INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING, MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

Asymptotic Convergence Properties of Autoregressive Moving Average (Arma) Model Estimators

Dayo, Kayode Vincent, Olanrewaju Samuel Olayemi, Nasiru Mukaila Olakorede

Department of Statistics, University of Abuja.

DOI: https://doi.org/10.51583/IJLTEMAS.2026.15020000096

Received: 18 February 2026; Accepted: 23 February 2026; Published: 19 March 2026

ABSTRACTS

This study investigates the empirical convergence thresholds of the Gaussian Estimation Procedure (GEP), Generalised Least Squares (GLS), and Exact Maximum Likelihood (EML) for ARMA processes. While large-sample asymptotic equivalence is theoretically established, the specific data requirements for numerical reconciliation in higher-order models remain under-researched. Using a generalised Fibonacci-based sampling recurrence ($\mathrm{T_i} = \mathrm{T_{i-1}} + \mathrm{T_{i-2}}$) to determine the non-linear sample interval from $\mathrm{T} = 75$ to 850, estimator stability across six distinct data-generating processes was evaluated. The findings demonstrated a 'complexity-dependent convergence': while lower-order processes achieved numerical reconciliation at $\mathrm{T} = 75$, higher-order ARMA (2,2) specifications require $\mathrm{T} \geq 525$ to achieve harmonization. These results identify a critical transition zone where estimator choice become neural, providing a structural blueprint for selection based on modal dimensionality and available sample size.

Keywords: ARMA Models, Likelihood, Asymptotic Theory, Stationarity, Score, Approximation

INTRODUCTION

Several existing procedures exhibit distinctive asymptotic properties while estimating the parameters of time series models. This study examines the comparative asymptotic convergence properties of alternative estimation procedures for estimating the parameters of an Autoregressive Moving Average (ARMA) process.

$$\phi_0 x_t - \phi_1 x_{t-1} - \cdots - \phi_p x_{t-p} = \theta_0 \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}, \qquad [1.1]$$

Where $\phi_0 = \theta_0 = 1$, and ε_t is a white noise process, typically assumed to be independent and identically distributed (i.i.d.) with mean zero and variance δ^2. If the backshift operator B, is considered $B^k x_t$ denotes the random variable $\{x_t\}$ at lag k, such that, $B^k x_t = x_{t-k}$, then [1.1] can compactly be written as:

$$\Phi(B) x_t = \Theta(B) \varepsilon_t \qquad [1.2]$$

The asymptotic properties of the parameter estimators of [1,1] are intimately tied to the concepts of stationarity and invertibility in time series analysis. A process is strictly stationary if the joint distribution of $(X_{t1}, \ldots, X_{tk})$ is identical to that of $(X_{t1+h}, \ldots, X_{tk+h})$ for all h. In the context of linear ARMA models, however, weak stationarity (or covariance stationarity) is usually sufficient, requiring that the first two moments are time-invariant. This condition is met iff all roots of the characteristic equation $\Phi(z) = 0$, lie outside the unit circle in the complex plane. Where:

$$\Phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p, \qquad [1.3]$$

and

$$\Theta(z) = 1 + \theta_1 z + \theta_q z^q. \qquad [1.4]$$

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

Conversely, invertibility requires that the innovations, ε_t admit a convergent linear combination of current and past observations of $\{x_t\}$, which requires that all roots of the moving average polynomial $\Theta(z) = 0$, lie outside the unit circle. An ARMA process is stable if the AR polynomial satisfies $\phi(z) \neq 0$ for all $|z| \leq 1$; equivalently, in state-space (companion) form, stability holds if all eigenvalues of the companion matrix lie strictly inside the unit circle, i.e., formally, if ϕ is the companion matrix, $\rho\ (\phi) = max|\lambda_i(\phi)| < 1$.

Generally, patterns in a random set of data can be modelled to draw inferences about the process or population being observed. The science that deals with the collection and interpretation of these patterns is popularized in time series. As conceptualised in the recent literature on stochastic processes (Montgomery et al., 2024), a time series is not merely a data sequence but a realisation of a generating mechanism where temporary dependency dictates the information value of each successive observation. In a discrete-time series, these observations are recorded at a specified time interval, where $t = 1, 2, \dots$, the continuous-time series, on the other hand, observations are recorded continuously over some time interval, e.g., $t_0 = [0, 1]$.

The analysis of random variables, $\{x_{1\,,}\ x_2, \dots\}$ as outlined in (Kitagawa, 2020), is captured in the selection of a suitable probability model for the data. A complete probabilistic time series model for the sequential x_t would specify all the joint distributions of the random vector $(X_{t1}, \dots, X_{tk})$, $k = 1, 2, \dots, \infty$ or equivalently, all the probabilities $p[X_1 \leq x_1, \dots, X_n \leq x_n]$.

Exact Maximum Likelihood (EML) has long been the gold standard for ARMA estimation (Wei, 2006). Recent studies by Shumway, Robert H., and Stoffer (2025) have highlighted its computational limitations in high-dimensional settings. Consequently, there has been a resurgence in investigating Generalised Estimation Procedure (GEP), as seen in the work of (Athanasopoulos et al., 2024). An ARMA (p, q) process is a sequence of random variables $\{x_t\}$, capturing the present and future dynamics.

Several existing procedures exhibit distinctive asymptotic properties; recent comparative studies have focused primarily on large-scale datasets, often overlooking the nuanced transition from small to moderate sample sizes Chen & Yao, (2024). The focus of this research is to bridge this gap by examining the comparative asymptotic convergence properties of GEP, GLS, and EML estimators specifically within the Fibonacci sequential framework. By identifying the precise thresholds where these estimators achieve numerical reconciliation, particularly in higher-order ARMA (2, 2) specifications, this study establishes a structural blueprint for estimator selection in environments characterised by varying data availability.

The use of finitely parameterized models to characterize the behaviuors of time series data assumed to be generated by a stochastic process has received a wide coverage in both statistical and engineering literature (Salau, 2000). An important class of such stochastic models is the autoregressive moving average (ARMA) models. One reason for the great empirical importance of ARMA models is that every regular stationary process can be approximated with arbitrary accuracy by an ARMA process, and that only a finite number of (real-valued) parameters is needed to describe (up to second moment) an ARMA process.

Arma Representation

In modelling applications, such as forecasting and control, it is commonplace to represent such a process using a stationary and invertible autoregressive moving average (ARMA) model of the form:

$$\sum_{j=0}^{p} \Theta_j x_{t-j} = \sum_{j=0}^{q} \Phi_j \varepsilon_{t-j} \qquad [1.5]$$

Where ε_t is an unobservable stochastic disturbance, assumed to be white noise or more generally, a sequence of stationary martingale differences, so that, almost surely

$$\mathbb{E}\{\varepsilon_t / F_{t-1}\} = 0,\ \mathbb{E}\{\varepsilon_t^2 / F_{t-1}\} = \delta^2,\ \delta > 0 \qquad [1.6]$$

And symbol $\mathbb{E}$ is the expectation operator. Here, F_t is the $\sigma - algebra$ of events determined by the $\{\mathbb{E}(\varepsilon_s)\}$, $s < t$ and $\delta^2$is the prediction variance. The condition in [1.6] provides a natural relaxation of the model assumption that the innovation sequence, $\{\varepsilon_s\}$, are serially independent. For a complete exposition of these moment assumptions, see, for example, (Box et. al., 1994), (Tsay, 2019). The essential feature of [1.6] in our

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

context is that the classical theory still goes through, even though the innovation sequence, $\{\varepsilon_s\}$, is no longer serially independent and identically distributed. Thus, [1.6] is a natural condition with which the linear model [2.4] has little meaning.

