Journal of Quality Engineering and Management

www.pqprc.ir

J. Qual. Eng. Manag. Vol. 16, No. 2 (2026) 186–208.

Paper Type: Original Article

Parameter Estimation of the Weighted Composite Quantile Autoregressive Model Using the SEM Algorithm and Its Application in Statistical Quality Control

Mohammad Bahmani[1], **Mahdi Kalantari**[2,*]

[1] Department of Statistics, Faculty of Mathematics, Statistics and Computer Science, University of Tabriz, Tabriz, Iran; bahmanimo74@znu.ac.ir.
[2] Department of Statistics, Payame Noor University, Tehran, Iran; kalantarimahdi@pnu.ac.ir.

Citation:

Received: 10 February 2026 *Revised: 16 April 2026* *Accepted: 26 May 2026*	Bahmani, M., & Kalantari, M. (2026). Parameter estimation of the weighted composite quantile autoregressive model using the SEM algorithm and its application in statistical quality control. *Journal of Quality Engineering and Management, 16*(2), 186-208.

Abstract

Purpose: Linear autoregressive models are among the most widely used models in time series analysis. However, the presence of outliers and data with heavy-tailed distributions significantly reduces the accuracy of parameter estimation in these models. Therefore, the use of quantile models as an effective solution to address this challenge and improve analyses in fields such as statistical quality control is essential.

Methodology: In this article, the parameters of the Weighted Composite Quantile Autoregressive (WCQAR) model are estimated using a frequentist approach, and its application in statistical quality control is presented. The performance of the proposed method is evaluated through a comprehensive simulation study and the analysis of a real dataset (the percentage weight of a material in the output of a distillation column).

Findings: The results of the simulation and real data analysis show that the Stochastic Expectation-Maximization (SEM) algorithm, compared to conventional methods, provides more accurate and stable estimates and effectively overcomes problems caused by outliers and heavy-tailed distributions.

Originality/Value: The main innovation of this research is the use of the SEM algorithm instead of conventional algorithms such as EM to compute the maximum likelihood estimators for the parameters of the weighted mixture quantile autoregressive model. This approach can serve as a powerful and reliable tool in time series analysis and statistical quality control.

Keywords: Stochastic expectation-maximization algorithm, Asymmetric Laplace Composite distribution, Time series, Statistical quality control, Weighted mixture quantile autoregressive model.

Corresponding Author: kalantarimahdi@pnu.ac.ir
10.48313/jqem.2026.575020.1600
Licensee System Analytics. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0).

مهندسی و مدیریت کیفیت

www.pqprc.ir

دوره ۱۶، شماره (۲)، (۱۴۰۵)، ۲۰۸-۱۸۶

نوع مقاله: پژوهشی

برآورد پارامترهای مدل اتورگرسیو چندکی ترکیبی وزنی با استفاده از الگوریتم SEM و کاربرد آن در کنترل کیفیت آماری

محمد بهمنی[۱]، مهدی کلانتری[۲،*]

[۱] گروه آمار، دانشکده ریاضی، آمار و علوم کامپیوتر، دانشگاه تبریز، تبریز، ایران.

[۲] گروه آمار، دانشگاه پیام نور، تهران، ایران.

چکیده

هدف: مدل‌های اتورگرسیو خطی از پرکاربردترین مدل‌ها در تحلیل سری‌های زمانی هستند، اما وجود مشاهدات دورافتاده و داده‌هایی با توزیع‌های دم‌کلفت، دقت برآورد پارامترهای این مدل‌ها را به‌طور محسوسی کاهش می‌دهد. از این‌رو، استفاده از مدل‌های چندکی به‌عنوان یک راه‌حل موثر برای مقابله با این چالش و بهبود تحلیل‌ها در حوزه‌هایی مانند کنترل کیفیت آماری ضروری است.

روش‌شناسی پژوهش: در این مقاله، پارامترهای مدل اتورگرسیو چندکی ترکیبی وزنی با رویکرد فراوانی‌گرا برآورد شده و کاربرد آن در کنترل کیفیت آماری ارایه شده است. عملکرد روش پیشنهادی از طریق یک مطالعه شبیه‌سازی جامع و نیز تحلیل یک مجموعه داده واقعی (درصد وزن یک ماده در خروجی برج تقطیر) مورد ارزیابی قرار گرفته است.

یافته‌ها: نتایج شبیه‌سازی و تحلیل داده واقعی نشان می‌دهد که الگوریتم *SEM* در مقایسه با روش‌های متداول، برآوردهای دقیق‌تر و پایدارتری ارایه می‌دهد و به‌طور موثری بر مشکلات ناشی از داده‌های پرت و توزیع‌های دم‌کلفت غلبه می‌کند.

اصالت/ارزش افزوده علمی: نوآوری اصلی این پژوهش، به‌کارگیری الگوریتم *SEM* به‌جای الگوریتم‌های مرسوم مانند *EM* برای محاسبه برآوردگرهای حداکثر درست‌نمایی پارامترهای مدل اتورگرسیو چندکی ترکیبی وزنی است. این رویکرد می‌تواند به‌عنوان یک ابزار قدرتمند و قابل‌اعتماد در تحلیل سری‌های زمانی و کنترل کیفیت آماری مورد استفاده قرار گیرد.

کلیدواژه‌ها: الگوریتم *SEM*، توزیع لاپلاس نامتقارن ترکیبی، سری‌های زمانی، کنترل کیفیت آماری، مدل اتورگرسیو چندکی ترکیبی وزنی.

۱- مقدمه

مدل‌های اتورگرسیو[۱] به‌عنوان یکی از مهم‌ترین و پرکاربردترین ابزارها در تحلیل سری‌های زمانی، توجه گسترده محققان را به خود جلب کرده‌اند. این مدل‌ها به‌دلیل ترکیب سادگی و انعطاف‌پذیری، در حوزه‌های متنوعی از جمله اقتصاد، مالی، مهندسی و علوم محیطی به‌طور گسترده مورد استفاده

[1] Autoregressive (AR)

قرار می‌گیرند. روش‌های رایج برای برآورد پارامترهای مدل *AR* عبارتند از روش حداقل مربعات معمولی[1] و روش حداکثر درست‌نمایی[2]. اگر توزیع خطا را در مدل *AR* توزیع نرمال در نظر بگیریم، مشاهدات دورافتاده قابل‌توجهی در مدل نخواهیم داشت. در این حالت، روش‌های *OLS* یا *ML* کارایی لازم را در برآورد پارامترهای مدل دارند. ولی اگر توزیع خطا از یک توزیع دم‌کلفت[3] پیروی کند، مشاهدات دورافتاده قابل‌توجهی تولید می‌شوند و برآورد پارامترهای مدل *AR* به روش *OLS* یا *ML* تحت تاثیر قرار می‌گیرد.

به‌طورکلی می‌توان گفت وجود مشاهدات دورافتاده در یک سری زمانی می‌تواند تاثیر منفی بر تحلیل‌ها ازجمله استنباط درمورد پارامترهای مدل و پیش‌بینی مقادیر آینده سری زمانی داشته باشد. به‌عنوان‌مثال، وجود این نوع مشاهدات می‌تواند مانایی سری زمانی را که یک شرط اساسی در تحلیل کلاسیک سری‌های زمانی محسوب می‌شود، از بین ببرد [1]. همچنین مشاهدات دورافتاده تاثیر منفی بر ساختار خودهمبستگی در سری زمانی و درنتیجه روی نمودارهای خودهمبستگی و خودهمبستگی جزیی داشته و این موضوع سبب می‌شود که مرتبه مدل‌های *ARIMA* به‌درستی تشخیص داده نشود [2]. علاوه‌بر این، مشاهدات دورافتاده می‌توانند در برآورد چگالی طیفی نیز نقشی مخرب ایفا کنند [3]. در [4] انواع مختلف مشاهدات دورافتاده معرفی شده و روش‌های شناسایی و برآورد اثر این نقاط، در قالب چند مثال مورد بحث قرار گرفته است. برای کسب اطلاعات بیشتر و مروری در مورد مشاهدات دورافتاده در سری‌های زمانی می‌توان به [5] مراجعه کرد.

در کنترل کیفیت آماری نیز با مشاهدات دورافتاده مواجه می‌شویم، به‌ویژه زمانی که انحرافات با دلیل[4] در فرایند وجود داشته و مشاهدات مستقل نبوده و خودهمبستگی داشته باشند. در این حالت، استفاده از نمودارهای کنترل مرسوم که با فرض مستقل‌بودن مشاهدات طراحی شده‌اند، به دو دلیل کارایی خود را از دست می‌دهند [6]:

۱. افزایش زنگ خطرهای اشتباهی[5]: یعنی تعداد دفعاتی که محصولات مرغوب، به اشتباه جزو محصولات نامرغوب به‌حساب آورده می‌شود؛ افزایش خواهد یافت و این امر موجب می‌شود که کارخانه سازنده متحمل هزینه‌های اضافی و سنگینی شود.
۲. کاهش توانایی نمودار کنترل نسبت به کشف مشاهدات خارج از کنترل: یعنی محصولاتی نامرغوب تولید می‌شوند ولی نمودار کنترل قادر به شناسایی همه این محصولات نیست؛ درنتیجه برخی از محصولات نامرغوب به اشتباه جزو محصولات مرغوب به حساب آورده می‌شود و این امر موجب نارضایتی مشتری و ازبین‌رفتن اعتبار کارخانه سازنده خواهد شد.

یکی از روش‌های طراحی نمودارهای کنترل برای مشاهدات خودهمبسته، استفاده از مدل‌های سری زمانی است تا بتوان خودهمبستگی بین مشاهدات را حذف نموده و نمودارهای کنترل مرسوم را با استفاده از مانده‌های حاصل از برازش مدل‌های سری زمانی (که ناهمبسته هستند) رسم کرد [7]. بنابراین اگر مشاهدات خودهمبسته باشند، مدل‌های سری زمانی کاربرد زیادی در کنترل کیفیت آماری خواهند داشت، به‌ویژه اگر با مشاهدات دورافتاده مواجه شویم.

یک راه حل برای کاهش اثرات منفی مشاهدات دورافتاده روی نتایج حاصل از به‌کارگیری مدل‌های مرسوم سری زمانی، استفاده از روش‌های چندکی است. مدل اتورگرسیو چندکی[6]، اولین بار توسط کونکر و شیائو [8] مورد بررسی قرار گرفت. این مدل اطلاعات جزیی‌تری درخصوص تاثیر متغیرهای مستقل بر چندک‌های متغیر پاسخ را ارایه می‌دهد. مدل *QAR* در مقایسه با مدل *AR* دارای سه مزیت مهم است:

۱. مدل *QAR* نسبت به مشاهدات دورافتاده، استوار[7] است.
۲. مدل *QAR* اطلاعات جزیی‌تری درخصوص تاثیر متغیرهای مستقل در چندک‌های مختلف توزیع متغیر پاسخ ارایه می‌دهد.

[1] Ordinary Least Squares (OLS)
[2] Maximum Likelihood (ML)
[3] Heavy tail
[4] Assignable causes
[5] False alarms
[6] Quantile Autoregressive (QAR)
[7] Robust

۳. در مدل *QAR* هیچ فرضی درمورد توزیع خطاهای مدل وجود ندارد، به‌جز اینکه فرض می‌شود چندک τth خطاهای مدل برابر صفر است، به عبارت دیگر:

$$\int_{-\infty}^{0} f(\varepsilon_t) d\,\varepsilon_t = \tau,$$

به‌طوری‌که ε_t و $f(\varepsilon_t)$ به‌ترتیب نشان‌دهنده خطای مدل و تابع چگالی خطای مدل است. از این‌رو مدل *QAR* را می‌توان یک مدل توزیع آزاد در نظر گرفت. به‌دلیل کاربرد فراوان مدل‌های *QAR*، پژوهشگران زیادی این مدل‌ها را از دیدگاه آمار بیزی و فراوانی‌گرا مورد مطالعه قرار داده‌اند. به‌عنوان‌مثال، ژنگ و لی [9] مدل اتورگرسیو میانه بیزی را برای پیش‌بینی در سری‌های زمانی به‌کار برده‌اند و تیان و همکاران [10] برآورد پارامترهای *QAR* تاخیری را با رویکرد فراوانی‌گرا و با استفاده از الگوریتم بیشینه‌سازی امید ریاضی[1] به‌دست آوردند. همچنین، بهمنی [11] به بررسی مدل *QAR* خطی پرداخته و از الگوریتم بیشینه‌سازی امید ریاضی تصادفی[2] به‌عنوان یک روش تکراری برای محاسبه برآوردهای *ML* استفاده کرده و کارایی آن را از طریق مطالعات شبیه‌سازی و تحلیل داده‌های واقعی مورد ارزیابی قرار داده است. این پژوهش‌ها به‌وضوح اهمیت و کاربرد مدل‌های *QAR* در تحلیل داده‌های سری زمانی را نشان می‌دهند.

