*, vol. 41, pp. 257–284, Dec. 2003.
- [10] L. Wang, "Testing for change-point of the first-order autoregressive time series models," *Chin. J. Appl. Probabil. Stat.*, vol. 24, pp. 28–36, Mar. 2008.
- [11] E. Gombay, "Sequential change-point detection with likelihood ratios. Stat," *Probabil. Lett.*, vol. 49, no. 2, pp. 195–204, Aug. 2000.
- [12] N. Keriven, D. Garreau, and I. Poli, "NEWMA: A new method for scalable model-free online change-point detection," *IEEE Trans. Signal Process.*, vol. 68, pp. 3515–3528, Jan. 2020.
- [13] A. Omid, R. Kamranrad, and E. Mardan, "A heuristic approach to change point estimate in a nonparametric processes parameter," *Int. J. Prod. Q. Manag.*, vol. no. 4, pp. 496–512, Dec. 2021.
- [14] Y. C. Yao and R. A. Davis, "The asymptotic behavior of the likelihood ratio statistic for testing a shift in mean in a sequence of independent normal variates," *Sankhyā*, vol. 48, no. 3, pp. 339–353, Oct. 1986.
- [15] E. Gombay and L. Horváth, "On the rate of approximations for maximum likelihood tests in change-point models," *J. Mult. Anal.*, vol. 56, no. 1, pp. 120–152, Jan. 1996.
- [16] R. A. Manzano Sanchez, K. Naik, A. Albasir, M. Zaman, and N. Goel, "Detection of anomalous behavior of smartphone devices using changepoint analysis and machine learning techniques," *Digit. Threats*, vol. 4, no. 1, pp. 1–28, Feb. 2022.
- [17] R. K. Ahir, B. Chakraborty, and P. Mitra, "Informed change-point detection approach for solar prosumer detection and statistical verification in smart grid," *IEEE Trans. Smart Grid*, vol. 15, no. 1, pp. 987–998, Jan. 2024.
- [18] S. Inoue and S. Yamada, "Software shipping time estimation for Markovian change-point reliability model," *Int. J. Reliab. Qual. Saf. Eng.*, vol. 28, no. 3, p. 2150021, Jan. 2021.
- [19] O. O. Awe and A. A. Adepoju, "Change-point detection in CO2 emission-energy consumption nexus using a recursive bayesian estimation approach," *Stat. Trans. New Ser.*, vol. 21, no. 1, pp. 123–136, Jan. 2020.
- [20] K. Kim, J. H. Park, M. Lee, and J. W. Song, "Unsupervised Change Point Detection and Trend Prediction for Financial Time-Series Using a New CUSUM-Based Approach," *IEEE Access*, vol. 10, pp. 34690–34705, Apr. 2022.
- [21] S. Takahashi, K. Takeshita, K. Yamagishi and A. Shiozu, "Change Point Detection Based on Cluster Transition Distributions," *IEEE Access*, vol. 12, pp. 125145–125159, Sept. 2024.
- [22] M. Stasolla and X. Neyt, "A Novel Change Point Detection Method for Data Cubes of Satellite Image Time Series," *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium*, pp. 5696–5699, Oct. 2023.
- [23] A. Aue and L. Horváth, "Structural breaks in time series," *J. Time Ser. Anal.*, vol. 34, no. 1, pp. 1–16, Jan. 2013.
- [24] V. Jandhyala, S. Fotopoulos, I. MacNeill, and P. Liu, "Inference for single and multiple change-points in time series," *J. Time Ser. Anal.*, vol. 34, no. 4, pp. 423–446, Jul. 2013.
- [25] K. Atashgar, "Identification of the change point: an overview," *The International Journal of Advanced Manufacturing Technology*, 64.9-12:1663-1683, Jun. 2013.
- [26] T. S. Lee, "Change-point problems: bibliography and review," *J. Stat. Theory Pract.*, vol. 4, pp. 643–662, Nov. 2010.