These later assumptions ensure that x_t is realized from a stationary and invertible ARMA (p,q) process. Thus, [1.5] can be represented as an infinite autoregression.

$$\sum_{j=0}^{\infty} \varphi_j x_t = \varepsilon_t, \varphi_0 = 1 \quad [1.7]$$

Where the generating polynomial $\varphi(z) = \sum_{j=0}^{\infty} \varphi_j z^j \neq 0, |z| \leq 1, \varphi(z) = \Theta(z)^{-1}\Phi(z)$ and that $\sum_{j=0}^{\infty} |\varphi_j| < \infty$ since φ_j decreases to zero at a geometric rate. The significance of [2.5] in the analysis of linear time series has been thoroughly discussed in the literature, see for example, (Koreisha & Pukkila, 1990). The assumption $\mathbb{E}\{\varepsilon_t\} = 0$ means that in practice, the data will need to be mean-corrected; however, since that adjustment will not affect the asymptotic results we shall obtain later, it is ignored for simplicity. The integers, p, q (or other integers introduced to replace them), will be called the order of the system. The coefficients, $\Phi_j j = 1, 2, \dots, p;\ \Theta_j j = 1, 2, \dots, q$, specifying model [2.4] will be, called system parameters.

Interpreting B as a unit backshift operator, that is $Bx_t = x_{t-1}$ and $B\varepsilon_t = \varepsilon_{t-1}$, [2.4] can be succinctly rewritten in the form of [1.2], where $\Phi(B) = \sum_{j=0}^{p} \phi_j B^j$ and $\Theta(B) = \sum_{j=0}^{q} \theta_j B^j, \phi_j = \theta_j = 1$. The generating polynomial (or z-transforms), that is, $\Phi(z) = \sum_{j=0}^{p} \phi_j z^j$ and $\Theta(z) = \sum_{j=0}^{q} \theta_j z^j$ defined in the convention of Box, Jenkins & Reinsel (1994), are assumed to be relatively prime and also satisfy:

$$\Phi(z) \neq 0, \Theta(z) \neq 0, |z| \leq 1 \quad [1.8]$$

LITERATURE REVIEW

The estimation of Autoregressive Moving Average (ARMA) models is grounded in classical asymptotic theory. Foundational results by Whittle, (1953) and Walker, (1964) established that, under regularity conditions, likelihood-based and Least Squares type estimators are consistent and asymptotically normal, and converge to the same probability limit. This asymptotic equivalence underpins the conventional preference for Exact Maximum Likelihood (EML), which fully exploits the innovation covariance structure (See William W. S. Wei 2006).

However, asymptotic equivalence does not imply finite-sample equivalence. In moderate samples, particularly in higher-order ARMA specifications, the curvature of the likelihood surface may exhibit multinormality or near-flat regions, especially when parameters approach the boundary of invertibility or stationarity regions. As discussed in Time Series and Its Applications, numerical evaluation of the likelihood function becomes increasingly sensitive to nonlinear root configurations.

This distinction between asymptotic theory and finite-sample realization is critical. While asymptotic efficiency rankings are well established, the rate at which estimators reconcile numerically as sample size increases remains less precisely characterized. In practice, forecasters operate in finite samples, and the transition from estimator divergence to effective equivalence may be nonlinear and model-dependent.

Monte Carlo investigation of ARMA estimation traditionally evaluates performance at a small set of discrete sample sizes. Early simulation studies, such as Ansley & Newbold, (1980) and Burnside, (1994), provide important comparative insights, yet their experimental designs typically contrast widely spaced values of T. Such binary or coarse benchmarking implicitly assumes smooth convergence toward asymptotic equivalence.

For higher-order models such as ARMA (2, 2), this assumption may be restrictive. The parameter-to-sample bias analytical treatments of finite-sample properties (e.g., M. O. Salau 2001; Zinde-walsh, (1994) indicate that estimator dispersion may decay at a rate slower than asymptotic approximations suggest, particularly when roots lie close to the unit circle.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

METHODOLOGY

This section presents some procedures for estimating the parameters of the postulated ARMA model. For our exposition, the integers (order) p and q in [2.1] are assumed known, and after observing $x_t, t = 1, 2, \dots, T$, the system parameters, Φ_j, and δ^2 are to be estimated. Thus, this section is devoted to the straightforward part of the modelling procedure, namely, the estimations, for fixed values p and q of the parameters $\Phi = (\phi_1, \phi_2, \dots, \phi_p)$; $\theta = (\theta_1, \theta_2, \dots, \theta_q)$ are white noise variance δ^2. It will be assumed throughout that the data have been adjusted for the means. A variety of estimation procedures have been suggested in the literature (Wei, 2006). For ease of presentation, we have organized our discussion in this section into three subsections: the Gaussian Estimation procedure, the Generalised Least Squares method, and the Exact Maximum Likelihood technique. First, we introduce some notations, employing the conventions adopted in Salau (1998), we set for any integer $\xi_r(B)' = (B^1, B^2, \dots, B^r)$ are the symbol $\otimes$ has been used (Kronecker) matrix product. To fix ideas, we introduce α as a general parameter vector, so that:

$$\alpha = \left(\phi_1, \phi_2, \dots, \phi_p : \theta_1, \theta_2, \dots, \theta_q\right). \qquad [3.1]$$

Gaussian Estimation Procedure

In ARMA theory, the Gaussian Estimation Procedure refers to conditional Gaussian Likelihood, where initial observations are conditional upon. In R, the procedure is implemented as conditional Sum of Squares (CSS) or CSS-ML.

Under the standard assumption that the innovations ε_t of an ARMA process, are independently distributed as Gaussian random variables with mean zero and variance δ^2, the probability density function of each innovation is given by

$$g(\varepsilon_t) = \frac{1}{\sqrt{2\pi\delta^2}} exp\left\{-\frac{\left(\varepsilon_t^2\right)^2}{2\delta^2}\right\}, -\infty < x < \infty \qquad [3.2]$$

This assumption provides a natural justification for least squares and likelihood-based estimation procedures, since minimizing the sum of squared residuals is equivalent to maximizing the Gaussian likelihood. Hardin & Hilbe (2021) has very reliable discussion on GEP .

Generalised Least Squares Method

In the standard linear model, estimating the parameters of an ARMA process may be approached via least squares or likelihood-based procedures under the assumption of Gaussian innovations. Let the linear model be given as:

$$y = X\beta + \varepsilon, \varepsilon \sim N_n(0, \delta^2 I_n), \qquad [3.3]$$

For which the ordinary least squares estimation procedure is obtainable

$$\beta = (X'X)^{-1}X'Y. \qquad [3.4]$$

More generally, when the distribution vector satisfies $\varepsilon \sim N_n(0, \Sigma)$, with a symmetric positive definite covariance matrix Σ, the Generalised Least Squares estimator provides an efficient gain by accounting for the covariance structure included by serial dependence.