یکی از محدودیت‌های اصلی رگرسیون چندکی این است که برآوردها تنها بر اساس یک سطح مشخص از چندک محاسبه می‌شوند که ممکن است در برخی موارد دقت کمتری نسبت به رگرسیون میانگین داشته باشد. برای رفع این محدودیت، مدل رگرسیون چندکی ترکیبی[3] که توسط زو و یوان [12] معرفی شده است، با ترکیب اطلاعات از چندین سطح چندکی، برآوردهایی جامع‌تر و دقیق‌تر ارایه می‌دهد. این مدل، با بهره‌گیری از داده‌های کل توزیع و ادغام چندک‌های مختلف، برآوردهایی کارآمدتر و استوارتر نسبت به روش‌های سنتی مانند *OLS* و *ML* فراهم می‌کند در سال‌های اخیر، تحقیقات گسترده‌ای در زمینه مدل *CQR* انجام شده است. به‌عنوان‌مثال، آقا محمدی و بهمنی [13] مدل رگرسیون چندکی ترکیبی وزنی با خطاهای اتورگرسیو مرتبه p را با رویکرد بیزی مورد بررسی قرار دادند و نتایج تجربی آن‌ها نشان داد که این مدل از کارایی بیشتری نسبت به سایر مدل‌ها برخوردار است. همچنین، وای یو ژو و همکاران [14] با استفاده از الگوریتم نمونه‌گیری گیبز، برآورد پارامترهای مدل رگرسیون چندکی ترکیبی بیزی را به‌دست آوردند. علاوه‌بر این، ژائو و همکاران [15] به مطالعه *CQR* با خطاهای همبسته پرداخته و به تحلیل اثرات این نوع خطاها بر دقت برآوردها پرداختند. این مطالعات نشان‌دهنده پیشرفت‌های قابل‌توجه در کاربرد مدل‌های *CQR* و توانایی آن‌ها در ارایه نتایج دقیق‌تر در تحلیل داده‌های پیچیده هستند.

در مدل *CQR*، فرض بر این است که تاثیر تمامی چندک‌ها در برآورد پارامترها یکسان است. بااین‌حال، شواهد تجربی نشان می‌دهد که این فرض همیشه صحیح نیست، زیرا چندک‌های مختلف ممکن است نقش متفاوتی در توصیف ویژگی‌های داده‌ها ایفا کنند. به‌طور خاص، چندک‌های میانی اغلب در مقایسه با چندک‌های ابتدایی یا انتهایی، نماینده بهتری برای ساختار کلی داده‌ها هستند. به همین دلیل، وزن‌دهی به هر سطح از چندک می‌تواند به انعطاف‌پذیری بیشتر مدل و لحاظ‌کردن تفاوت تاثیر چندک‌ها در برآورد پارامترها منجر شود. برای رفع این محدودیت، ژائو و شیائو [16] مدل رگرسیون چندکی ترکیبی وزنی را معرفی کردند. آن‌ها نشان دادند که با اختصاص وزن‌های مناسب به هر سطح از چندک، می‌توان دقت و کارایی مدل را به‌طور قابل‌توجهی افزایش داد. نتایج مطالعات آن‌ها حاکی از آن است که مدل رگرسیون چندکی ترکیبی وزنی در مقایسه با مدل رگرسیون چندکی ترکیبی معمولی، برآوردهای دقیق‌تر و کاراتری ارایه می‌دهد و قابلیت بیشتری در تحلیل داده‌های پیچیده دارد.

در این مقاله، با الهام از ایده‌های مطرح‌شده توسط [12]، [16]، به بررسی برآورد پارامترهای مدل اتورگرسیو مرتبه p می‌پردازیم. این برآوردها با فرض توزیع لاپلاس نامتقارن ترکیبی[4] بر روی خطاها و با ترکیب اطلاعات چندک‌های مختلف متغیر پاسخ انجام می‌شود. ما از الگوریتم *SEM* به‌عنوان رویکرد اصلی خود استفاده کرده و بر اساس آمار فراوانی‌گرا به تحلیل داده‌ها می‌پردازیم. نوآوری اصلی این مقاله در استفاده از الگوریتم *SEM* به‌جای سایر الگوریتم‌های رایج مانند الگوریتم *EM* و نمونه‌گیری گیبز است که می‌تواند به بهبود کارایی برآوردها کمک کند. الگوریتم *SEM*

[1] Expectation–Maximization (EM)
[2] Stochastic Expectation-Maximization (SEM)
[3] Composite Quantile Regression (CQR)
[4] Composite Asymmetric Laplace (CAL)

که توسط سلو و دیبل [17] معرفی شد، به‌عنوان یک جایگزین قدرتمند و کارآمد برای الگوریتم *EM* شناخته می‌شود. این الگوریتم یک روش تکراری برای برآورد پارامترهای مدل به روش *ML* است که به‌طور خاص در مسایلی شامل متغیرهای پنهان یا داده‌های گم‌شده کاربرد دارد. هر تکرار این الگوریتم شامل دو مرحله اصلی است: مرحله (*Stochastic*) *S* و مرحله (*Maximization*) *M* که فرایند با یک مقدار اولیه مانند $\Psi^{(0)}$ آغاز می‌شود. الگوریتم *SEM* با جایگزین‌کردن شبیه‌سازی تصادفی متغیرهای پنهان در گام *S* به‌جای محاسبه امید ریاضی این متغیرها در گام *E* الگوریتم *EM*، ماهیتی تصادفی پیدا می‌کند.

در این چارچوب، گام *M* به‌صورت قطعی و گام *S* به‌صورت تصادفی انجام می‌شود و ترکیب این دو گام، یک زنجیره مارکوف از مقادیر پارامترها ایجاد می‌کند. همان‌گونه‌که در مطالعات [17]، [18] مطرح شده است، رفتار تصادفی ایجادشده در الگوریتم *SEM* بخش اساسی این الگوریتم را تشکیل می‌دهد و زنجیره حاصل، تحت شرایط مناسب، به یک توزیع ایستا همگرا می‌شود. همچنین باید در نظر داشته باشیم زنجیر تصادفی حاصل با نمونه‌های حاصل از توزیع پسینی در استنباط بیزی متفاوت است. از دیدگاه الگوریتمی، یکی از ویژگی‌های مهم الگوریتم *SEM* این است که تصادفی‌سازی موجود در گام *S* می‌تواند امکان اجتناب از گرفتارشدن در برخی ماکزیمم‌های محلی و نقاط زینی لگاریتم تابع درست‌نمایی را فراهم کند [19]. همچنین، از آن‌جاکه الگوریتم *SEM* لگاریتم تابع درست‌نمایی داده‌های شبه‌کامل حاصل از گام *S* را مستقیما در گام *M* بیشینه می‌کند، این الگوریتم می‌تواند به‌ویژه در مسایلی مفید باشد که محاسبه تحلیلی انتگرال مورد نیاز در گام *E* الگوریتم *EM* دشوار یا امکان‌پذیر نباشد.

از نظر نظری، دی بولت و رابرت [18] نشان دادند که ترکیب گام *M* و گام *S* تصادفی در الگوریتم *SEM*، زنجیره‌ای مارکوف ایجاد می‌کند که تحت شرایط مناسب به توزیع ایستا همگرا می‌شود. پس از سپری‌شدن یک دوره سوخت[1]، اثر مقادیر اولیه زنجیره کاهش یافته و تکرارهای متعاقب می‌توانند برای مطالعه ویژگی‌های توزیع ایستای زنجیره مورد استفاده قرار گیرند. در این چارچوب، میانگین توزیع ایستا با برآوردگر *ML* ارتباط دارد و در شرایط مجانبی، توزیع ایستا حول برآوردگر *ML* متمرکز می‌شود. از این‌رو، تکرارهای پس از *Burn-in* صرفا مجموعه‌ای از مقادیر تصادفی فاقد ارتباط با مساله برآورد پارامتر نیستند، بلکه اطلاعاتی درباره رفتار تصادفی و عدم‌قطعیت ناشی از الگوریتم *SEM* در اختیار قرار می‌دهند. برای مباحث نظری و جزئیات بیشتر درباره ویژگی‌های الگوریتم *SEM*، می‌توان به [20]، [21] نیز مراجعه کرد.

در مجموع، الگوریتم *SEM* نه‌تنها محدودیت‌های روش *EM* را برطرف کرده است، بلکه با ارایه یک رویکرد ساده‌تر و سریع‌تر، توانسته است جایگاه ویژه‌ای در مسایل آماری پیچیده پیدا کند. در بخش ۲ ابتدا مدل *QAR* و اتورگرسیو چندکی ترکیبی وزنی[2] به‌صورت اجمالی معرفی می‌شوند. سپس تابع درست‌نمایی برای مدل *WCQAR* با فرض توزیع *CAL* برای خطاهای مدل را به‌دست می‌آوریم. در بخش ۳، پارامترهای مدل *WCQAR* به روش *ML* با استفاده از الگوریتم *SEM* برآورد می‌شوند.

در بخش ۴ نیز، کارایی مدل *WCQAR* با استفاده از شبیه‌سازی و به‌کارگیری این روش برای تحلیل یک مجموعه داده در کنترل کیفیت آماری (درصد وزن یک نوع ماده در خروجی برج تقطیر) با سایر روش‌ها مقایسه می‌شود. علاوه‌بر این، به‌منظور دسترسی به کدهای پیاده‌سازی مقاله و بازتولید نتایج، تمامی کدهای مربوطه در مخزن *GitHub* نویسندگان به نشانی *https://github.com/Mohammad-bahmani/WCQAR* دردسترس قرار گرفته است.

۲- معرفی مدل‌های مختلف

در این بخش، به‌صورت اجمالی مدل‌های مختلف اتورگرسیو شامل *QAR*، اتورگرسیو چندکی ترکیبی و *WCQAR* معرفی خواهند شد.

[1] Burn-in

[2] Weighted Composite Quantile Autoregressive (WCQAR)

۱-۲- مدل QAR

مدل اتورگرسیو خطی مرتبه p را به‌صورت زیر در نظر بگیرید:

$$y_t = \sum_{j=1}^{p} \phi_j y_{t-j} + \varepsilon_t = y_t^T \Phi + \varepsilon_t \qquad (۱)$$

$$t = p+1, \dots, n,$$

که در آن y_t، مقدار سری زمانی در لحظه t، $y_t = (y_{t-1}, \dots, y_{t-p})^T$، $\Phi = (\phi_1, \dots, \phi_p)^T$ بردار پارامترهای اتورگرسیو و ε_t خطای تصادفی مدل را نشان می‌دهد که مستقل و هم‌توزیع هستند. چندک شرطی τth، y_t به شرط بردار $p-$بعدی مقادیر سری زمانی در p گام قبل‌تر به‌صورت زیر به‌دست می‌آید:

$$Q_\tau(y_t | F_{t-1}) = y_t^T \Phi, \qquad (۲)$$

که در آن F_{t-1} سیگما میدانی است که بر اساس تمام اطلاعات قبل از زمان $t-1$ ایجاد می‌شود. با تکیه بر چارچوب رگرسیون چندکی که کونکر و بست [22] معرفی کرده‌اند و همچنین ایده‌های مطرح‌شده توسط کونکر و شیائو [8]، برآورد پارامترها در مدل *QAR* برای چندک τth به‌صورت

$$\widehat{\Phi} = arg\ \min_{\Phi} \sum_{t=p+1}^{n} \rho_\tau (y_t - y_t^T \Phi) \qquad (۳)$$

به‌دست می‌آید که در آن $\rho_\tau(u)$ تابع زیان بوده و به‌صورت $\rho_\tau(u) = \frac{|u|+(2\tau-1)u}{2}$ تعریف می‌شود. اگر τ را برابر ۰/۵ در نظر بگیریم، آنگاه مدل اتورگرسیو میانه به‌دست می‌آید. باید توجه داشت که در مدل *QAR*، برآورد پارامترها از طریق کمینه‌سازی مجموع قدرمطلق انحرافات موزون[1] محاسبه می‌شود. این ویژگی باعث می‌شود که این روش در برابر مشاهدات دورافتاده استواری بالایی داشته باشد و نتایج حاصل از آن کمتر تحت تاثیر این نوع داده‌ها قرار گیرد. به همین دلیل، مدل *QAR* به‌عنوان یک روش استوار و قابل‌اعتماد در تحلیل داده‌هایی با توزیع‌های غیرنرمال یا دارای داده‌های دورافتاده شناخته می‌شود.

۲-۲- مدل اتورگرسیو چندکی ترکیبی

تابع زیان مورد استفاده در **رابطه (۳)**، برآوردگرها را بر اساس یک سطح چندک مشخص از توزیع مشاهدات به‌دست می‌دهد که در بسیاری از موارد، این برآوردگرها کارایی کمتری نسبت به روش‌های کلاسیک دارند. به‌منظور رفع این مشکل، تیان و همکاران [23] مدل *CQR* را معرفی کردند. در این مدل، برآورد پارامترها بر اساس ترکیبی از سطوح مختلف چندک محاسبه می‌شود. به عبارت دیگر، به‌جای تکیه بر یک سطح چندک خاص، مدل از اطلاعات حاصل از سطوح چندکی گوناگون بهره می‌برد و برآوردها را بر اساس آن‌ها به‌دست می‌آورد. این رویکرد امکان استفاده از اطلاعات بیشتری از توزیع داده‌ها را فراهم می‌کند و انتظار می‌رود که برآوردهای حاصل، کارایی بالاتری نسبت به روش‌های مبتنی بر یک سطح چندک داشته باشند. برآورد پارامترهای **مدل (۱)** در این روش با فرض K سطح چندک $0 < \tau_1 < \cdots < \tau_K < 1$ به‌صورت

$$\widehat{\Phi}, \hat{\alpha} = arg\ \min_{\Phi,\alpha} \sum_{t=p+1}^{n} \sum_{k=1}^{K} \rho_{\tau_k} \left(y_t - y_t^T \Phi - \alpha_{\tau_k}\right) \qquad (۴)$$

[1] Least Absolute Deviation (LAD)

به‌دست می‌آیند که در آن $\alpha = (\alpha_{\tau_1}, \dots, \alpha_{\tau_K})$ و α_{τ_k} چندک $\tau_k th$ مولفه خطا یعنی ε_t است. همچنین τ_k برابر $\frac{k}{K+1}$ برای $k = 1, \dots, K$ است. توجه داشته باشید که اگر $K = 1$ آنگاه مدل اتورگرسیو میانه به‌دست می‌آید. در **رابطه (۴)**، برآوردگرها از طریق ترکیب مجموعه‌ای از توابع هدف تحت چندک‌های مختلف محاسبه می‌شوند.

