

International Journal of Statistics and Applied Mathematics



ISSN: 2456-1452
NAAS Rating (2026): 4.49
Maths 2026; 11(8): 145-159
© 2026 Stats & Maths
<https://www.mathsjournal.com>
Received: 12-06-2026
Accepted: 27-07-2026

Iman Jamal Abdul Al Kadhimi
Department of Statistics, College
of Administration and
Economics, University of Sumer,
Dhi-Qar, Iraq

Saif Sabbar Hatem
Al-Rifa'i Department of
Education, Dhi Qar Directorate
of Education, Ministry of
Education, Iraq

Abbas Salem Kazim
Department of Statistics, College
of Administration and
Economics, University of Sumer,
Dhi-Qar, Iraq

Corresponding Author:
Iman Jamal Abdul Al Kadhimi
Department of Statistics, College
of Administration and
Economics, University of Sumer,
Dhi-Qar, Iraq

Evaluating the statistical efficiency of robust and classical estimators for ARIMA models under contaminated time series

Iman Jamal Abdul Al kadhimi, Saif Sabbar Hatem and Abbas Salem Kazim

DOI: <https://www.doi.org/10.22271/math.2026.v11.i8b.2675>

Abstract

This study models natural gas production volumes in Iraq for the period (2000-2025) using an annual time series of 26 observations, addressing outliers and external shocks that undermine the accuracy of classical estimation. To achieve this, data normality and stationarity were evaluated using the Kolmogorov-Smirnov and Shapiro-Wilk tests, the autocorrelation and partial autocorrelation functions (ACF & PACF), and box plots. MATLAB (v2021) was utilized for statistical programming to construct and estimate ARIMA models, evaluating the trade-off between the classical Maximum Likelihood (ML) method and the Robust Iteratively Reweighted Least Squares (RLS) method based on MAE, AIC, and BIC criteria. The results revealed non-normality in the data distribution, requiring first-differencing ($d = 1$) to achieve stationarity. Model evaluation criteria demonstrated the superiority of the robust estimation method (RLS) in isolating the impact of outliers compared to the conventional approach, highlighting the ARIMA(0,1,1) model as optimal by recording the lowest error and information criteria values (MAE = 4.490651, AIC = 90.38945, BIC = 92.8272). The estimated moving average parameter was $\theta_1 = -0.18344$ with the absence of autoregressive parameters, economically confirming that natural gas production is influenced by sudden shocks, geopolitical conditions, and accumulated technical shutdowns rather than historical production values. The study recommends that planning authorities adopt this model as a reference tool for forecasting and decision-making.

Keywords: ARIMA models, Robust Least Squares (RLS), Robust estimation, Outliers, Natural gas production

1. Introduction

Autoregressive Integrated Moving Average (ARIMA) models are among the most prevalent and efficient standard tools for time series analysis and forecasting, widely applied across various economic, financial, and social sciences. The development of this integrated methodology is credited to Box and Jenkins (1976) [14], who established an inferential framework based on combining two primary components: the Autoregressive (AR) component, which relies on the historical values of the series, and the Moving Average (MA) component, which tracks past forecast errors, while addressing non-stationary behavior through the mechanism of differencing (d). Despite the high statistical efficiency of ARIMA models under ideal standard conditions, they face severe challenges when time series are contaminated by outliers. Outliers represent a natural and inherent phenomenon in real-world data, arising from sudden structural shifts, external shocks, or measurement errors. Extreme observations cause fundamental distortions in parameter estimates and inflate error variance when relying on conventional estimators such as the Maximum Likelihood (ML) method, due to its high sensitivity and breakdown point in the presence of minor statistical contamination, ultimately leading to misleading inferences and forecasts [24]. Consequently, recent research has shifted toward developing and adopting robust estimation methods as an effective alternative. These estimators possess the capacity to downweight contaminated observations and yield stable, highly durable parameters compared to classical approaches [21].

1.1 Research Problem

Numerous applied studies face fundamental challenges regarding the accuracy of statistical modeling and forecasting for future time series trends, negatively impacting the reliability of decisions derived from these estimates. These challenges are particularly exacerbated when dealing with time series affected by statistical contamination—specifically the presence of outliers—which causes violations of underlying classical assumptions, such as the normality of residual distributions. Consequently, the efficiency of classical estimation methods degrades significantly due to their extreme sensitivity to contaminated data, leading to biased parameter estimates and inflated standard errors. Hence, the core research problem centers on how to establish robust predictive methodologies within such contexts. This necessitates the implementation of robust estimation techniques capable of downweighting the influence of contaminated observations, thereby refining forecasting models and enhancing their performance under complex data conditions.

1.2 Research Objectives

This research aims to achieve several primary objectives, formulated as follows:

- **Testing and diagnosing** the impact of statistical contamination and non-normality resulting from outliers on the behavior and stability of ARIMA model parameters.
- **Comparing the efficiency** of ARIMA model estimators using the classical Maximum Likelihood (ML) estimation method against the performance of the Robust Iteratively Reweighted Least Squares (RLS) estimation method to determine the optimal model and estimator.
- **Evaluating the goodness-of-fit and forecasting accuracy** of the estimated models across different estimators using comparative criteria, specifically the Mean Absolute Error (MAE), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC).
- **Applying the estimated classical and robust models** to real-world time series contaminated with outliers to validate their practical performance and empirical efficiency.