In-depth discussion on this subject can be seen in the approaches of (Hamilton, 1994), (Koutsoyiannis, 2009), (Greene, 2024), among others.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

Exact Maximum Likelihood Technique

The Exact Maximum Likelihood estimation of ARMA models is derived by constructing the joint density of the conditional innovations on initial observations. Under the assumption of Gaussian and independent innovations, the conditional likelihood for observations $t = k + 1, \ldots, T.$ is given as:

$$f(\varepsilon_{k+1}, \ldots, \varepsilon_T) = \left(\frac{1}{2\pi\delta_\varepsilon^2}\right)^{(T-k)/2} Exp\left(-\frac{1}{2\delta_\varepsilon^2}\left(\sum_{t=k+1}^{T} \varepsilon_t^2\right)\right) \quad [3.5]$$

Maximization of this likelihood yields the exact maximum likelihood estimators of the ARMA parameters, see (Olajide et al., 2012), (Box et at., 2015), and (Shumway & Stoffer, 2025).

Stability Consideration in Arma

To introduce the notion of invertibility (stability), consider the MA (1) process,

$$y_t = \varepsilon_t + \theta\varepsilon_{t-1} \quad [3.6]$$

The only non-zero autocorrelation associated with this process is

$$\rho_1 = \frac{\theta}{(1+\theta^2)}. \quad [3.7]$$

Suppose that θ is replaced with θ^{-1}, then [3.4] becomes

$$\rho_1 = \frac{\theta^{-1}}{(1+\theta^{-2})}. \quad [3.8]$$

Defining [3.3] in terms of the back shift operator, B,

$$y_t = (1 + \theta B)\varepsilon_t \quad [3.9]$$

By inverting [3.6], we have:

$$\varepsilon_t = (1 + \theta B)^{-1} y_t \quad [3.10]$$

Expressing ε_t as a linear filter of y_t and taking a formal power series expansion of the term, $(1 + \theta B)^{-1}$, gives :

$$\varepsilon_t = \sum_{j=0}^{\infty} \theta^j y_{t-j} \quad [3.11]$$

The easiest way to obtain the autocorrelation function in a specific class is to express y_t as a general linear process.

$$y_t = \{\phi(B)\}^{-1}\theta(B)\varepsilon_t \quad [3.12]$$

$$y_t = \left(\sum_{j=1}^{\infty} a_j\theta_j\right)\varepsilon_t \quad [3.13]$$

The coefficient a_j being determined by a formal power series expansion of $\{\phi(B)\}^{-1}$.

Thus if ε_t is thought of as being determined by the present and the past y_t, it is noticeable that the remote past has a vanishingly small influence if and only if $-1 < \theta < 1$.

In a similar discussion to (Salau, 2000), a simple adaptation of the technique in spectral analysis can be used to derive the second-order properties of an ARMA process y_t defined equivalently by [2.4] or [1.2]. It can immediately be deduced from [3.8] that the spectrum, $f(\omega)$, must satisfy:

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

$$\left|\phi\left(e^{i\omega}\right)\right|^2 f(\omega) = \frac{\delta^2}{2\pi}\left|\theta\left(e^{i\omega}\right)\right|^2 \quad [3.14]$$

Thus,

$$f(\omega) = \frac{\delta^2}{2\pi}\left|\theta\left(e^{i\omega}\right)\right|^2\left|\phi\left(e^{i\omega}\right)\right|^{-2} \quad [3.15]$$

$$f(\omega) = \frac{\delta^2}{2\pi}\left[\left\{1+\sum_{i=1}^{p}\theta_i cos(i\omega)\right\}^2 + \left\{\sum_{i=1}^{p}\theta_i sin(i\omega)\right\}^2\right] X \left[\left\{1-\sum_{j=1}^{q}\phi_j cos(j\omega)\right\}^2 + \left\{\sum_{j=1}^{q}\phi_j sin(j\omega)\right\}^2\right]^{-1} \quad [3.16]$$

Assuming that y_t is stationary, the form of $f(\omega)$ suggests that the values of $\phi_j$are critical in determining stationarity, and this is indeed the case. The precise condition is the same as for the autoregressive process $\phi(B)y_t = \varepsilon_t$, that all roots of $\phi(z) = 0$ must be greater than 1 in absolute value. More specifically, from [4.10], the stationarity condition follows from the requirement that $\left|\phi\left(e^{i\omega}\right)\right|^2$ be bounded away from zero for all $\omega\, \epsilon |-\pi,\pi|$, which equivalently satisfies that all roots of $\phi\ (z) = 0$ lying outside the unit circle.

Theorem I:

Let $f(\omega) = \frac{\delta^2}{2\pi}\left|k_0\left(e^{i\omega}\right)\right|^2$ be the true spectral density of y_t and, $k_0(z) = \alpha_0 z^{-1}\beta_0(z)$ under the stationarity and invertibility conditions, $k_0(z)^{-1}y_t = \sum_{j\geq 0}\phi_{0,j}y_{t-1}$. Then there exists an integer N and a constant k dependent on $f(\omega)$ such that $h \geq N(T) = \frac{logT}{-2log\,\rho_0}$, and the truncated autoregressive estimator satisfies:$\sum_{k=1}^{h}\left|\hat{\phi}_{h,k} - \phi_{0k}\right| = o\left(T^{-\frac{1}{2}}\right)$, where ρ_0 is the modulus of the zeros of $\beta_0(z)$ nearest$|z| = 1$.

Proof:

By Baxter's inequality (Baxter, 1962), $\sum_{k=1}^{h}\left|\phi_{h,k} - \phi_{0,k}\right| \leq \sum_{k\geq h+1}\left|\phi_{o,k}\right|$ [3.17]

Obtaining the bound for the term on the RHS of [4.12], and by the invertibility property of the stationary ARMA process, ρ_0 satisfies the inequality $0 < \rho_0^{-1} < 1$. The roots of the power series expansion of $\alpha_0(z)^{-1}$ shows that $\left|\phi_{0,j}\right|$ decreases geometrically at a rate determined by ρ_0. That is $\left|\phi_{0,j}\right| \leq k\rho_0^{j}$. For large T however, $h \geq N(T)$. We may now write that:

$$\rho_0^h = exp(h\ log\rho_0) \quad [3.18]$$

$$\rho_0^h = exp\left\{\frac{\log T}{-2\log\rho_0} X\ \log\rho_0\right\} \quad [3.19]$$

$$\rho_0^h = T^{-\frac{1}{2}} \quad [3.20]$$

Hence, it is concluded that [4.12] above is bounded by $kT^{-\frac{1}{2}}$.

Model Selection

Correct specification of the ARMA (p, q) order is a prerequisite for meaningful estimator comparison. Following the iterative model identification procedure described in (Box et al., 2015), and recent developments in automated order selection (Lin et al., 2024), the determination of p and q precedes parameter estimation.