۳-۲- مدل WCQAR

در مدل اتورگرسیو چندکی ترکیبی تاثیر همه چندک‌ها بر برآورد پارامترها به‌صورت یکسان در نظر گرفته می‌شود. بااین‌حال، شواهد تجربی نشان می‌دهد که این فرض همیشه صحیح نیست و تاثیر هر چندک در برآورد پارامترها ممکن است متفاوت باشد. به عبارت دیگر، برخی از چندک‌های میانی ممکن است نماینده بهتری برای داده‌ها باشند و اطلاعات بیشتری نسبت به چندک‌های ابتدایی یا انتهایی ارایه دهند. به همین دلیل، با وزن‌دهی مناسب به هر سطح از چندک، می‌توان تاثیرات متفاوت آن‌ها را در برآورد پارامترهای مدل لحاظ کرد. در این راستا، ژائو و شیائو [16] مدل رگرسیون چندکی ترکیبی وزنی را معرفی کردند تا کارایی مدل *CQR* را افزایش دهند. نتایج تجربی آن‌ها نشان داد که مدل رگرسیون چندکی ترکیبی وزنی برآوردهای دقیق‌تری نسبت به مدل رگرسیون چندکی ترکیبی استاندارد ارایه می‌دهد. با استفاده از ایده‌های ارایه شده توسط [12] و [16]، برآورد پارامترهای مدل *WCQAR* به‌صورت زیر به‌دست می‌آید.

$$\widehat{\Phi}, \hat{\alpha} = arg \min_{\Phi,\alpha} \sum_{p+1}^{n} \sum_{k=1}^{K} w_k \rho_{\tau_k} \left(y_t - y_t^T \Phi - \alpha_{\tau_k}\right), \qquad (۵)$$

که در آن w_k وزن چندک kth بوده و $\sum_{k=1}^{K} w_k = 1$ است. بر این اساس، هرچه پارامتر مقیاس متناظر با یک سطح چندکی کوچک‌تر باشد، وزن آن سطح چندکی بیشتر خواهد بود و درنتیجه، سهم بیشتری در برآورد پارامترهای مشترک مدل خواهد داشت. درمقابل، با افزایش پارامتر مقیاس، وزن متناظر کاهش می‌یابد [23]. در بخش بعد، باتوجه‌به [23]، تابع درست‌نمایی را بر اساس فرض توزیع *CAL* برای خطاها و با استفاده از آمیخته مقیاس این توزیع تشکیل می‌دهیم.

۳- تابع درست‌نمایی مدل WCQAR بر اساس توزیع CAL

گوییم متغیر تصادفی Y دارای توزیع *CAL* است و با نماد $Y \sim CAL(\mu, \sigma, \tau)$ نشان می‌دهیم، هر گاه تابع چگالی احتمال آن متناسب با عبارت زیر باشد:

$$\prod_{k=1}^{K} \frac{1}{\sigma_k} exp\left\{-\rho_{\tau_k}\left(\frac{y - \mu_k}{\sigma_k}\right)\right\},$$

که در آن $\rho_\tau(u) = \frac{|u| + (2\tau - 1)u}{2}$ ، $\mu = (\mu_1, \cdots, \mu_K)^T$ بردار مکان، $\sigma = (\sigma_1, \cdots, \sigma_K)^T > 0$ بردار مقیاس و $0 < \tau_1 < \cdots < \tau_K < 1$ سطوح مختلف چندک بوده و می‌توان پارامترهای چولگی نیز در نظر گرفت. سریرام و همکاران [24] اولین بار این توزیع را در مبحث *CQR* مطرح کرده و نشان دادند از طریق این توزیع می‌توان برآوردهای کارایی را به‌دست آورد. حال در **مدل (۱)** فرض می‌کنیم که $y_t \sim CAL(\mu, \sigma, \tau)$ است که در آن:

$$\mu = \left(\mu_1 = y_t^T \Phi - \alpha_{\tau_1}, \cdots, \mu_K = y_t^T \Phi - \alpha_{\tau_K}\right)^T,$$

$$\sigma = (\sigma_1, \cdots, \sigma_K)^T, \quad \tau = (\tau_1, \cdots, \tau_K)^T$$

است. درنتیجه تابع چگالی احتمال y_t عبارت است از:

$$f(y_t|.) \propto \prod_{k=1}^{K} \frac{1}{\sigma_k} exp\left\{-\rho_{\tau_k}\left(\frac{y_t - y_t^T \Phi - \alpha_{\tau_k}}{\sigma_k}\right)\right\}. \qquad (۶)$$

برآوردهای *ML* پارامترهای Φ به‌دست‌آمده بر اساس تابع درست‌نمایی توزیع *CAL* معادل با مینیمم‌کردن تابع زیان زیر است:

$$\hat{\Phi}, \hat{\alpha} = arg \min_{\Phi,\alpha} \sum_{p+1}^{n} \sum_{k=1}^{K} \frac{1}{\sigma_k} \rho_{\tau_k} \left(y_t - y_t^T \Phi - \alpha_{\tau_k}\right),$$

که در آن $w_k = \frac{1}{\sigma_k}$ برای $k = 1, \cdots, K$ وزن‌ها را نشان می‌دهند. زمانی که $\sigma_1 = \cdots = \sigma_K$ مدل اتورگرسیو چندکی ترکیبی به‌دست می‌آید. به دلیل وجود تابع قدرمطلق در تابع درست‌نمایی مدل *WCQAR*، نمی‌توان برآوردگرها را به‌صورت یک شکل بسته محاسبه کرد. برای رفع این مشکل، به پیروی از [25]، توزیع لاپلاس نامتقارن را به‌صورت آمیخته‌ای از توزیع نرمال بیان می‌کنیم. به عبارت دیگر اگر متغیر تصادفی Y دارای توزیع لاپلاس نامتقارن باشد، آنگاه

$$Y|v \sim N(\theta_1 v, \theta_2 \sigma v), \quad v \sim Exp\left(\frac{1}{\sigma}\right), E(v) = \sigma, \qquad (۷)$$

درنتیجه با استفاده از **رابطه (۷)** تابع چگالی احتمال شرطی متغیر y_t به‌صورت سلسله‌مراتبی به شکل زیر بیان می‌شود:

$$\begin{cases} g(y_t|v_t, \Phi, \sigma, \tau) \propto \prod_{k=1}^{K} \frac{1}{\sqrt{\theta_{2k}\sigma_k v_{tk}}} exp\left\{\frac{(y_t - \eta_{tk} - \theta_{1k} v_{tk})^2}{2\theta_{2k}\sigma_k v_{tk}}\right\} \\ v_{tk} \sim Exp\left(\frac{1}{\sigma_k}\right), E(v_{tk}) = \sigma_k, \; k=1,\cdots, K, \end{cases} \qquad (۸)$$

که در آن

$$\eta_{tk} = y_t^T \Phi - \alpha_{\tau_k},$$

$$v = \{v_{p+1}, \cdots, v_n\}, \quad v_t = (v_{t1}, \cdots, v_{tK}),$$

و

$$\theta_{1k} = \frac{1\text{-}2\tau_k}{\tau_k(1-\tau_k)} \; ; \; \theta_{2k} = \frac{2}{\tau_k(1-\tau_k)}.$$

باتوجه‌به **رابطه (۷)** و **رابطه (۸)**، تابع درست‌نمایی توام در مدل *WCQAR* برای داده‌های کامل $\{y, v\}$ را می‌توان به‌صورت زیر نوشت:

$$L_c(\Psi|y, v) = \prod_{t=p+1}^{n} \left[\prod_{k=1}^{K} \frac{1}{\sqrt{2\pi}\sqrt{\theta_{2k}\sigma_k v_{tk}}} \; exp\left\{-\frac{(y_t - \eta_{tk} - \theta_{1k} v_{tk})^2}{2\theta_{2k}\sigma_k v_{tk}}\right\} \prod_{k=1}^{K} \frac{1}{\sigma_k} \; exp\left\{-\frac{1}{\sigma_k} v_{tk}\right\}\right], \qquad (۹)$$

که در آن v_{tk} متغیر آمیخته‌کننده را نشان می‌دهد. توجه داشته باشید که تابع درست‌نمایی در **رابطه (۹)** بر اساس توزیع نرمال تعریف شده است. باتوجه‌به وجود متغیرهای پنهان v_{tk} در این تابع درست‌نمایی، الگوریتم *SEM* به‌عنوان یک روش کارآمد برای محاسبه برآوردهای *ML* پارامترهای مدل *WCQAR* محسوب می‌شود. با درنظرگرفتن مزایای الگوریتم *SEM* نسبت به الگوریتم *EM* که در مقدمه به آن اشاره شد، در این مقاله از الگوریتم *SEM* به‌جای *EM* برای محاسبه برآوردهای مدل *WCQAR* استفاده خواهیم کرد.

۴- برآورد پارامترهای مدل WCQAR به روش SEM

در این بخش برای برآورد پارامترهای مدل از الگوریتم *SEM* استفاده می‌کنیم. برای مدل *WCQAR* مرتبه p، لگاریتم تابع درست‌نمایی برای داده‌های کامل به‌صورت زیر به‌دست می‌آید:

$$log[L_c(\Psi|y,v)] \propto \qquad (۱۰)$$

$$-\frac{(n-p)}{2}\sum_{k=1}^{K} log(\theta_{2k}) - \frac{3(n-p)}{2}\sum_{k=1}^{K} log\,\sigma_k - \frac{1}{2}\sum_{t=p+1}^{K}\sum_{k=1}^{K} log\,v_{tk}$$

$$-\sum_{t=p+1}^{n}\sum_{k=1}^{K}\frac{1}{\sigma_k}\left[\frac{(y_t-\eta_{tk})^2}{2\theta_{2k}}v_{tk}^{-1} + \frac{\theta_{2k}^2+2\theta_{1k}}{2\theta_{2k}}v_{tk} - \frac{\theta_{1k}(y_t-\eta_{tk})}{\theta_{1k}}\right].$$

مرحله *S*: در این مرحله ابتدا توزیع شرطی متغیر پنهان v_{tk} را به‌صورت زیر به‌دست می‌آوریم:

$$f(v_{tk}|y_t) \propto \frac{1}{\sqrt{v_{tk}}}\,exp\left\{-\frac{1}{2}\frac{(y_t-\eta_{tk}-\theta_{1k}v_{tk})^2}{2\theta_{2k}\sigma_k v_{tk}}\right\}.exp\left\{-\frac{1}{\sigma_k}v_{tk}\right\} \qquad (۱۱)$$

$$\propto \frac{1}{\sqrt{v_{tk}}}\,exp\left\{-\frac{1}{2}\left[\frac{(y_t-\eta_{tk})^2}{\theta_{2k}\sigma_k}v_{tk}^{-1} + \frac{\theta_{1k}^2+2\theta_{2k}}{\theta_{2k}\sigma_k}v_{tk}\right]\right\}$$

$$\sim GIG\left(\frac{1}{2}, \frac{(y_t-\eta_{tk})^2}{\theta_{2k}\sigma_k}, \frac{\theta_{1k}^2+2\theta_{2k}}{\theta_{2k}\sigma_k}\right),\ t=p+1,\ \dots,n,\ \ k=1,\ \dots,K$$

نماد $GIG(\lambda,\chi,\psi)$، $\chi>0$، $\psi>0$ نشان‌دهنده توزیع گاوسی وارون تعمیم یافته[1] با تابع چگالی زیر است:

$$f(x;\lambda,\chi,\psi) = \frac{\chi^{-\lambda}\left(\sqrt{\chi\psi}\right)^{\lambda}}{2K_\lambda\left(\sqrt{\chi\psi}\right)}x^{\lambda-1}exp\left\{-\frac{1}{2}(\chi x^{-1}+\psi x)\right\},\ x>0,$$

به‌طوری‌که $-\infty<\lambda<\infty$ همچنین نماد $K_\lambda\left(\sqrt{\chi\psi}\right)$ نشان‌دهنده تابع بسل تعدیل‌یافته است که به‌صورت زیر تعریف می‌شود:

$$K_\lambda\left(\sqrt{\chi\psi}\right) = \frac{1}{2}\int_0^\infty x^{\lambda-1}exp\left\{-\frac{1}{2}(\chi\psi)\{\chi-x^{-1}\}\right\}dx.$$

مرحله *M*: در این مرحله، مقدار v_{tk} را با استفاده از توزیع گوسی وارون تعمیم‌یافته که در مرحله *S* به‌دست آمد، شبیه‌سازی کرده و در **رابطه (۱۰)** جایگذاری کرده و لگاریتم تابع درست‌نمایی را نسبت به $\Psi=(\Phi,\sigma,\alpha)$ ماکزیمم می‌کنیم. برای ماکزیمم‌سازی، از **رابطه (۱۰)** نسبت به پارامترها مشتق گرفته و مساوی صفر قرار می‌دهیم. برای پارامتر مقیاس σ_k در تکرار h*th* داریم:

$$\frac{\partial Q}{\partial\sigma_k} = -\frac{3(n-p)}{2\sigma_k} + \frac{1}{\sigma_k^2}\sum_{t=p+1}^{n}\left[\frac{(y_t-\eta_{tk})^2}{2\theta_{2k}}v_{tk}^{-1} + \frac{\theta_{1k}^2+2\theta_{2k}}{2\theta_{2k}}v_{tk} - \frac{\theta_{1k}(y_t-\eta_{tk})}{\theta_{2k}}\right] = 0. \qquad (۱۲)$$

[1] Generalized Inverse Gaussian Distribution

لذا

$$\hat{\sigma}_k^{(h)} = \frac{2}{3(n-p)} \sum_{t=p+1}^{n} \left[\frac{(y_t - \eta_{tk})^2}{2\theta_{2k}} v_{tk}^{-1} + \frac{\theta_{1k}^2 + 2\theta_{2k}}{2\theta_{2k}} v_{tk} - \frac{\theta_{1k}(y_t - \eta_{tk})}{\theta_{2k}} \right], \quad (۱۳)$$

$$t = p+1, \ \dots, n, \ \ k = 1, \ \dots, K.$$

برای پارامتر α_{τ_k} در تکرار hth داریم:

$$\frac{\partial Q}{\partial \alpha_{\tau_k}} = \frac{\sum_{t=p+1}^{n} (y_t - y_t^T \Phi - \alpha_{\tau_k}) v_{tk}^{-1}}{\theta_{2k}\sigma_k} - \frac{(n-p)\theta_{1k}}{\theta_{2k}\sigma_k} = 0. \quad (۱۴)$$

بنابراین

$$\hat{\alpha}_{\tau_k}^{(h)} = \frac{\sum_{t=p+1}^{n} (y_t - y_t^T \hat{\Phi}^{(h-1)}) \, v_{tk}^{-1} - (n-p)\theta_{1k}}{\sum_{t=p+1}^{n} v_{tk}^{-1}}. \quad (۱۵)$$

برای پارامترهای اتورگرسیو Φ در تکرار hth داریم:

$$\frac{\partial Q}{\partial \Phi} = -\sum_{t=p+1}^{n} \sum_{k=1}^{K} \left[y_t \frac{v_{tk}^{-1}}{\theta_{2k}\sigma_k} y_t^T \right] \Phi + \sum_{t=p+1}^{n} \sum_{k=1}^{K} \left[\frac{v_{tk}^{-1}}{\theta_{2k}\sigma_k} y_t y_t - \frac{\theta_{1k+}\alpha_{\tau_k} v_{tk}^{-1}}{\theta_{2k}\sigma_k} y_t \right] = 0. \quad (۱۶)$$

رابطه (۱۶) را می‌توان به شکل ماتریسی به‌صورت زیر نوشت:

$$\sum_{t=p+1}^{n} y_t \sum_{k=1}^{K} \frac{v_{tk}^{-1}}{\theta_{2k}\sigma_k} y_t^T \cdot \Phi = \sum_{t=p+1}^{n} y_t \sum_{k=1}^{K} \frac{v_{tk}^{-1}}{\theta_{2k}\sigma_k} y_t - \sum_{t=p+1}^{n} y_t \sum_{k=1}^{K} \frac{\theta_{1k+}\alpha_{\tau_k} v_{tk}^{-1}}{\theta_{2k}\sigma_k}. \quad (۱۷)$$

بنابراین

$$(E^T W^{(h-1)} E)\Phi = E^T W^{(h-1)} Y - E^T D. \quad (۱۸)$$

از **رابطه (۱۸)**، برآورد Φ در تکرار hth به‌صورت زیر خواهد بود:

$$\hat{\Phi}^{(h)} = (E^T W^{(h-1)} E)^{-1} E^T W^{(h-1)} \bar{Y}, \quad (۱۹)$$

که در آن

$$\bar{Y} = Y - \left[W^{(h-1)} \right]^{-1} D^{(h-1)},$$

و

$$E^T = \begin{pmatrix} y_p & y_{p+1} & \cdots & y_{n-1} \\ y_{p-1} & y_p & \cdots & y_{n-2} \\ & & \vdots & \\ y_1 & y_2 & \cdots & y_{n-p} \end{pmatrix}, Y = \left(y_{p+1}, \ldots, y_n\right)^T,$$

$$D^{(h-1)} = \left(d_{p+1}^{(h-1)}, \ldots, d_n^{(h-1)}\right)^T, W^{(h-1)} = diag\left(w_{p+1}^{(h-1)}, \cdots, w_n^{(h-1)}\right),$$

$$w_t^{(h-1)} = \sum_{k=1}^{K} \frac{v_{tk}^{-1}}{\theta_{2k}\hat{\sigma}_k^{(h-1)}}, \; d_t^{(h-1)} = \sum_{k=1}^{K} \frac{\theta_{1k}+\hat{\alpha}_{\tau_k}^{(h-1)} v_{tk}^{-1}}{\theta_{2k}\hat{\sigma}_k^{(h-1)}},$$

$$t = p+1, \cdots, n, \quad k = 1, \cdots, K$$

است. حال مرحله‌های S و M را تا رسیدن به همگرایی مطلوب ادامه داده و برآوردها را به‌دست می‌آوریم. به‌طور خلاصه نحوه‌ی اجرای الگوریتم *SEM* برای مدل *WCQAR* را می‌توان به‌صورت زیر ارایه کرد:

گام ۱- مقادیر آغازین پارامترها را برابر $\Psi^0 = (\Phi^0, \sigma^0, \alpha^0)$ قرار می‌دهیم.

گام ۲- مقدار متغیر پنهان v_{tk} را با استفاده از توزیع گوسی وارون تعمیم‌یافته شبیه‌سازی می‌کنیم.

$$v_{tk} \sim GIG\left(\frac{1}{2}, \frac{(y_t-\eta_{tk})^2}{\theta_{2k}\sigma_k}, \frac{\theta_{1k}^2+2\theta_{2k}}{\theta_{2k}\sigma_k}\right), \; t = p+1, \ldots, n, \; k = 1, \ldots, K.$$

گام ۳- باتوجه‌به **گام ۲**، $\hat{\sigma}_k$ را به‌صورت زیر به‌دست می‌آوریم:

$$\hat{\sigma}_k^{(h)} = \frac{2}{3(n-p)} \sum_{t=p+1}^{n} \left[\frac{(y_t-\eta_{tk})^2}{2\theta_{2k}} v_{tk}^{-1} + \frac{\theta_{1k}^2+2\theta_{2k}}{2\theta_{2k}} v_{tk} - \frac{\theta_{1k}(y_t-\eta_{tk})}{\theta_{2k}}\right],$$

$$t = p+1, \ldots, n, \; k = 1, \ldots, K.$$

گام ۴- $\hat{\alpha}_{\tau_k}$ را به‌صورت زیر محاسبه می‌کنیم:

$$\hat{\alpha}_{\tau_k}^{(h)} = \frac{\sum_{t=p+1}^{n}\left(y_t - y_t^T\hat{\Phi}^{(h-1)}\right) v_{tk}^{-1} - (n-p)\theta_{1k}}{\sum_{t=p+1}^{n} v_{tk}^{-1}}, \qquad k = 1, \cdots, K.$$

گام ۵- با استفاده از مقادیر محاسبه‌شده در مراحل قبل، برآورد پارامترهای اتورگرسیو را به‌صورت زیر به‌دست می‌آوریم:

$$\hat{\Phi} = (E^T W^{(h-1)} E)^{-1} E^T W^{(h-1)} \bar{Y}.$$

گام ۶- گام ۲ تا **گام ۵** را T بار تکرار می‌کنیم و برآورد پارامترها را بعد از حذف نمونه‌های غیرموثر به‌دست می‌آوریم.

۵- مطالعات عددی

در این بخش، به‌منظور ارزیابی عملکرد روش پیشنهادی (*WCQAR-SEM*)، از مطالعات شبیه‌سازی مونت‌کارلو و یک مجموعه داده واقعی استفاده شده است. نتایج این روش با روش برآورد پارامترهای مدل *WCQAR* مبتنی بر الگوریتم *EM* (*WCQAR-EM*) و همچنین روش‌های کلاسیک *OLS* و *ML* مقایسه می‌شود.

۱-۵- شبیه‌سازی

برای ارزیابی عملکرد الگوریتم *SEM* و کارایی مدل *WCQAR-SEM* که در بخش‌های قبلی به آن اشاره شد، مطالعه شبیه‌سازی را با استفاده از سه مدل مختلف و با حجم نمونه $n = 100$ انجام می‌دهیم. جزئیات این مدل‌ها به شرح زیر است:

۱. $y_t = 0.5y_{t-1} + \varepsilon_t, \ \ t = 2, \cdots, 100$
۲. $y_t = 0.4y_{t-1} + 0.4y_{t-2} + \varepsilon_t, \ \ t = 3, \cdots, 100$
۳. $y_t = 0.3y_{t-1} + 0.05y_{t-2} + 0.1y_{t-3} + 0.1y_{t-4} + \varepsilon_t, \ \ t = 5, \cdots, 100$

در مدل‌های ارایه‌شده، برای ε_t توزیع‌های زیر در نظر گرفته می‌شود تا تاثیر انواع مختلف توزیع خطا بر عملکرد مدل بررسی شود:

۱. توزیع نرمال استاندارد، $N(0,1)$
۲. توزیع t با سه درجه آزادی، $t_{(3)}$
۳. توزیع لاپلاس استاندارد، $Laplace(0,1)$
۴. آمیخته دو توزیع نرمال، $0.1\ N(0,1) + 0.9\ N(0,9)$

در مدل *WCQAR*، مشابه [26]، تعداد سطوح چندکی برابر با $K = 5$ و $K = 9$ در نظر گرفته شد. علاوه‌بر این به‌منظور انتخاب مقدار مناسب K به‌صورت تجربی، مقادیر مختلف این پارامتر در بازه ۲ تا ۱۹ مورد بررسی قرار گرفت. نتایج این بررسی‌ها نشان داد که مقادیر $K = 5$ و $K = 9$، در مقایسه با سایر مقادیر، ضمن حفظ هزینه محاسباتی در سطحی قابل‌قبول، عملکرد مطلوب‌تری از نظر معیارهای ارزیابی ارایه می‌کنند. از این‌رو، این دو مقدار در شبیه‌سازی‌ها و تحلیل داده‌های واقعی مورد استفاده قرار گرفتند. به‌منظور ایجاد شرایط یکسان برای اجرای مدل‌های *WCQAR-SEM* و *WCQAR-EM*، در مرحله مقداردهی اولیه پارامترها در هر دو الگوریتم *SEM* و *EM*، از برآوردهای حاصل از روش‌های *OLS* یا *ML* استفاده می‌شود. همچنین، داده‌های ورودی، ساختار مدل، تابع درست‌نمایی، تعداد و محل سطوح چندکی، وزن‌های مربوط به چندک‌ها در هر دو روش یکسان در نظر گرفته شده است و تنها تفاوت میان آن‌ها در الگوریتم مورد استفاده برای برآورد پارامترها است. به‌منظور بررسی تغییرپذیری نتایج حاصل از شبیه‌سازی و افزایش اعتبار آن‌ها، فرایند شبیه‌سازی را ۱۰۰ بار تکرار می‌کنیم. در هر تکرار از شبیه‌سازی، مراحل S و M الگوریتم *SEM* را مشابه [27]، ۴۰۰۰ مرتبه تکرار می‌کنیم. سپس، ۲۰۰۰ نمونه اول را به‌عنوان دوره سوخت حذف می‌کنیم تا از تاثیر نمونه‌های اولیه نامعتبر جلوگیری شود. از ۲۰۰۰ نمونه باقی‌مانده برای استنباط درباره پارامترهای مدل استفاده می‌کنیم. در ادامه برای ۱۰۰ تکرار شبیه‌سازی، متوسط مقدار پارامترها (*Mean*)، اریبی (*Bias*) و مجذور میانگین مربعات خطا[1] برای پارامترهای مدل مورد نظر به‌صورت زیر به‌دست می‌آوریم:

$$Mean = \frac{1}{100}\sum_{i=1}^{100}\hat{\phi}_{i,j}\,, \qquad i = 1, \cdots, 100, \quad j = 1, \cdots, p,$$

$$RMSE = \sqrt{\frac{1}{100}\sum_{i=1}^{100}\left(\hat{\phi}_{i,j} - \phi_j^{True}\right)^2}.$$

مقادیر *Mean*، *Bias* و *RMSE* برای **مدل (۱)** تا **مدل (۳)**، به‌ترتیب در **جدول ۱** تا **جدول ۳** به‌ازای $K = 9$ ارایه شده است. این مقادیر به ما کمک می‌کنند تا دقت، اعتبار و پایداری مدل *WCQAR-SEM* را به‌طور جامع ارزیابی کنیم.

[1] Root Mean Square Error (RMSE)

جدول ۱- مقادیر میانگین، اریبی و RMSE برای پارامترها در مدل شبیه‌سازی اول به‌ازای $K = 9$.

Table 1– Mean values, bias, and RMSE for the parameters in the first simulation model for K=9.

خطا	پارامترها	$\phi_1 = 0.5$
$N(0,1)$	Mean	0.4767
	Bias	0.0233
	RMSE	0.0759
$t_{(3)}$	Mean	0.4653
	Bias	0.0347
	RMSE	0.0709
$Laplace(0,1)$	Mean	0.4710
	Bias	0.0290
	RMSE	0.0764
$0.1N(0,1) + 0.9N(0,9)$	Mean	0.4616
	Bias	0.0384
	RMSE	0.0791

جدول ۲- مقادیر میانگین، اریبی و RMSE برای پارامترها در مدل شبیه‌سازی دوم به‌ازای $K = 9$.

Table 2– Mean values, bias, and RMSE for the parameters in the second simulation model for K=9.