Underfitting induces asymptotic misspecification bias, as estimators converge to pseudo-true values that fail to represent the full dynamic structure. Overfitting, by contrast, increases estimators variance and may introduce near-cancelation between autoregression and moving average roots, compromising numerical stability through weak identification.

Information criteria, such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) provide a likelihood-based selection rule.

$$AIC = -2\ell(\widehat{\Theta}) + 2k \quad [3.21]$$

$$BIC = -2\ell(\widehat{\Theta}) + kln(n) \quad [3.22]$$

Where $k = p + q$. The asymptotic properties of these criteria diverge significantly. The BIC is a consistent selector when the true model is contained within the candidate set (the probability of selection approaches 1 as $n \to \infty$), whereas AIC is asymptotically efficient for prediction but not consistent, whether the maintained specification reflects the true data-generating process.

Sampling Framework

To trace estimator reconciliation across increasing sample sizes, this study employs a Fibonacci-type sequential sampling selection defined recursively by:

$$T_i = T_{i-1} + T_{i-2} \quad [3.23]$$

This nonlinear expression produces dense convergence in smaller samples, where bias and curvature effects are strongest, while gradually widening intervals as convergence stabilizes. Such adaptive spacing enables more precise identification of the transition region between finite-sample divergence and effective asymptotic equivalence.

Recent discussions in stochastic process and time series estimation literature (e.g., Franq & Zakoian, 2024; Zhang and Ling, 2024; Shumway & Stoffer, 2025) emphasized the importance of examining convergence behaviour beyond fixed sample benchmarks. The present design operationalizes this principle within a structural framework.

Law of Large Numbers

The most fundamental requirement for an estimator, as discussed in (Hamilton, 1994), is the property that the estimator converges to the true parameter value as more data are collected (consistency). Specifically, for a stationary and ergodic ARMA process $\{x_t\}$ with mean μ.

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} x_t = \mu \; a.s. \quad [3.24]$$

The consistency of an estimator is typically established using the concept of uniform convergence. See more on uniform convergence in (Sanchez, 2010).

The normalized log-likelihood $T^{-1}\ell(\theta)$ converges uniformly in probability to a non-stochastic function $L(\theta)$. For ARMA models, the compactness of the parameter space Θ is usually satisfied by restricting the roots of $\phi(z)$ and $\theta(z)$ to lie in a closed set outside the unit circle, ensuring stationarity and invertibility.

Equivalence of Estimators

Theorem II

Let $\{x_t\}$ be a causal and invertible ARMA (p. q) process with true parameter vector

$$\alpha = \left(\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q\right)', \quad [3.25]$$

where the autoregressive and moving-average polynomials $\phi(z)$ and $\theta(z)$ have no common zeros, and $\{\varepsilon_t\}\sim iid(0, \delta^2)$. Supposed that $\hat{\alpha}$ is a preliminary estimator of $\left(\alpha_1, \dots, \alpha_{p+q}\right)'$ such that $\sqrt{T}\|\hat{\alpha} - \alpha\| = o_p(1)$, and that $\hat{\alpha}$ is the estimator constructed from α_0 as outlined in the estimation procedures of Section Three earlier.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

If $\hat{\alpha}_{\text{GEP}}$, $\hat{\alpha}_{\text{GLS}}$, and $\hat{\alpha}_{\text{EML}}$ denote the estimators obtainable via the Gaussian Estimation Procedure, Generalized Least Squares, and Exact Maximum Likelihood, respectively, then

i. $T^{\frac{1}{2}} \|\hat{\phi}_{GEP} - \hat{\phi}_{GLS}\| = o_p(1)\ a.s.$

ii. $T^{\frac{1}{2}} \|\hat{\phi}_{GLS} - \hat{\phi}_{EML}\| = o_p(1)\ a.s.$ [3.26]

iii. $T^{\frac{1}{2}} \|\hat{\phi}_{EML} - \hat{\phi}_{GEP}\| = o_p(1)\ a.s.$

Here $\|.\|$ denotes the Euclidean norm. Since all norms on a finite-dimensional space are equivalent, the choice of norm does not affect the asymptotic order.

Proof:

Each of the estimators $\hat{\alpha}_{\text{GEP}}$, $\hat{\alpha}_{\text{GLS}}$, and $\hat{\alpha}_{\text{EML}}$ may be expressed as a solution to estimating equations of the form

$$S_T(\alpha) = 0 \qquad [3.27]$$

Where $S_T(\alpha)$ denotes the corresponding score or normal equations.

By construction, these estimating equations coincide up to terms of order $o_p\left(T^{-\frac{1}{2}}\right)$ when evaluated in a $T^{-\frac{1}{2}}$-neighbourhood of the true parameter α. Expanding each estimator around the preliminary estimator $\hat{\alpha}$ using a first-order Taylor expansion yields

$$\sqrt{T}\|\hat{\alpha} - \alpha\| = \left[\frac{1}{T}\frac{\partial S_T(\alpha)}{\partial \alpha'}\right]^{-1} \frac{1}{\sqrt{T}} S_T(\alpha) + o_p(1) \qquad [3.28]$$

Under the assumed regularity conditions (i.e., causality, invertibility, finite fourth moment and non-singularity of information matrix), the leading terms of the expressions are identical for GEP, GLS, and EML. Consequently, their pairwise differences satisfy:

$$\sqrt{T}\|\hat{\alpha}_i - \hat{\alpha}_j\| = o_p(1). \qquad [3.29]$$

$i, j \in (\text{GEP}, \text{GLS}, \text{EML})$, which establishes asymptotic equivalence.

Numerical Application

The procedure for various estimation procedures for the investigation is applied to the simulated data points. For each of these processes, the sample size T was determined via the Fibonacci sequence, $T_i = T_{i-1} + T_{i-2},\ \forall\, i \geq 3$, the range of $T_i = 75, 125, \ldots, 850$. Replication for each T is given by the relation $R = \frac{10{,}000}{T}$ for the integer part of the relation.

; As a measure of convergence, the average squared differences between GEP and GLS, GLS and EML, and GEP and EML were captured by:

$$C_R = \frac{1}{R}\sum_{r=1}^{R}\left\{\sum_{j=1}^{h_T}\left(\hat{\alpha}_{js}^{r} - \hat{\alpha}_{j\lambda}^{r}\right)^2\right\}, R = p + q \qquad [4.1]$$

And is compared with the criterion measure $C_T = T^{-1}$, where, for convenience, we have used s and λ to denote the pair of estimates used in the evaluation of $C_R$while $\hat{\alpha}_{(j)}^{r}, j = 1, \ldots, h_T$, as before, denotes the parameter estimates at the rth replication.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

Data-Generating Process

The data-generating process used in the simulation study is presented in **Error! Reference source not found.**.