خطا	پارامترها	$\phi_1 = 0.4$	$\phi_2 = 0.4$
$N(0,1)$	Mean	0.3643	0.3561
	Bias	0.0357	0.0439
	RMSE	0.0476	0.0796
$t_{(3)}$	Mean	0.3657	0.3464
	Bias	0.0343	0.0536
	RMSE	0.0490	0.0770
$Laplace(0,1)$	Mean	0.4028	0.3547
	Bias	-0.0028	0.0453
	RMSE	0.0833	0.0769
$0.1N(0,1) + 0.9N(0,9)$	Mean	0.3671	0.3562
	Bias	0.0329	0.0438
	RMSE	0.0382	0.0892

با بررسی مقادیر میانگین (*Mean*) در **جدول ۱** تا **جدول ۳**، مشخص می‌شود که برآورد پارامترها به‌طور چشمگیری به مقادیر واقعی نزدیک است. همچنین، مقادیر کوچک *RMSE* نشان‌دهنده عملکرد مطلوب الگوریتم *SEM* در برآورد پارامترهای مدل *WCQAR-SEM* است و دقت بالای برآوردگرها را تایید می‌کند.

جدول ۳- مقادیر میانگین، اریبی و RMSE برای پارامترها در مدل شبیه‌سازی سوم به‌ازای $K = 9$.

Table 3– Mean values, bias, and RMSE for the parameters in the third simulation model for K=9.

خطا	پارامترها	$\phi_1 = 0.3$	$\phi_2 = -0.05$	$\phi_3 = 0.1$	$\phi_4 = 0.1$
$N(0,1)$	Mean	0.3183	-0.0909	0.0642	0.0846
	Bias	-0.0183	0.0409	0.0358	0.0154
	RMSE	0.0966	0.0857	0.0989	0.1060
$t_{(3)}$	Mean	0.2884	-0.0595	0.0884	0.0718
	Bias	0.0116	0.0095	0.0116	0.0282
	RMSE	0.0721	0.0725	0.0637	0.0713
$Laplace(0,1)$	Mean	0.2938	-0.0758	0.0927	0.0960
	Bias	0.0062	0.0258	0.0073	0.0040
	RMSE	0.0790	0.0762	0.0908	0.0665
$0.1N(0,1) + 0.9N(0,9)$	Mean	0.2714	-0.0525	0.0736	0.0648
	Bias	0.0286	0.0025	0.0264	0.0352
	RMSE	0.0849	0.0779	0.0765	0.0739

برای مقایسه‌ی کارایی مدل *WCQAR-SEM* با روش‌های *OLS*، *WCQAR-EM* و *MLE* از معیار خطای مدل (*ME*) که به‌صورت زیر تعریف می‌شود، استفاده می‌کنیم:

$$ME(\hat{\Phi}) = (\hat{\Phi} - \Phi^{true})^T(\hat{\Phi} - \Phi^{true}).$$

درواقع $ME(\hat{\Phi})$ را ۱۰۰ مرتبه تحت هر مدل محاسبه کرده و سپس متوسط مقدار و انحراف استاندارد آن‌ها به‌عنوان معیار مقایسه گزارش می‌شوند. میانگین مقادیر *ME* و انحراف استاندارد مربوطه برای مدل‌های *WCQAR-SEM* به‌ازای $K = 5$ با مدل‌های *WCQAR-EM*، *OLS* و *MLE* در **جدول ۴** گزارش شده است. همچنین مشابه **جدول ۴** میانگین $ME(\hat{\Phi})$ و انحراف استانداردهای آن برای مدل‌های *WCQAR-SEM* به‌ازای $K = 9$ با مدل‌های *WCQAR-EM*، *OLS* و *MLE* در **جدول ۵** گزارش شده است.

جدول ۴- متوسط مقادیر ME و انحراف استاندارد مربوطه تحت روش‌های مختلف برای $K = 5$.

Table 4– Mean ME values and corresponding standard deviations under different methods for K=5.

OLS	ML	WCQAR-EM	WCQAR-SEM	پارامتر	خطا	مدل شبیه‌سازی
0.0058	0.0059	0.0067	0.0066	mean	N(0,1)	1
0.0065	0.0066	0.0092	0.0086	sd		
0.0062	0.0059	0.0060	0.0055	mean	$t_{(3)}$	
0.0081	0.0075	0.0083	0.0081	sd		
0.0097	0.0098	0.0093	0.0091	mean	Laplace(0,1)	
0.0151	0.0150	0.0146	0.0141	sd		
0.0074	0.0075	0.0091	0.0087	mean	0.1N(0,1) + 0.9N(0,9)	
0.0097	0.0097	0.0143	0.0136	sd		
0.0236	0.0237	0.0279	0.0284	mean	N(0,1)	2
0.0268	0.0271	0.098	0.095	sd		
0.0183	0.0179	0.0153	0.0147	mean	$t_{(3)}$	
0.0190	0.0180	0.0136	0.0144	sd		
0.0216	0.0228	0.0201	0.0193	mean	Laplace(0,1)	
0.0253	0.0263	0.0263	0.0244	sd		
0.0207	0.0204	0.0203	0.0199	mean	0.1N(0,1) + 0.9N(0,9)	
0.0215	0.0209	0.0220	0.0181	sd		
0.0449	0.0453	0.0509	0.0611	mean	N(0,1)	3
0.0275	0.0291	0.0314	0.0459	sd		
0.0400	0.0419	0.0398	0.0361	mean	$t_{(3)}$	
0.0380	0.0383	0.0423	0.0404	sd		
0.0400	0.0408	0.0362	0.0368	mean	Laplace(0,1)	
0.0317	0.0318	0.0310	0.0321	sd		
0.0355	0.0347	0.0359	0.0357	mean	0.1N(0,1) + 0.9N(0,9)	
0.0227	0.0223	0.0229	0.0220	sd		

بررسی نتایج **جدول‌های ۴** و **۵** نشان می‌دهد که در شرایطی که خطاها از توزیع نرمال پیروی می‌کنند، اختلاف عملکرد چهار روش چندان چشمگیر نیست و در برخی سناریوها، خطاهای مدل روش‌های *ML* و *OLS* مقادیر اندکی کمتر از *WCQAR-SEM* و *WCQAR-EM* است. این نتیجه با مبانی نظری نیز سازگار است، زیرا در شرایط نرمال، روش‌های کلاسیک معمولا از کارایی مناسبی برخوردارند و انتظار برتری قابل‌توجه برای روش‌های استوار وجود ندارد. درمقابل، با تغییر توزیع خطا به توزیع‌های دم پهن مانند $t_{(3)}$ و لاپلاس استاندارد و توزیع آمیخته عملکرد روش *WCQAR-SEM* بهبود نسبی پیدا می‌کند. در اغلب این سناریوها، میانگین خطای مدل برآورد این روش کمتر یا هم‌تراز با سایر روش‌ها است که بیانگر پایداری بیشتر آن در برابر انحراف از فرض نرمال‌بودن و حضور داده‌های پرت است. همچنین، نتایج نشان می‌دهد که روش *WCQAR-EM* در اکثر سناریوها عملکردی بسیار نزدیک به *WCQAR-SEM* دارد و اختلاف بین این دو روش معمولا ناچیز است

این نتایج نشان می‌دهد که استفاده از الگوریتم *SEM* در روش پیشنهادی، کاهش محسوسی در دقت برآورد ایجاد نمی‌کند و از نظر کیفیت برآورد، نتایجی قابل‌مقایسه با روش *WCQAR-EM* ارایه می‌دهد. لازم به ذکر است که همان‌گونه که در مقدمه مقاله نیز اشاره شده است، انگیزه اصلی استفاده از الگوریتم *SEM* صرفا دستیابی به دقت بیشتر در برآورد پارامترها نبوده، بلکه بهره‌گیری از ویژگی‌های محاسباتی و الگوریتمی این روش نیز

مدنظر بوده است. از سوی دیگر، هر دو روش *WCQAR-SEM* و *WCQAR-EM* به مقادیر اولیه پارامترها حساس هستند. از این‌رو، به‌منظور ایجاد شرایط یکسان و قابل‌مقایسه، در هر دو روش، مقادیر اولیه پارامترها بر اساس برآوردهای حاصل از روش‌های *OLS* یا *ML* تعیین شده‌اند. همچنین، تعداد سطوح چندکی و سایر شرایط شبیه‌سازی در دو روش یکسان در نظر گرفته شده و تفاوت اصلی میان آن‌ها به الگوریتم مورد استفاده برای برآورد پارامترها مربوط می‌شود.

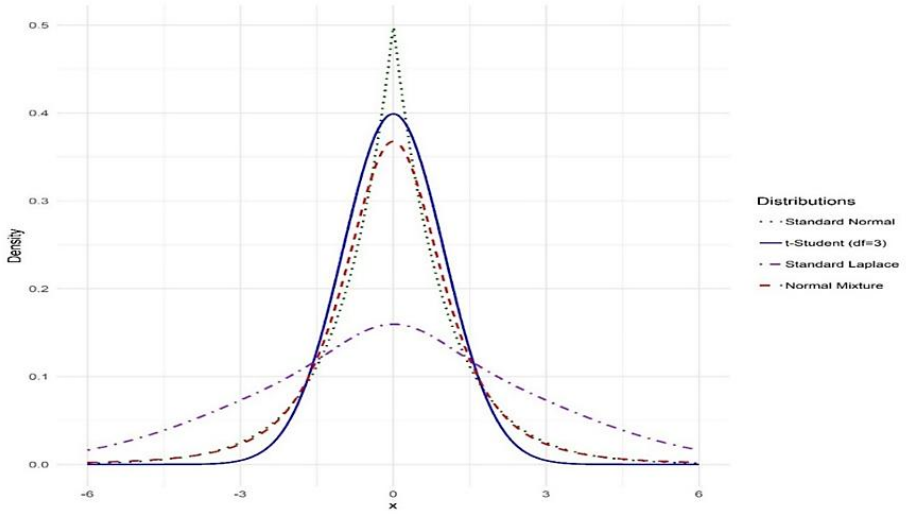
از نظر زمان محاسباتی نیز تفاوت مشخصی میان دو الگوریتم مشاهده شد. در روش *WCQAR-EM*، تعداد تکرارها تا زمان برآورده‌شدن معیار همگرایی ادامه می‌یابد و در آزمایش‌های انجام‌شده، الگوریتم معمولا پس از حدود ۱۰ تا ۱۵ تکرار به همگرایی می‌رسد. درمقابل، در الگوریتم *SEM*، باتوجه‌به ماهیت روش، تعداد تکرارها از پیش تعیین می‌شود. در مطالعه حاضر، زمان اجرای روش مبتنی بر *SEM*، بسته به سناریوی شبیه‌سازی، تقریبا بین ۵۰ تا ۷۰ ثانیه متغیر بوده است، درحالی‌که زمان اجرای روش مبتنی بر *EM* در شرایط مشابه تقریبا ۰/۵ تا ۱ ثانیه بوده است. این محاسبات بر روی سیستمی مجهز به پردازنده *12th Gen Intel(R) Core(TM) i5-1235U (1.30 GHz)* انجام شده است.

بنابراین، اگرچه روش مبتنی بر الگوریتم *SEM* از نظر زمان محاسباتی هزینه بیشتری نسبت به روش مبتنی بر *EM* دارد، نتایج شبیه‌سازی نشان می‌دهد که این افزایش هزینه محاسباتی با کاهش محسوسی در کیفیت برآورد همراه نبوده است. بنابراین، باتوجه‌به نتایج **جدول‌های ۴** و **۵**، می‌توان نتیجه گرفت که روش پیشنهادی ضمن حفظ دقت برآورد در سطحی مشابه روش مبتنی بر *EM*، از مزایای محاسباتی الگوریتم *SEM* نیز بهره‌مند است. همچنین، به‌منظور بررسی اینکه برتری مشاهده‌شده برای روش پیشنهادی از نظر آماری نیز قابل‌تایید باشد، مقایسه زوجی عملکرد روش *WCQAR-SEM*با روش‌های *WCQAR-EM*، *ML* و *OLS* انجام شد.

بدین منظور، ابتدا اختلاف معیار عملکرد بین روش *WCQAR-SEM* و هر یک از روش‌های مقایسه‌ای در تمامی تکرارهای شبیه‌سازی محاسبه گردید. سپس نرمال‌بودن اختلاف‌ها با استفاده از آزمون نرمالیتی بررسی شده است. در مواردی که فرض نرمال‌بودن اختلاف‌ها تایید شده، از آزمون tزوجی[1] و در غیر این‌صورت از آزمون ناپارامتری ویلکاکسون زوجی[2] برای مقایسه میانگین عملکرد روش‌ها استفاده شده است. به‌منظور ارایه تصویری روشن از نتایج، خروجی آزمون‌های آماری به‌صورت یک نمودار حرارتی[3] در **شکل ۲** گزارش شده است. در این نمودار، رنگ آبی بیانگر وجود اختلاف آماری معنی‌دار بین عملکرد روش *WCQAR-SEM* و روش مورد مقایسه و رنگ قرمز نشان‌دهنده عدم‌وجود اختلاف معنی‌دار در سطح خطای ۵% است. همان‌گونه‌که در شکل مشاهده می‌شود، در مدل‌های با مرتبه پایین و تعداد کم سطوح چندکی (مانند $AR(1)$ با $K = 5$)، اختلاف عملکرد روش *WCQAR-SEM* با روش‌های مقایسه‌ای در اغلب موارد از نظر آماری معنی‌دار نیست. با افزایش مرتبه مدل و همچنین افزایش تعداد سطوح چندکی، تعداد مواردی که اختلاف عملکرد از نظر آماری معنی‌دار است، به‌ویژه در مقایسه با روش‌های *OLS* و *ML*، به‌طور محسوسی افزایش می‌یابد.

این نتایج نشان می‌دهد که مزیت روش پیشنهادی در شرایط پیچیده‌تر و به‌ویژه در حضور توزیع‌های غیرنرمال، با شواهد آماری بیشتری همراه است. از سوی دیگر، تعداد اندک اختلاف‌های معنی‌دار در مقایسه با روش *WCQAR-EM* بیانگر آن است که عملکرد روش *WCQAR-SEM* و*WCQAR-EM*در بسیاری از سناریوهای مورد بررسی به یکدیگر نزدیک است. بااین‌حال، در اکثر موارد، روش *WCQAR-SEM* همچنان عملکرد رقابتی و در برخی شرایط عملکرد بهتری نسبت به سایر روش‌های مورد مقایسه ارایه می‌دهد.

[1] Paired t-test
[2] Wilcoxon Signed-Rank Test
[3] Heatmap



شکل ۱- نمودار توابع چگالی توزیع‌های $N(0,1)$، $t_{(3)}$، $Laplace(0,1)$ و $0.9N(0,9) + 0.1N(0,1)$.

Figure 1– Plot of the density functions for the N(0,1), $t_{(3)}$, Laplace(0,1), and 0.1N(0,1) + 0.9N(0,9) distributions.

جدول ۵- متوسط مقادیر ME و انحراف استاندارد مربوطه تحت روش‌های مختلف برای $K = 9$.

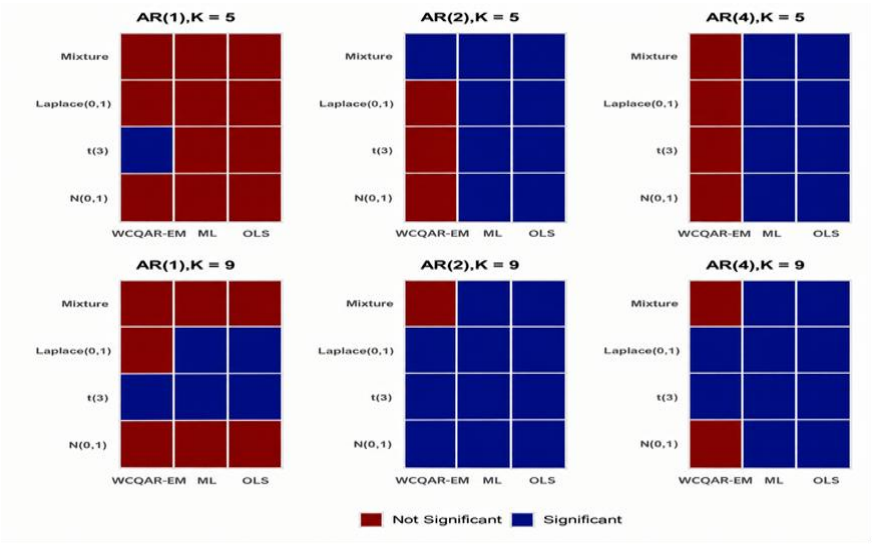
Table 5– Mean ME values and corresponding standard deviations under different methods for K=9.

OLS	ML	WCQAR-EM	WCQAR-SEM	پارامتر	خطا	مدل شبیه‌سازی
0.0083	0.0082	0.0095	0.0092	mean	N(0,1)	1
0.0125	0.0124	0.0135	0.0120	sd		
0.0097	0.0097	0.0085	0.0082	mean	$t_{(3)}$	
0.0149	0.0152	0.0109	0.0119	sd		
0.0103	0.0103	0.0089	0.0088	mean	Laplace(0,1)	
0.0151	0.0153	0.0140	0.0134	sd		
0.0112	0.0113	0.0105	0.0104	mean	0.1N(0,1) + 0.9N(0,9)	
0.0197	0.0198	0.0182	0.0178	sd		
0.0167	0.0180	0.0143	0.0139	mean	N(0,1)	2
0.0251	0.0256	0.0180	0.0155	sd		
0.0167	0.0180	0.0143	0.0134	mean	$t_{(3)}$	
0.0182	0.0200	0.0263	0.0259	sd		
0.0244	0.0244	0.0235	0.0210	mean	Laplace(0,1)	
0.0446	0.0437	0.0420	0.0382	sd		
0.0208	0.0211	0.0169	0.0152	mean	0.1N(0,1) + 0.9N(0,9)	
0.0221	0.0222	0.0231	0.0220	sd		
0.0552	0.0542	0.0567	0.0585	mean	N(0,1)	3
0.0353	0.0347	0.0392	0.0347	sd		
0.0372	0.0395	0.0348	0.0330	mean	$t_{(3)}$	
0.0322	0.0346	0.0297	0.0265	sd		
0.0464	0.0447	0.0419	0.0339	mean	Laplace(0,1)	
0.0500	0.0481	0.0490	0.0459	sd		
0.0456	0.0449	0.0383	0.0399	mean	0.1N(0,1) + 0.9N(0,9)	
0.0302	0.0340	0.0376	0.0359	sd		

باتوجه‌به نتایج ارایه‌شده در **جدول‌های ۴** و **۵**، می‌توان جمع‌بندی زیر را ارایه کرد:

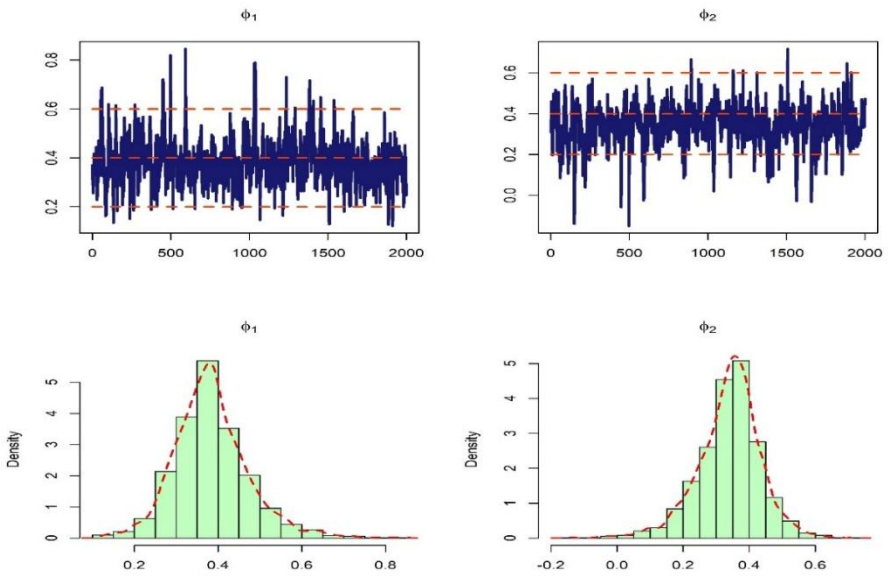
۱. با فاصله‌گرفتن توزیع خطا از توزیع نرمال و حرکت به‌سمت توزیع‌های دم پهن، مانند توزیع $t_{(3)}$ ، لاپلاس و توزیع آمیخته نرمال روش پیشنهادی *WCQAR-SEM* در بسیاری از سناریوها عملکرد مناسب‌تری نسبت به سایر روش‌های مورد مقایسه از خود نشان می‌دهد. این نتایج بیانگر آن است که روش *WCQAR-SEM* در برابر انحراف از فرض نرمال‌بودن خطاها و وجود مشاهدات دورافتاده، از استواری بیشتری برخوردار است.

۲. همچنین، با افزایش تعداد سطوح چندکی مورد استفاده (افزایش مقدار K از ۵ به ۹)، عملکرد روش *WCQAR-SEM* بهبود می‌یابد. این موضوع نشان می‌دهد که بهره‌گیری از تعداد بیشتری از سطوح چندکی، اطلاعات جامع‌تری از توزیع شرطی داده‌ها در اختیار مدل قرار داده و درنتیجه، دقت برآورد پارامترها را افزایش می‌دهد.



شکل ۲- نمودار حرارتی نتایج آزمون‌های زوجی مقایسه عملکرد روش WCQAR-SEM با روش‌های WCQAR-EM، ML و OLS. رنگ آبی نشان‌دهنده وجود اختلاف آماری معنی‌دار و رنگ قرمز بیانگر عدم وجود اختلاف معنی‌دار در سطح ۵%.

Figure 2– Heatmap of the results from pairwise performance comparisons between the WCQAR-SEM method and the WCQAR-EM, ML, and OLS methods. The blue color indicates the existence of a statistically significant difference and the red color indicates the absence of a significant difference at the 5% level.



شکل ۳- ردیف ۱: نمودارهای مسیر پارامترها در ۲۰۰۰ نمونه موثر برای $K = 9$ برای مدل شبیه‌سازی دوم زمانی‌که $\varepsilon_t \sim N(0,1)$. ردیف ۲: نمودارهای تابع چگالی و هیستوگرام ۲۰۰۰ نمونه موثر مربوطه.

Figure 3- Row 1: Path diagrams of parameters in 2000 effective samples for K=9 for the second simulation model when ε_t~N(0,1). Row 2: Density function graphs and histograms of 2000 relevant effective samples.

شکل ۳ شامل نمودارهای مسیر پارامترها در ۲۰۰۰ تکرار و همچنین نمودارهای هیستوگرام است. این شکل نشان می‌دهد که ۲۰۰۰ نمونه موثر به‌خوبی به توزیع هدف همگرا شده‌اند و بیشتر نمونه‌ها حول مقدار واقعی پارامترها قرار گرفته‌اند. ردیف دوم **شکل ۲**، نمودارهای هیستوگرام و منحنی‌های تابع چگالی نمونه‌های موثر را نمایش می‌دهد. این نمودارها علاوه‌بر ارایه یک توصیف بصری از شکل توزیع نمونه‌های موثر، همگرایی آن‌ها به توزیع هدف را نیز نشان می‌دهند. این نتایج حاکی از عملکرد و کارایی مناسب الگوریتم *SEM* در برآورد پارامترهای مدل *WCQAR* است.

۲-۵- تحلیل داده‌های برج تقطیر

یک فرض اساسی در طراحی نمودار کنترل شوهارت این است که مشاهدات به‌دست‌آمده از یک فرایند تولید، باید از لحاظ آماری مستقل باشند. باید توجه داشت که شرط استقلال، شرطی بسیار ایده‌آل است و در عمل به‌ندرت برقرار می‌شود. به تجربه ثابت شده است که اگر فاصله بین نمونه‌گیری‌ها کاهش یابد، مشاهدات دیگر از نظر آماری مستقل نخواهند بود و از خود، نوعی خود همبستگی نشان می‌دهند [6]. متاسفانه، در بعضی از فرایندهای تولیدی فرض وجود داده‌های مستقل حتی به‌صورت تقریبی نیز برقرار نیست. به‌عنوان‌مثال، می‌توان به فرایندهای شیمیایی که در آن‌ها مشخصه‌های محصول یا فرایند از همبستگی زیادی برخوردار هستند و به‌طور متوالی نیز اندازه‌گیری می‌شوند و یا به روش‌های آزمایش و بازرسی خودکار که در آن‌ها هر مشخصه‌ی کیفی محصول در فواصل زمانی معین اندازه‌گیری می‌شود، اشاره نمود [7]. بنابراین فرض پایه‌ای طراحی نمودار کنترل نقض می‌شود. همان‌طورکه در بخش مقدمه نیز اشاره شد، استفاده از نمودارهای کنترل شوهارت مرسوم برای مشاهدات خود همبسته، دو پیامد منفی به‌دنبال دارد: افزایش زنگ خطرهای اشتباهی و کاهش توانایی نمودار کنترل نسبت به کشف مشاهدات خارج از کنترل. این دو پیامد منفی نشان می‌دهند که وجود خود همبستگی در داده‌ها، تاثیر عمیقی بر نمودارهای کنترل شوهارت مرسوم می‌گذارد؛ بنابراین درنظرگرفتن خود همبستگی در مشاهدات و طراحی نمودارهای کنترل مناسب برای این نوع داده‌ها ضروری است. یکی از روش‌های طراحی نمودارهای کنترل برای داده‌های خودهمبسته، استفاده از مدل‌های سری زمانی است. در این روش ابتدا مدل مناسبی برای ساختار همبستگی بین داده‌ها تعیین شده و سپس از این مدل به‌منظور حذف خود همبستگی موجود در داده‌ها استفاده می‌گردد. درنهایت از نمودارهای کنترل مرسوم برای مانده‌های مدل که به نمودار کنترل مانده‌ها[1] معروف است، استفاده می‌شود.

در این بخش، روش پیشنهادی این مقاله برای تحلیل و پیش‌بینی داده‌های مربوط به یک فرایند شیمیایی به‌کار گرفته می‌شود. تیمر و همکاران [28] مطالعه‌ای را در یک کارخانه تولید مواد شیمیایی انجام دادند و هدف از انجام این بررسی، کنترل کیفیت محصول خروجی یک برج تقطیر با استفاده از نمودارهای کنترل برای متغیرهای انفرادی بود. لازم به ذکر است که محصول خروجی این برج تقطیر یکی از مواد اولیه بسیار مهمی است که در ساخت نوعی پلیمر به‌کار می‌رود، به همین دلیل پی‌بردن به وجود انحرافات با دلیل در فرایند تولید این محصول، از اهمیت ویژه‌ای برخوردار است. مشخصه‌ی کیفی مورد نظر، درصد وزن ماده‌ای به‌خصوص در خروجی برج تقطیر است. درصد وزن این ماده باید در حد معینی باشد و اگر تغییر قابل‌ملاحظه‌ای در آن به‌وجود آید، پلیمر ساخته‌شده از خروجی برج تقطیر کیفیت مطلوب را نخواهد داشت. پنجاه مشاهده با فاصله‌ی زمانی چهار دقیقه، از درصد وزن ماده به‌خصوص جمع‌آوری شده است. این داده‌ها در **جدول ۶** درج شده‌اند.