Table 1: Data Structure

Process	Structure	Parameter Value
P1	ARMA (1, 0)	$(0.7; 0.0)$
P2	ARMA (0, 1)	$(0.0; 0.4)$
P3	ARMA (1,1)	$(0.7;\ 0.4)$
P4	ARMA (2,1)	$(0.5, 0.2; 0.4)$
P5	ARMA (1,2)	$(0.7;\ 0.3, 0.1)$
P6	ARMA (2,2)	$(0.5, 0.2; 0.3,0.1)$

Table 2: Estimated Parameter Values and Sample Sizes

	Process	Parameter	True Value	Estimator	T=75(Mean)	T=125(Mean)	T=200(Mean)	T=325(Mean)	T=525(Mean)	T=850(Mean)
1	$P_1(1,0)$	ϕ_1	0.7	GEP	0.66117944	0.676854126	0.67973859	0.6884713	0.6955236	0.6972699
2				GLS	0.6854949	0.69131673	0.6886923	0.69327329	0.6994132	0.6990736
3				EML	0.66253758	0.67763406	0.6801831	0.68805141	0.6961638	0.69707419
4	$P_2(0,1)$	θ_1	0.4	GEP	0.41646304	0.405868294	0.39195877	0.4005283	0.4022999	0.3966196
5				GLS	0.4299250	0.41097835	0.3955579	0.40263239	0.4035263	0.3973998
6				EML	0.42044837	0.40604285	0.3924096	0.40072115	0.4023788	0.39668798
7	$P_3(1,1)$	ϕ_1	0.7	GEP	0.65555884	0.672510009	0.68838887	0.6976498	0.6904089	0.6947891
8				GLS	0.6778366	0.68728898	0.6977506	0.70272793	0.6937755	0.6970958
9				EML	0.65085997	0.67112715	0.6877177	0.69660846	0.6900087	0.69476456
10		θ_1	0.4	GEP	0.42870028	0.412090866	0.3982597	0.3994185	0.4088747	0.4026696
11				GLS	0.4258462	0.40883597	0.3958234	0.39798633	0.4086827	0.4018921
12				EML	0.43581313	0.41481648	0.3994608	0.4001733	0.4099883	0.40272220
13	$P_4(2,1)$	ϕ_1	0.5	GEP	0.74729542	0.698346975	0.60270501	0.5830821	0.5800875	0.5607545
14				GLS	0.6196919	0.6257646	0.5579713	0.52206532	0.5218179	0.5303176
15				EML	0.60268130	0.59386990	0.5185011	0.52503356	0.5116059	0.53424541
16		ϕ_2	0.2	GEP	-0.0608097	0.005661657	0.09550125	0.1154162	0.1298494	0.1496292
17				GLS	0.0682709	0.07761254	0.1428440	0.16994660	0.1800150	0.1773297
18				EML	0.05731817	0.08893925	0.1643769	0.16104726	0.1843502	0.17161058
19		θ_1	0.4	GEP	0.15279453	0.182474081	0.29494190	0.3126248	0.3114914	0.3414853
20				GLS	0.3021933	0.27480916	0.3454526	0.37795349	0.3718250	0.3719911
21				EML	0.29957737	0.29616436	0.3788311	0.37155445	0.3793551	0.36674232
22	$P_5(1,2)$	ϕ_1	0.7	GEP	0.64209803	0.663612821	0.66949443	0.6887492	0.6822076	0.6949135
23				GLS	0.6925140	0.68811499	0.6832942	0.69550947	0.6863627	0.6978062
24				EML	0.63586773	0.65627115	0.6677945	0.68712898	0.6811649	0.69461578
25		θ_1	0.3	GEP	0.33374164	0.324440565	0.32626738	0.3138208	0.3201734	0.3007742

INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING, MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)
ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

26				GLS	0.1292696	0.12299916	0.1279157	0.11972310	0.1207768	0.1067530
27				EML	0.33863283	0.33260504	0.3282570	0.31479910	0.3212135	0.30141462
28		θ_2	0.1	GEP	0.13942311	0.110499635	0.11607845	0.1097090	0.1116613	0.1006137
29				GLS	0.4437183	0.48081055	0.4582843	0.43088931	0.4347899	0.4006575
30				EML	0.14393842	0.11573372	0.1176162	0.11068653	0.1124858	0.10109678
31	$P_6(2,2)$	ϕ_1	0.5	GEP	0.49919022	0.553030094	0.60742611	0.5883734	0.6044689	0.5754640
32				GLS	0.4564667	0.58491049	0.4726875	0.49454986	0.4823419	0.5400178
33				EML	0.38451568	0.45953029	0.3591297	0.40020291	0.4468448	0.46238272
34		ϕ_2	0.2	GEP	0.08642173	0.114701816	0.09719208	0.1188177	0.1127602	0.1339194
35				GLS	0.1719496	0.11220136	0.2124507	0.19773765	0.2161714	0.1636760
36				EML	0.18640371	0.19629044	0.2921534	0.26858634	0.2408711	0.22552130
38		θ_1	0.3	GEP	0.27921655	0.235098686	0.18571668	0.2101895	0.1876904	0.2198918
39				GLS	0.3271343	0.20466686	0.3244308	0.30773466	0.3121061	0.2561541
40				EML	0.38603628	0.32275311	0.4341254	0.39867916	0.3453440	0.33223466
41		θ_2	0.1	GEP	0.18842684	0.139641002	0.11858267	0.1000197	0.1022251	0.1034287
42				GLS	0.1640100	0.14500102	0.1174242	0.09987044	0.1005205	0.1029733
43				EML	0.17013016	0.14333271	0.1084024	0.09977264	0.1000577	0.09989274

Table 3: Bias and MSE for Each Estimation Procedure

T	Process	Bias_GEP	Bias_GLS	Bias_EML	MSE_GEP	MSE_GLS	MSE_EML
75	P1	-0.0488	-0.0232	-0.0461	0.00239	0.000540	0.00212
125	P1	-0.0118	0.00170	-0.0120	0.000139	0.00000291	0.000144
200	P1	-0.0124	-0.00368	-0.0122	0.000154	0.0000135	0.000149
325	P1	-0.0155	-0.00942	-0.0146	0.000242	0.0000888	0.000215
525	P1	-0.00667	-0.00340	-0.00664	0.0000444	0.0000116	0.0000441
850	P1	0.000246	0.00217	0.000169	0.0000000607	0.00000472	0.0000000286
75	P2	0.00352	0.0114	0.00282	0.0000124	0.000131	0.00000795
125	P2	-0.00837	-0.00299	-0.00805	0.0000700	0.00000897	0.0000647
200	P2	0.00143	0.00505	0.00204	0.00000204	0.0000255	0.00000415
325	P2	0.00272	0.00485	0.00295	0.00000738	0.0000235	0.00000870
525	P2	-0.00531	-0.00411	-0.00529	0.0000282	0.0000169	0.0000280
850	P2	0.00230	0.00312	0.00240	0.00000528	0.00000972	0.00000576
75	P3	-0.0539	-0.0296	-0.0568	0.00290	0.000877	0.00323
125	P3	-0.0400	-0.0257	-0.0417	0.00160	0.000659	0.00174
200	P3	-0.0220	-0.0135	-0.0235	0.000485	0.000182	0.000553
325	P3	-0.00526	0.000549	-0.00560	0.0000277	0.000000302	0.0000313
525	P3	0.00356	0.00681	0.00304	0.0000127	0.0000464	0.00000921
850	P3	-0.00788	-0.00604	-0.00837	0.0000621	0.0000365	0.0000700
75	P4	0.153	0.108	0.115	0.0235	0.0118	0.0131
125	P4	0.166	0.0346	0.0148	0.0274	0.00119	0.000219
200	P4	0.153	0.0971	0.0853	0.0235	0.00943	0.00728
325	P4	0.144	0.0967	0.0941	0.0209	0.00935	0.00886
525	P4	0.0736	0.00888	0.0113	0.00542	0.0000788	0.000128
850	P4	0.0830	0.0447	0.0516	0.00689	0.00200	0.00266
75	P5	-0.0695	-0.0206	-0.0855	0.00484	0.000424	0.00731
125	P5	-0.0420	-0.0211	-0.0468	0.00176	0.000444	0.00219
200	P5	-0.0312	-0.0194	-0.0340	0.000974	0.000377	0.00115

INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING,
MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)
ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

325	P5	-0.00534	0.00147	-0.00699	0.0000285	0.00000215	0.0000489
525	P5	-0.0208	-0.0170	-0.0221	0.000431	0.000287	0.000488
850	P5	-0.00682	-0.00438	-0.00756	0.0000465	0.0000192	0.0000571
75	P6	0.176	-0.0392	-0.0446	0.0310	0.00154	0.00199
125	P6	0.158	0.0296	-0.0632	0.0249	0.000876	0.00400
200	P6	0.195	0.0850	-0.0142	0.0379	0.00723	0.000201
325	P6	0.0439	-0.0245	-0.0563	0.00192	0.000601	0.00317
525	P6	0.0951	0.0348	-0.0320	0.00904	0.00121	0.00102
850	P6	0.0305	-0.0568	-0.100	0.000931	0.00323	0.0101

Table 4: Convergence Thresholds (D vs C)

Process	T_val	Param	D	C	Converged
P1	75	ar1	1.493106e-04	0.0094220126	YES
P1	125	ar1	5.289990e-05	0.0033124115	YES
P1	200	ar1	3.522878e-05	0.0025225950	YES
P1	325	ar1	1.179219e-05	0.0016164380	YES
P1	525	ar1	3.422185e-06	0.0010587146	YES
P1	850	ar1	1.042063e-06	0.0006540241	YES
P2	75	ma1	1.052965e-04	0.0137012959	YES
P2	125	ma1	6.452275e-05	0.0068361929	YES
P2	200	ma1	8.109736e -06	0.0052452223	YES
P2	325	ma1	4.732902e-06	0.0020500065	YES
P2	525	ma1	1.333455e-06	0.0019661002	YES
P2	850	ma1	1.006547e-06	0.0008515338	YES
P3	75	ar1	2.755184e-04	0.0160599796	YES
P3	75	ma1	2.889659e-04	0.0146185451	YES
P3	125	ar1	9.171017e-05	0.0071936179	YES
P3	125	ma1	2.503606e-04	0.0106358788	YES
P3	200	ar1	2.879243e-05	0.0038605966	YES
P3	200	ma1	1.101488e-04	0.0053895968	YES
P3	325	ar1	1.034999e-05	0.0024219560	YES
P3	325	ma1	1.490384e-05	0.0037135382	YES
P3	525	ar1	6.874336e-06	0.0011912826	YES
P3	525	ma1	1.321776e-05	0.0023214317	YES
P3	850	ar1	1.646092e-06	0.0007264899	YES
P3	850	ma1	4.537375e-06	0.0010187526	YES
P4	75	ar1	3.420026e-01	0.3072018256	NO
P4	75	ar2	1.996112e-01	0.1957135467	NO
P4	75	ma1	4.152028e-01	0.3028746536	NO
P4	125	ar1	2.669891e-01	0.2164452655	NO
P4	125	ar2	1.726361e-01	0.1446557172	NO
P4	125	ma1	2.851509e-01	0.2042712564	NO
P4	200	ar1	2.890677e-01	0.2371224787	NO
P4	200	ar2	1.883616e-01	0.1647818128	NO
P4	200	ma1	2.954760e-01	0.2473319448	NO
P4	325	ar1	4.369652e-02	0.1607622777	YES
P4	325	ar2	2.694771e-02	0.1087164706	YES
P4	325	ma1	4.610328e-02	0.1496021705	YES

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

P4	525	ar1	4.239244e-02	0.1283222189	YES
P4	525	ar2	2.787307e-02	0.0843635641	YES
P4	525	ma1	4.102681e-02	0.1213877662	YES
P4	850	ar1	2.741151e-02	0.0738167978	YES
P4	850	ar2	1.784693e-02	0.0482923042	YES
P4	850	ma1	2.822225e-02	0.0729458484	YES
P5	75	ar1	6.425307e-03	0.0298653266	YES
P5	75	ma1	9.198733e-03	0.0379876074	YES
P5	75	ma2	7.442244e-03	0.0287535307	YES
P5	125	ar1	8.004057e-04	0.0199091435	YES
P5	125	ma1	1.063077e-03	0.0245159348	YES
P5	125	ma2	8.476869e-04	0.0241792397	YES
P5	200	ar1	1.566959e-04	0.0104618020	YES
P5	200	ma1	1.727998e-04	0.0127901935	YES
P5	200	ma2	1.315822e-04	0.0100145274	YES
P5	325	ar1	4.386340e-05	0.0048297281	YES
P5	325	ma1	4.971498e-05	0.0054908511	YES
P5	325	ma2	4.027359e-05	0.0048265603	YES
P5	525	ar1	1.031019e-05	0.0028525433	YES
P5	525	ma1	2.028162e-05	0.0042154993	YES
P5	525	ma2	1.549325e-05	0.0038667733	YES
P5	850	ar1	4.793569e-06	0.0017050822	YES
P5	850	ma1	7.854964e-06	0.0021204804	YES
P5	850	ma2	7.187170e-06	0.0023276628	YES
P6	75	ar1	4.470805e-01	0.3292313312	NO
P6	75	ar2	2.654512e-01	0.2334360308	NO
P6	75	ma1	5.138451e-01	0.3924679603	NO
P6	75	ma2	2.264582e-02	0.0442964111	YES
P6	125	ar1	2.239885e-01	0.2488812293	YES
P6	125	ar2	1.436606e-01	0.1743962483	YES
P6	125	ma1	2.460867e-01	0.2861316257	YES
P6	125	ma2	1.041010e-02	0.0222345687	YES
P6	200	ar1	3.270347e-01	0.2290212239	NO
P6	200	ar2	2.175085e-01	0.1550506697	NO
P6	200	ma1	3.389386e-01	0.2484301869	NO
P6	200	ma2	2.724779e-03	0.0128367576	YES
P6	325	ar1	1.811885e-01	0.1784909371	NO
P6	325	ar2	1.160543e-01	0.1189935996	YES
P6	325	ma1	1.849996e-01	0.1811415612	NO
P6	325	ma2	9.926453e-04	0.0058761309	YES
P6	525	ar1	1.839834e-01	0.1838276324	NO
P6	525	ar2	1.203909e-01	0.1214142733	YES
P6	525	ma1	1.885004e-01	0.1914593077	YES

INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING,
MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)
ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

Table 5: Convergence Thresholds (D vs C)

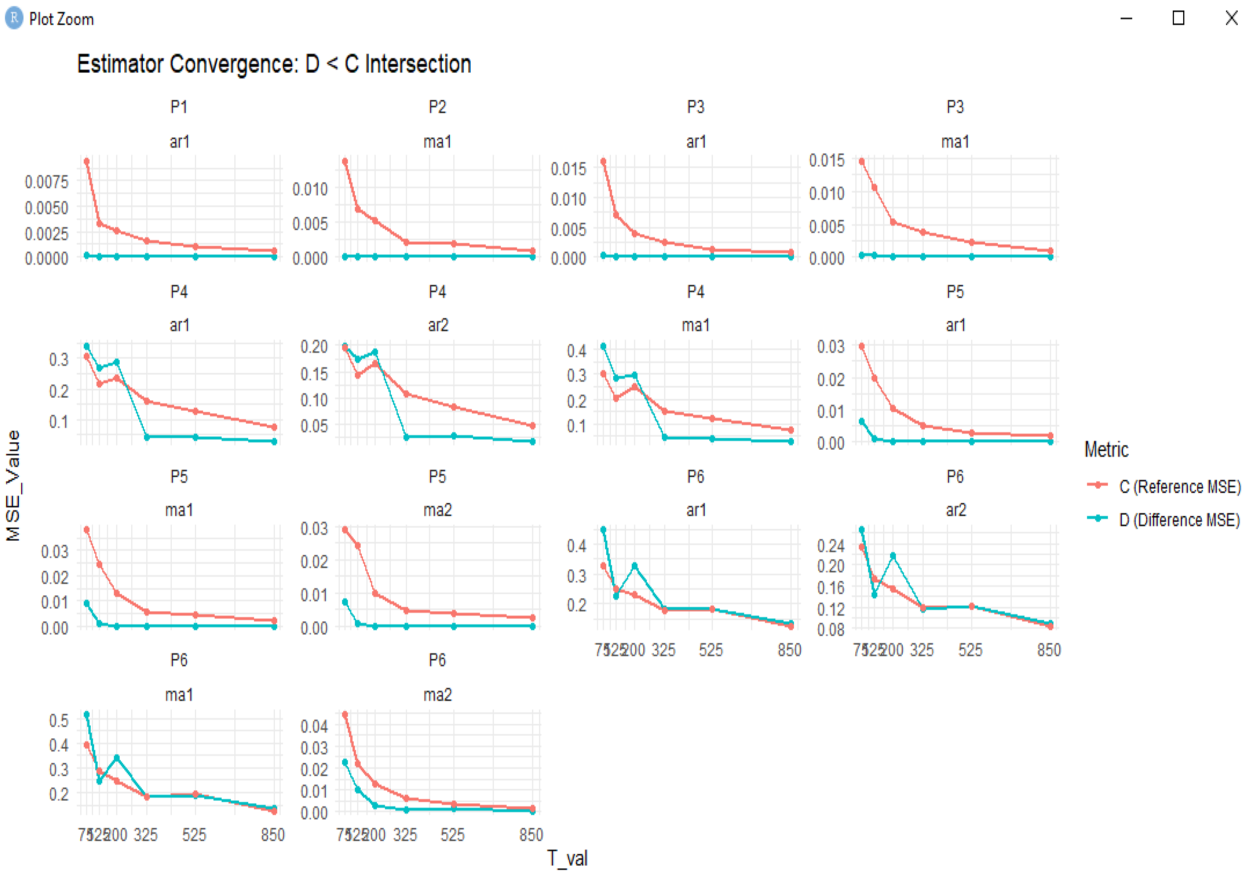

Figure 1: Estimator Convergence: D < C intersection

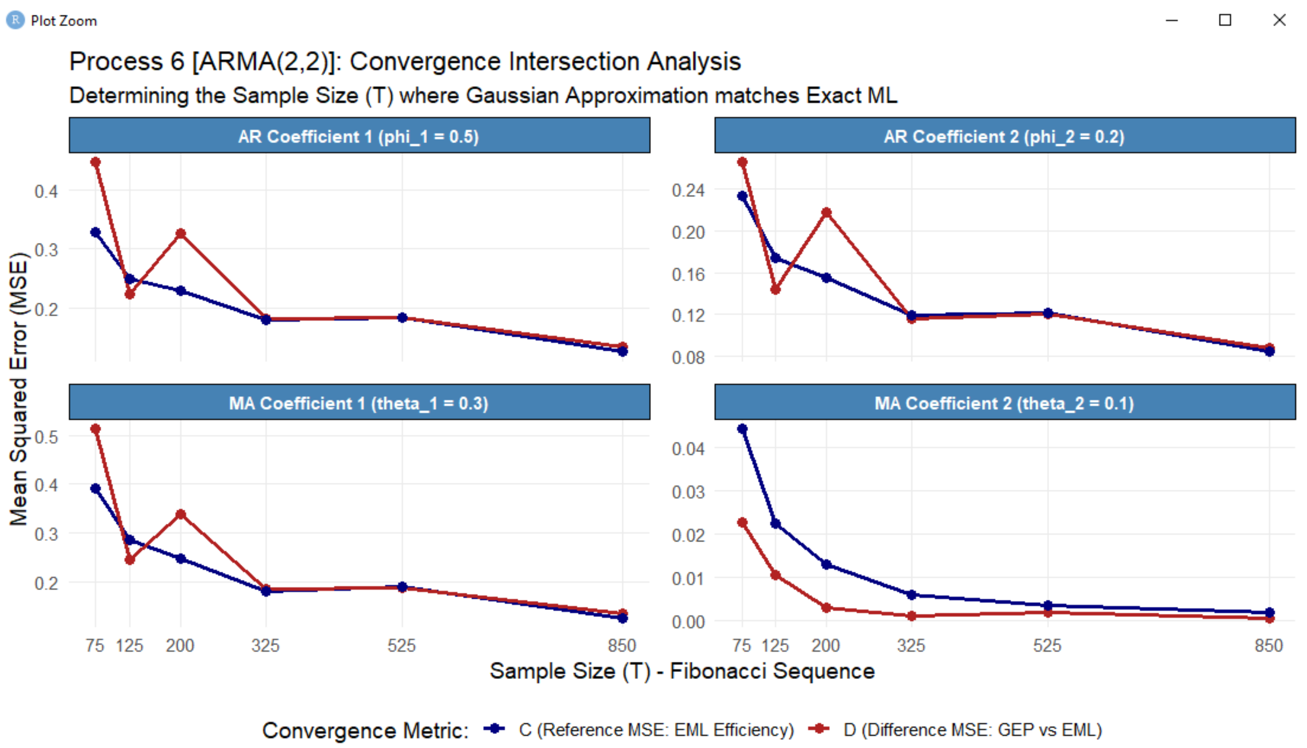

Figure 2: ARMA (2,2) Convergence Intersection Analysis

Design of Experiment

For each estimator, the sample size, parameter estimates, standard error, bias, and mean square error were obtained by averaging the estimates over the number of replications, where $\delta(\hat{\phi}_h) = \left\{\frac{1}{R}\sum_{r=1}^{R}(\hat{\phi}_h^r - \phi_0)^2\right\}^{\frac{1}{2}}$ with $\hat{\phi}_h^r$ denoting the estimate of ϕ_h in the r[th] replication. Following the standard practise, the accuracy of the parameter estimate is judged by the mean square error. The outcomes of this investigation are summarized inError! Reference source not found. Table 2 and Error! Reference source not found., respectively.