جدول ۶- داده‌های برج تقطیر.

Table 6- Distillation column data.

مشاهده	شماره نمونه	مشاهده	شماره نمونه
0.8017	26	0.7995	1
0.8017	27	0.8017	2
0.8017	28	0.7995	3
0.7995	29	0.8017	4
0.8017	30	0.8017	5
0.8017	31	0.8017	6
0.7995	32	0.7995	7

[1] Residuals Control Chart

جدول ۶- ادامه.

Table 6- Continued.

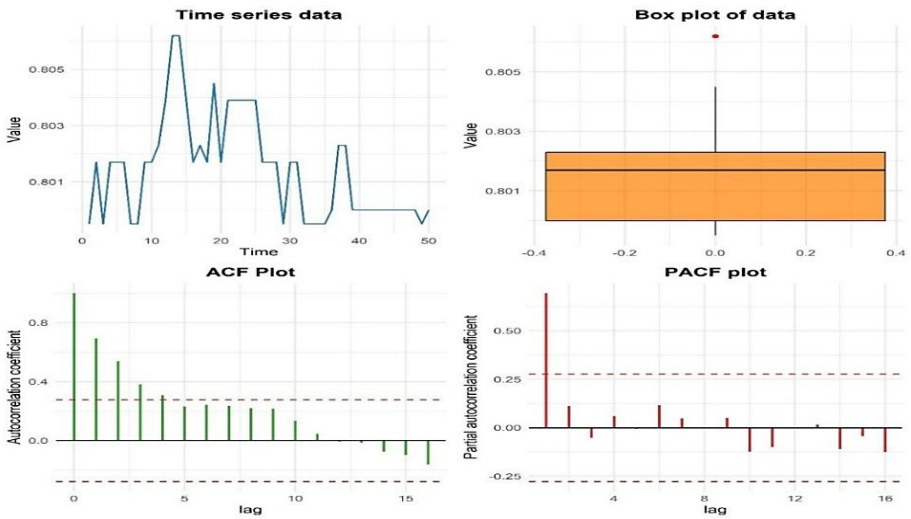
شماره نمونه	مشاهده	شماره نمونه	مشاهده
8	0.7995	33	0.7995
9	0.8017	34	0.7995
10	0.8017	35	0.7995
11	0.8023	36	0.8000
12	0.8039	37	0.8023
13	0.8062	38	0.8023
14	0.8062	39	0.8000
15	0.8039	40	0.8000
16	0.8017	41	0.8000
17	0.8023	42	0.8000
18	0.8017	43	0.8000
19	0.8045	44	0.8000
20	0.8017	45	0.8000
21	0.8039	46	0.8000
22	0.8039	47	0.8000
23	0.8039	48	0.8000
24	0.8039	49	0.7995
25	0.8039	50	0.8000

علاوه‌بر این، میانگین، میانه، انحراف معیار و چندک‌های ۲۵، ۵۰ و ۷۵ برای داده‌های برج تقطیر در **جدول ۷** خلاصه شده است. همچنین، نمودار جعبه‌ای به‌عنوان یک ابزار تحلیلی در **شکل ۴** نشان داده شده است که وجود مشاهدات دورافتاده در مجموعه داده‌ها را نمایان می‌سازد. باتوجه‌به استواری مدل *WCQAR-SEM* در مواجهه با داده‌های دارای مشاهدات دورافتاده و همچنین قابلیت روش برآورد مبتنی بر الگوریتم *SEM* در تحلیل روابط پیچیده، ترکیب این دو روش رویکردی بهینه و کارآمد برای تحلیل آماری داده‌های حاضر محسوب می‌شود.

جدول ۷- شاخص‌های آماری توصیفی داده‌های برج تقطیر.

Table 7– Descriptive statistical indicators for the distillation column data.

پارامتر	میانگین	میانه	انحراف معیار	چندک ۲۵	چندک ۵۰	چندک ۷۵
مقادیر	0.8014	0.8017	0.0018	0.8000	0.8017	0.8023



شکل ۴- نمودار جعبه‌ای، سری زمانی داده‌ها و نمودار توابع خودهمبستگی و خودهمبستگی جزیی داده‌های برج تقطیر.

Figure 4– Box plot, time series of the data, and plots of the autocorrelation and partial autocorrelation functions for the distillation column data.

نمودار سری زمانی مشاهدات در **شکل ۴** نشان داده شده است. به‌نظر می‌رسد داده‌ها روند مشخصی ندارند و بنابراین سری زمانی ایستا محسوب می‌شوند. همچنین، **شکل ۴** نمودارهای توابع خودهمبستگی و خودهمبستگی جزیی سری زمانی را به‌ترتیب نمایش می‌دهد. باتوجه‌به اینکه خودهمبستگی به‌تدریج کاهش یافته و خودهمبستگی جزیی نیز بعد از تاخیر ۱ معنی‌دار نیست، مدل اتورگرسیو مرتبه اول $AR(1)$ برای داده‌ها به‌صورت زیر در نظر گرفته می‌شود:

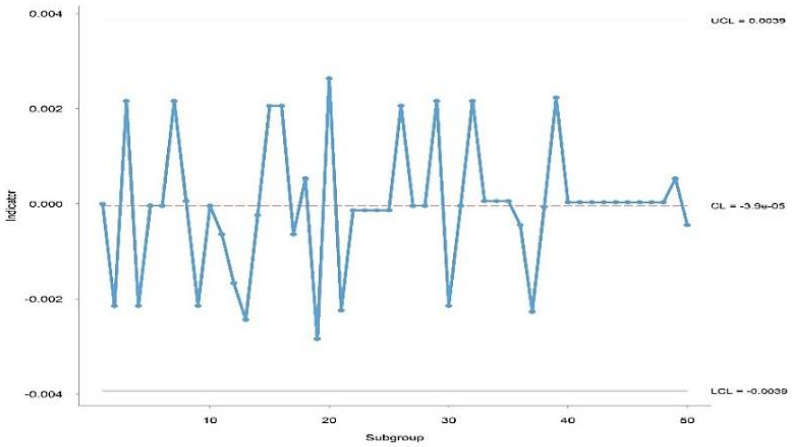
$$X_t = \phi_0 + \phi_1 X_{t-1} + \varepsilon_t, \;\; t = 2, \cdots, 50,$$

که در آن فرض می‌کنیم ε_t دارای توزیع $CAL(0, \sigma, \tau)$ است. برای برازش مدل *WCQAR-SEM* به داده‌ها، تعداد سطوح چندکی را برابر با ۵ و ۹ در نظر می‌گیریم. مقادیر آغازین پارامترها را بر اساس برآوردهای حاصل از روش *WCQAR-EM*، *OLS* یا *ML* تعیین می‌کنیم. در ادامه، مرحله‌های S و M را ۴۰۰۰ مرتبه تکرار می‌کنیم. سپس، ۲۰۰۰ نمونه اول را به‌عنوان نمونه‌های غیرموثر حذف کرده و از ۲۰۰۰ نمونه‌ی بعدی به‌عنوان نمونه‌های موثر برای استنباط درمورد پارامترها استفاده می‌کنیم. نتایج نشان می‌دهند که تقریبا پس از ۲۰۰۰ نمونه، همگرایی مطلوب برای همه‌ی پارامترها حاصل می‌شود. همچنین، به‌منظور تضمین بازتولیدپذیری نتایج و ایجاد امکان تکرار دقیق تحلیل‌های محاسباتی، بذر تصادفی *set.seed(123)* برای اجرای نهایی تحلیل داده‌ها در نظر گرفته شده است. همچنین بر اساس آنچه در [21] ارایه شده است، فاصله‌های اطمینان ۹۵% پارامترهای مدل بر اساس ویژگی نرمال‌بودن مجانبی برآوردگرهای حاصل از الگوریتم *SEM* محاسبه شدند. بدین منظور، واریانس هر برآوردگر از طریق توزیع نمونه‌گیری آن، حاصل از تکرار فرایند تولید داده‌ها و برازش مدل، برآورد و خطای استاندارد متناظر محاسبه شد. سپس، با استفاده از برآورد پارامترها و خطاهای استاندارد فاصله‌های اطمینان مجانبی ۹۵% به‌دست آمدند. **جدول ۸** شامل مقادیر برآورد پارامترها و اطلاعات مربوط به دقت برآوردها است که می‌تواند به تحلیل بهتر نتایج کمک کند و در **شکل ۵** نمودار کنترلی X برای مانده‌های مدل رسم شده است.

جدول ۸- میانگین مقادیر برآورد، خطای استاندارد و فاصله اطمینان ۹۵ درصدی برای داده‌های فرایند شیمیایی با استفاده از روش WCQAR به‌ازای $K = 5, 9$.

Table 8– Mean estimated values, standard errors, and 95% confidence intervals for chemical process data using the WCQAR method, for K = 5, 9.

ϕ_1	ϕ_0		K
0.9664	0.0268	Mean	5
0.0897	0.0718	Sd	
0.6912	-0.1938	2.5%	
1.2415	0.2476	97.5%	
0.9022	0.0782	Mean	9
0.1788	0.1431	Sd	
0.7020	-0.0823	2.5%	
1.1023	0.2388	97.5%	



شکل ۵- نمودار کنترلی X برای مانده‌های مدل.

Figure 5– X control chart for model residuals.

برای ارزیابی کارایی روش‌های مختلف، از پیش‌بینی بازگشتی خارج نمونه‌ای با مبدا زمانی t_0 استفاده می‌کنیم. این ارزیابی شامل روش پیشنهادی این مقاله و همچنین روش‌های *WCQAR-EM*، *OLS* و *ML* است. به این ترتیب، ابتدا مدل را بر اساس بخشی از داده‌ها (با حجم نمونه $n = 38$) برازش می‌دهیم و سپس یک گام بعدی را پیش‌بینی می‌کنیم. پس از آن، یک گام به جلو می‌رویم و داده جدید (با حجم نمونه $n = 39$) را قرار می‌دهیم و دوباره یک گام بعدی را پیش‌بینی می‌کنیم. این روند را تا زمان رسیدن به زمان $t = T$ یعنی $(n = 50)$ ادامه می‌دهیم. همچنین، همین فرایند برای پیش‌بینی دو، سه و چهار گام نیز انجام می‌شود. در مجموع، با این روش ۱۲ مورد پیش‌بینی تحت هر یک از روش‌ها انجام خواهد شد. برای ارزیابی دقت پیش‌بینی‌ها، مقادیر متوسط قدرمطلق خطای پیش‌بینی[1] و انحراف استاندارد میانگین خطای پیش‌بینی[2] را در مدل‌های *WCQAR-SEM*، *WCQAR-EM*، *OLS* و *MLE* را به‌صورت زیر محاسبه می‌کنیم:

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{X}_{t_0+i} - X_{t_0+i}^{true}\right| = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{e}_{t_0+i}\right|,$$

$$SDP = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\left|\hat{e}_{t_0+i}\right| - \mathrm{MAEP}\right)^2},$$

در اینجا، N تعداد پیش‌بینی‌ها است. به‌وضوح، هر چه مقادیر *MAP* و *SDP* تحت یک مدل کمتر باشد، کارایی آن مدل بیشتر است. مقادیر *MAP* و *SDP* برای مدل‌های مختلف در **جدول ۹** گزارش شده است. نتایج ارایه‌شده در **جدول ۹** نشان می‌دهد که با افزایش گام پیش‌بینی، مقدار خطای پیش‌بینی در اغلب روش‌ها افزایش می‌یابد که بیانگر کاهش نسبی دقت پیش‌بینی در گام‌های بلندتر است. مقایسه نتایج روش‌های *OLS*، *MLE* و *WCQAR-SEM*نشان می‌دهد که این سه روش در مجموع عملکرد نسبتا مشابهی دارند، بااین‌حال، روش *WCQAR-SEM* در بسیاری از گام‌های پیش‌بینی خطای کمتری داشته و از پایداری بیشتری برخوردار است. همچنین، در روش مورد بررسی، زمانی که مقدار $(K = 9)$ در نظر گرفته می‌شود، عملکرد پیش‌بینی به‌صورت نسبی بهبود می‌یابد که می‌تواند نشان‌دهنده مناسب‌تر بودن این مقدار از (K) برای ترکیب اطلاعات چندک‌های مختلف باشد. در مقابل، روش *WCQAR-EM* در مقایسه با سایر روش‌ها، مقادیر خطای بیشتری را در اغلب گام‌های پیش‌بینی نشان می‌دهد و درنتیجه عملکرد ضعیف‌تری دارد. نکته قابل‌توجه دیگر، روند تغییرات خطای روش *WCQAR-SEM* است، به‌گونه‌ای‌که با افزایش گام پیش‌بینی، افزایش خطای این روش نسبت به سایر روش‌ها محدودتر بوده و عملکرد آن از ثبات بیشتری برخوردار است. در مجموع، نتایج **جدول ۹** نشان می‌دهد که اگرچه افزایش گام پیش‌بینی موجب افزایش خطا در تمامی روش‌ها می‌شود، روش *WCQAR-SEM*، به‌ویژه در $(K = 9)$، از نظر دقت و پایداری پیش‌بینی عملکرد مطلوب‌تری نسبت به سایر روش‌های مورد مقایسه ارایه می‌دهد. همچنین، **جدول ۱۰** مقادیر پیش‌بینی‌شده توسط هر چهار روش را برای گام پیش‌بینی دوم ارایه می‌کند. **شکل ۶** نیز این مقادیر را به‌صورت نموداری نمایش می‌دهد تا امکان مقایسه و درک بصری بهتر نتایج پیش‌بینی فراهم شود.

جدول ۹- مقادیر MAP و SDP برای داده‌های فرایند شیمیایی تحت روش‌های مختلف.

Table 9– MAP and SDP values for chemical process data under various methods.

$WCQAR-SEM_{K=9}$	$WCQAR-SEM_{K=9}$	$WCQAR-EM_{K=9}$	MLE	OLS		h
0.0004	0.0004	0.0054	0.0006	0.0007	MAP	1
0.0007	0.0006	0.0040	0.0005	0.0005	SDP	
0.0006	0.0006	0.0090	0.0009	0.0011	MAP	2
0.0007	0.0008	0.0053	0.0006	0.0006	SDP	
0.0008	0.0008	0.0128	0.0012	0.0013	MAP	3
0.0008	0.0009	0.0852	0.0006	0.0006	SDP	
0.0009	0.0010	0.0161	0.0013	0.0014	MAP	4
0.0009	0.0009	0.0081	0.0006	0.0006	SDP	

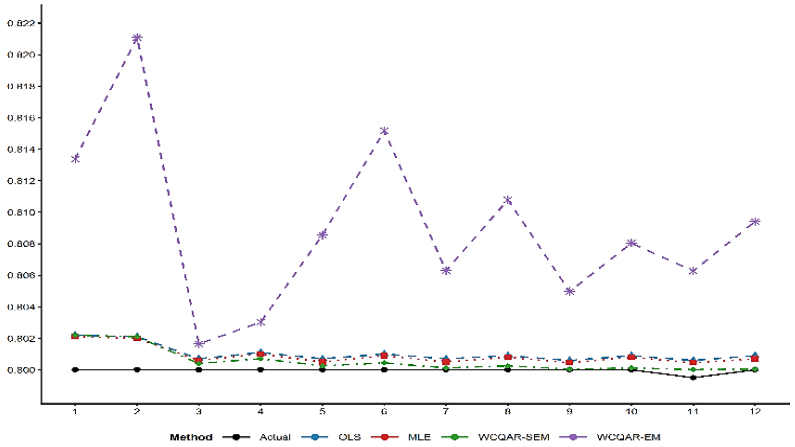
[1] Mean Absolute Prediction Error (MAPE)

[2] Standard Deviation Prediction (SDP)

جدول ۱۰- مقادیر پیش‌بینی‌شده در گام دوم تحت روش‌های مختلف به‌همراه مقادیر واقعی مشاهدات پیش‌بینی‌شده.

Table 10– Predicted values in the second step using various methods, along with the actual observed values.

WCQAR-EM	WCQAR-SEM	MLE	OLS	Actual
0.8133739	0.8021811	0.8021312	0.8022296	0.8000
0.8210922	0.8020898	0.8020224	0.8021851	0.8000
0.8016495	0.8003844	0.8005792	0.8006915	0.8000
0.8030274	0.8006838	0.8009543	0.8011313	0.8000
0.8085609	0.8002247	0.8005263	0.8006264	0.8000
0.8151712	0.8004199	0.8008752	0.8010358	0.8000
0.8062917	0.8001245	0.8004825	0.8005725	0.8000
0.8107840	0.8002394	0.8008085	0.8009551	0.8000
0.8049719	0.8000547	0.8004456	0.8005271	0.8000
0.8080396	0.8001073	0.8007515	0.8008859	0.8000
0.8062753	0.8000236	0.8004141	0.8004884	0.7995
0.8094096	0.8000459	0.8007022	0.8008260	0.8000



شکل ۶- مقادیر پیش‌بینی‌شده در گام دوم تحت روش‌های مختلف به‌همراه مقادیر واقعی مشاهدات پیش‌بینی‌شده.

Figure 6– Predicted values in the second step using various methods, along with the actual observed values.

۶- نتیجه‌گیری

در این مقاله، مدل *WCQAR* معرفی و مورد بررسی قرار گرفت و مزایای استفاده از آن بیان شد. برای برآورد پارامترهای این مدل، از الگوریتم *SEM* که یک روش تکرارشونده برای محاسبه برآوردهای *ML* است، استفاده شد. الگوریتم *SEM* که نسخه‌ای بهینه‌تر از الگوریتم *EM* محسوب می‌شود، توانایی بالاتری در حل مسایل پیچیده آماری دارد. یکی از جنبه‌های کلیدی این پژوهش، استفاده از توزیع *CAL* و آمیخته مقیاس آن است. فرض این توزیع برای خطاها امکان بهره‌گیری از الگوریتم *SEM* را برای برآورد پارامترها فراهم می‌کند. در ادامه، کارایی مدل از طریق مدل‌های شبیه‌سازی مختلف و با فرض توزیع‌های متنوع برای خطاها ارزیابی شد. همچنین، عملکرد مدل *WCQAR-SEM* با روش‌های *WCQAR-EM*، *ML* و *OLS* مقایسه گردید. نتایج شبیه‌سازی‌ها نشان داد که روش پیشنهادی این مقاله در برآورد پارامترها کارایی مناسب‌تری نسبت به روش‌های مورد اشاره دارد، به‌ویژه در شرایطی که توزیع خطاها دارای دم‌های پهن یا شامل مشاهدات دورافتاده باشد. علاوه‌بر شبیه‌سازی‌ها، برای ارزیابی کاربرد عملی این روش، مجموعه داده‌ای واقعی فرایند شیمیایی تحلیل شد. نتایج حاصل از برازش مدل پیشنهادی و مقایسه آن با روش‌های *WCQAR-EM*، *ML* و *OLS*نشان داد که روش ارایه‌شده علاوه‌بر عملکرد مناسب در مطالعات شبیه‌سازی در تحلیل داده‌های واقعی نیز عملکرد مناسبی دارد. این یافته‌ها نشان‌دهنده پتانسیل بالای روش پیشنهادی در تحلیل داده‌های واقعی و کاربردهای عملی است. ازجمله پیشنهادهای پژوهشی برای مطالعات آتی، می‌توان به درنظرگرفتن پارامترهای تاوان بر روی ضرایب مدل اتورگرسیو چندکی ترکیبی به‌منظور انتخاب متغیرها و برآورد همزمان پارامترها اشاره

کرد. همچنین، بررسی استواری روش پیشنهادی در سناریوهای مختلف، از جمله وجود خطاهای ناهمسانی واریانس، بدتعریفی مرتبه مدل اتورگرسیو و سایر شرایط پیچیده‌تر، می‌تواند در پژوهش‌های آینده مورد توجه قرار گیرد.

تشکر و قدردانی

نویسندگان مقاله مراتب قدردانی خود را از داوران محترم اعلام می‌دارند، بی‌شک دیدگاه ارزشمند آن‌ها در بهبود کیفیت مقاله نقش به‌سزایی داشته است.

منابع مالی

در طول انجام این پژوهش هیچ گونه کمک هزینه خاصی از هیچ موسسه، سرمایه‌گذار در بخش عمومی، خصوصی، تجاری یا غیرانتفاعی دریافت نشده است.

تعارض با منافع

هیچ تضادی در منافع در مورد انتشار این نسخه وجود ندارد و همه نویسندگان، نسخه نهایی ارسال‌شده را مشاهده و تایید کرده‌اند. نویسندگان تضمین می‌کنند که مقاله اثر اصلی آن‌ها بوده و پیش از این چاپ نشده و در حال حاضر تحت انتشار نیست.

منابع

[1] Tsay, R. S. (1988). Outliers, level shifts, and variance changes in time series. *Journal of forecasting*, *7*(1), 1–20. https://doi.org/10.1002/for.3980070102

[2] Deutsch, S. J., Richards, J. E., & Swain, J. J. (1990). Effects of a single outlier on ARMA identification. *Communications in statistics - Theory and methods*, *19*(6), 2207–2227. https://doi.org/10.1080/03610929008830316

[3] Kleiner, B., Martin, R. D., & Thomson, D. J. (1979). Robust estimation of power spectra. *Journal of the royal statistical society: Series b (Methodological)*, *41*(3), 313–338. https://doi.org/10.1111/j.2517-6161.1979.tb01087.x

[4] Peña, D. (2001). Outliers, influential observations and missing data. In *A course in time series analysis* (pp. 136–170). Wiley. https://doi.org/10.1002/9781118032978

[5] Tolvi, J. (1998). Outliers in time series: A review. University of Turku.

[6] Montgomery, D. C., & Mastrangelo, C. M. (1991). Some statistical process control methods for autocorrelated data. *Journal of quality technology*, *23*(3), 179–193. https://doi.org/10.1080/00224065.1991.11979321

[7] Harris, T. J., & Ross, W. H. (1991). Statistical process control procedures for correlated observations. *The canadian journal of chemical engineering*, *69*(1), 48–57. https://doi.org/10.1002/cjce.5450690106

[8] Koenker, R., & Xiao, Z. (2006). Quantile Autoregression. *Journal of the American statistical association*, *101*(475), 980–990. https://doi.org/10.1198/016214506000000672

[9] Zeng, Z., & Li, M. (2021). Bayesian median autoregression for robust time series forecasting. *International journal of forecasting*, *37*(2), 1000–1010. https://doi.org/10.1016/j.ijforecast.2020.11.002

[10] Tian, Y., Wang, L., Tang, M., Zang, Y., & Tian, M. (2020). Likelihood-based quantile autoregressive distributed lag models and its applications. *Journal of applied statistics*, *47*(1), 117–131. https://doi.org/10.1080/02664763.2019.1633285

[11] Bahmani, M. (2022). Linear quantile autoregressive model estimation using Stochastic EM algorithm. *Journal of advanced mathematical modeling*, *12*(3), 357-371. (**In Persian**). https://doi.org/10.22055/jamm.2022.39430.1988

[12] Zou, H., & Yuan, M. (2008). Composite quantile regression and the oracle model selection theory. *The annals of statistics*, *36*(3), 1108–1126. https://doi.org/10.1214/07-AOS507

[13] Aghamohammadi, A., & Bahmani, M. (2024). Bayesian weighted composite quantile regression estimation for linear regression models with autoregressive errors. *Communications in statistics - Theory and methods*, *53*(8), 2888–2907. https://doi.org/10.1080/03610926.2022.2150054

[14] Yuzhu, T., Liyong, W., Xinqian, W. U., & Maozai, T. (2019). Gibbs sampler algorithm of Bayesian weighted composite quantile regression. *Chinese journal of applied probability and statistics*, *35*(2), 178–192. https://dx.doi.org/10.3969/j.issn.1001-4268.2019.02.006

[15] Zhao, W., Lian, H., & Song, X. (2017). Composite quantile regression for correlated data. *Computational statistics & data analysis*, *109*, 15–33. https://doi.org/10.1016/j.csda.2016.11.015

[16] Zhao, Z., & Xiao, Z. (2014). Efficient regressions via optimally combining quantile information. *Econometric theory*, *30*(6), 1272–1314. https://doi.org/10.1017/S0266466614000176

[17] Celeux, G., & Diebolt, J. (1985). The SEM algorithm: A probabilistic teacher algorithm derived from the EM algorithm for the mixture problem. *Computational statistics quarterly*, *34*(2), 35-52. (**In French**). https://www.numdam.org/article/RSA_1986__34_2_35_0.pdf

[18] Diebolt, J., & Robert, C. P. (1994). Estimation of finite mixture distributions through Bayesian sampling. *Journal of the royal statistical society: Series b (methodological), 56*(2), 363–375. https://doi.org/10.1111/j.2517-6161.1994.tb01985.x
[19] Celeux, G., Chauveau, D., & Diebolt, J. (1996). Stochastic versions of the EM algorithm: An experimental study in the mixture case. *Journal of statistical computation and simulation, 55*(4), 287–314. https://doi.org/10.1080/00949659608811772
[20] Ip, E. H. S. (1994). *A stochastic EM estimator in the presence of missing data—Theory and applications*. https://statistics.stanford.edu/technical-reports/stochastic-em-estimator-presence-missing-data-theory-and-applications
[21] Feodor Nielsen, S. (2000). The stochastic EM algorithm: Estimation and asymptotic results. *Bernoulli, 6*(3), 457–489. https://doi.org/10.2307/3318671
[22] Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica, 46*(1), 33–50. https://doi.org/10.2307/1913643
[23] Tian, Y., Zhu, Q., & Tian, M. (2016). Estimation of linear composite quantile regression using EM algorithm. *Statistics & probability letters, 117*, 183–191. https://doi.org/10.1016/j.spl.2016.05.019
[24] Sriram, K., Ramamoorthi, R. V, & Ghosh, P. (2016). On Bayesian quantile regression using a pseudo-joint asymmetric laplace likelihood. *Sankhya a, 78*(1), 87–104. https://doi.org/10.1007/s13171-015-0079-2
[25] Kozumi, H., & Kobayashi, G. (2011). Gibbs sampling methods for Bayesian quantile regression. *Journal of statistical computation and simulation, 81*(11), 1565–1578. https://doi.org/10.1080/00949655.2010.496117
[26] Tian, Y., Lian, H., & Tian, M. (2017). Bayesian composite quantile regression for linear mixed-effects models. *Communications in statistics - Theory and methods, 46*(15), 7717–7731. https://doi.org/10.1080/03610926.2016.1161798
[27] Yang, F. (2018). A stochastic EM algorithm for quantile and censored quantile regression models. *Computational economics, 52*(2), 555–582. https://doi.org/10.1007/s10614-017-9704-6
[28] Timmer, D. H., Pignatiello Jr., J. J., & Longnecker, M. T. (2001). Applying an AR(1) cusum control chart to data from a chemical process. *Quality engineering, 13*(1), 107–114. https://doi.org/10.1080/08982110108918629