As a measure of convergence, the average square differences between GEP and GLS, GEP and EML, GLS and EML were computed by:

$$C_R = \frac{1}{R}\sum_{r=1}^{R}\left\{\sum_{j=1}^{p+q}(\hat{\phi}_{js}^r - \hat{\phi}_{j\lambda}^r)^2\right\} \quad [4.2]$$

And is compared with the criterion measure $C_T = T^{-1}$ where, for instance, we have used S and λ to denote the pair of estimates used in the evaluation of C_R while $\hat{\phi}_{(j)}^r, j = 1, \ldots, h_T$ as before, denotes the parameter estimates at the r[th] replication.

CONCLUSION

The ARMA (2, 2) process (P_6) constitutes the upper bound of structural complexity in the simulation study, comprising four parameters $(\phi_1, \phi_2, \theta_1, \theta_2)$, and provided a stringent test of asymptotic approximation. While the autoregressive components exhibited stable and accurate estimation even at small sample sizes, the moving average parameters displayed pronounced finite sample instability, characterized by elevated variance and delayed convergence. The convergence diagnostics showed that the Difference MSE (D) exceeded the Reference MSE (C) through T =200, with consistent interaction achieved for all parameters, only at $T \geq 525$. This identified a clear "indifference threshold," beyond which the Gaussian Estimation Procedure (GEP) became numerically indistinguishable from the Exact Likelihood Estimator (EML). More broadly, the results revealed a systematic complexity lag: as model dimensionality increased, the $D < C$ intersection shifted markedly to the right, indicating that estimator choice depends jointly on sample size and the parameter-to-sample ratio rather than on T alone.

REFERENCES

1. Ansley, C. F., & Newbold, P. (1980). Finite sample properties of estimators for autoregressive moving average models. Journal of Econometrics, 13(2), 159–183. https://doi.org/10.1016/0304-4076(80)90013-5
2. Athanasopoulos, G., Hyndman, R. J., Kourentzes, N., & Panagiotelis, A. (2024). Forecast reconciliation: A review. International Journal of Forecasting, 40(2), 430–456. https://doi.org/10.1016/j.ijforecast.2023.10.010
3. Burnside, C. (1994). Hansen-Jagannathan Asset-Pricing. 12(1), 57–79.
4. Chan, L., & Yao, Qiwei (2024). Small Sample Consistency. Journal of Business & Economic Statistics (Volume 42). https://doi10.1080/07350015.2023.2245611
5. Francq C., & Zakoian J. M. (2024). Asymptotic Theory for Time Series Models. CRC Press
6. Geore E. P. Box, Gwilym M. Jenkins, G. C. R. (2015). Time Series Analysis-Forecasting and Control (Fifth Edit). Prentice Hall.
7. Hannan, E. J., & Deistler, M. (1990). The Statistical Theory of Linear Systems. In U. of W. Robert E. O'Malley, Jr. (Ed.), Technometrics (Vol. 32, Issue 1). Society for Industrial and Applied Mathematics, Philadelphia (SIAM). https://doi.org/10.2307/1269869
8. Greene, W. H. (2024). Econometric Analysis (9th ed, Issue 457). Printice Hall. http://pubs.amstat.org/doi/abs/10.1198/jasa.2002.s458
9. Hamilton, J. D. (1994). Time Series Analysis. In The SAGE Encyclopedia of Communication Research Methods (1st Editio, Issue I). Princeton University Press. https://us.sagepub.com/sites/default/files/upm-assets/23658_book_item_23658.pdf

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

10. Harden, J. W., & Hilbe J. M., (2021). "Generalised Estimating Equation (3rd ed." Chapman and Hall/CRC
11. Koreisha, S., & Pukkila, T. (1990). A Generalized Least-Squares Approach for Estimation of Autoregressive Moving-Average Models. Journal of Time Series Analysis, 11(2), 139–151. https://doi.org/10.1111/j.1467-9892.1990.tb00047.x
12. Koutsoyiannis, D. (2009). The Hurst phenomenon and fractional Gaussian noise made easy. 6667. https://doi.org/10.1080/02626660209492961
13. Lin, Y., Li, W., Zhu, Q., & Li, G. (2024). On scalable ARMA models. http://arxiv.org/abs/2402.12825
14. Olajide, J. T., Ayansola, O. A., Odusina, M. T., & Oyenuga, I. F. (2012). Forecasting the Inflation Rate in Nigeria: Box-Jenkins Approach. 3(5), 15–19.
15. Robert H. Shumway, & Stoffer D. S. (2025). Time Series Analysis and Its Applications - With R Examples (G. C. S. F. I. Olkin (ed.); Fifth edit). Springer. https://doi.org/10.1007/978-1-4419-7865-3
16. Ruey S. Tsay, & R. C. (2019). Nonlinear Time Series Analysis. In R. S. T. David J. Balding, Noel A. C. Cressie, Garrett M. Fitzmaurice, GeofH. Givens, Harvey Goldstein, Geert Molenberghs, David W. Scott, Adrian F. M. Smith (Ed.), Sustainability (Switzerland) (First Edit, Vol. 11, Issue 1). John Wiley & Sons, Inc.
17. Salau, M. O. (2000). On the Accuracy and Asymptotic Convergence of Widely Used Estimators of Autoregressive Approximation of Mixed ARMA Models. 24th Annual Conference of the Nigerian Statistical Association, Lagos. November 2000, 18.
18. Sanchez, J. (2010). Introduction to modern time series analysis. In Journal of Applied Statistics (Vol. 37, Issue 6). https://doi.org/10.1080/02664760902899766
19. Walker, A. M. (1964). Asymptotic Properties of Least-Squares Estimates of Parameters of the Spectrum of a Stationary Non-Deterministic Time-Series. Journal of the Australian Mathematical Society, 4(3), 363–384. https://doi.org/10.1017/S1446788700024137
20. Wei, W. W. S. 63. (2006). Time Series Analysis Univariate and Multivariate Methods (2nd Editio). Pearson Addison Wesle
21. Whittle, P. (1953). Estimation and information in stationary time series. Arkiv För Matematik, 2(5), 423–434. https://doi.org/10.1007/BF02590998
22. Zhang, R., & Ling, S., (2024). "On the Convergence Rates of Maximum Likelihood Estimates in Higher-Order ARMA Processes." Journal of Time Series Analysis.
23. Zinde-walsh, V. (1994). A simple noniterative estimator for moving average models. 81(1), 143–155.