1.4 Literature Review

- Mohammed and Abd (2007) ^[23] investigated robust estimation methods for the mixed autoregressive moving average model ARMA(1,1) and compared them with conventional approaches through simulation experiments. Their findings revealed that the classical Least Squares (LS) method remains the optimal and most efficient choice in the absence of extreme observations. However, in the presence of outliers within the series, the robust estimation method (RAMB) demonstrated superior stability and parameter durability, highlighting the necessity of adopting robust estimators to preserve modeling accuracy.
- Muler, Peña, and Yohai (2009) ^[24] introduced a mathematical framework developing a novel class of robust estimators for ARMA models. The proposed approach was based on a specialized formulation of M-estimators combined with an innovative residual calculation mechanism designed to constrain the influence of severe outliers, ensuring that their statistical impact remains confined solely to the period of occurrence rather than propagating throughout the series. Monte Carlo simulations demonstrated that these newly developed estimators significantly outperformed standard M-estimators and traditional diagnostic procedures, owing to their consistent asymptotic properties.
- Flimmel *et al.* (2017) ^[17] addressed the econometric issues resulting from data contamination by outliers, which induce severe biases leading to inaccurate order selection and misleading ARMA models. Utilizing both simulation studies and empirical applications, the authors emphasized the critical role of robust treatments. Their results established that employing robust techniques serves as a resilient foundation for preventing parameter bias and ensuring accurate model identification and order determination in outlier-contaminated data.
- Arai (2020) ^[8] examined the implementation of Iteratively Reweighted Least Squares (RLS) algorithms for forecasting complex time series subject to statistical gaps (missing data). The author introduced two learning and updating algorithms incorporating the Kalman Filter with RLS via time and measurement update mechanisms. The algorithmic outputs indicated that robust iterative estimators with adaptive weighting structures offer rapid convergence and smooth tracking in forecasting and fitting, underscoring the role of iterative (RLS) estimators in stabilizing dynamic forecasting models.
- Zhao *et al.* (2022) ^[26] conducted a cross-sectional comparative study evaluating the predictive efficiency of classical time series models, specifically ARIMA and the Grey model GM(1,1), against Long Short-Term Memory (LSTM) deep neural networks. Predictive performance was assessed using MAE, RMSE, and MAPE criteria. The empirical findings demonstrated that advanced non-linear frameworks like LSTM exhibit superior forecasting ability in modeling complex dynamic trends compared to classical structures.
- Bianco *et al.* (2024) ^[13] directly addressed outlier management in ARIMA time series models, specifically distinguishing between Innovation Outliers (IO) and Additive Outliers (AO). The authors developed robust estimation procedures as an alternative to classical Maximum Likelihood (ML) estimation, which suffers efficiency losses under data contamination. Monte Carlo simulation results confirmed that utilizing residuals derived from robust estimation enhances outlier detection capabilities and prevents the inflation of error variance compared to sensitive conventional methods.
- Ardia *et al.* (2025) ^[9] focused on assessing the forecasting performance of standard models when time series experience severe shocks causing deviations from residual normality (e.g., market shocks and periods of high uncertainty). The authors implemented advanced robust estimators, such as the Huber M-estimator, comparing them with classical approaches using Mean Absolute Error (MAE) and information criteria. Their empirical results demonstrated that integrating robust methods mitigates parameter drift and achieves superior forecasting accuracy and statistical stability in contaminated series.
- Otopeleanu *et al.* (2025) ^[25] introduced an advanced algorithmic solution based on a robust regularized RLS algorithm using third-order tensor decomposition. The study aimed to reduce computational complexity while maintaining parameter estimation efficiency in high-dimensional dynamic systems. Experimental results verified that the proposed robust algorithm exhibits high durability and clear superiority in downweighting distortions caused by signal contamination relative to standard RLS algorithms.

- Mano *et al.* (2026) ^[20] evaluated the impact of stochastic outlier characteristics on parameter accuracy in ARIMA models. They conducted a comparative analysis between traditional models reliant on ideal normality assumptions and robust models incorporating error-correction factors and weighting adjustments for extreme observations. The study concluded that robust estimators guarantee parameter unbiasedness and prevent misleading inferences resulting from abrupt shocks in the data.

2. Theoretical Framework

2.1 Time Series

A time series is defined as a sequence of observations ordered chronologically over time. These observations are typically dependent rather than statistically independent, where each observation relies on preceding ones. This temporal dependency is effectively exploited to generate reliable forecasts that reflect the sequential ordering of the data. Thus, y_t represents the observation at time t , y_{t-1} denotes the preceding observation, and y_{t+1} represents the subsequent observation ^[1]. Time series analysis is a crucial statistical methodology that has undergone significant developments, establishing itself as an effective tool for forecasting. This approach relies on analyzing the behavior of a given phenomenon over a specified time horizon and leveraging the extracted historical patterns to project future trends ^[2].

2.1.1 Stationary Time Series

A time series representation over a given period $[t, t+h]$ may sometimes exhibit structural consistency with its representation over a different time period $[s, s+h]$, indicating homogeneity in the behavior of the series. This property is referred to as **stationarity**. In general, a time series is visually identified as stationary through its time plot, or formally defined as stationary if it possesses a constant mean and variance over time, remaining unaffected by systematic trends or seasonal shifts. Specifically, a time series is said to be weakly stationary if it satisfies the following conditions ^[4]:

Mean Stationarity

$$E(Y_t) = \mu \quad (1)$$

$$\bar{Y} = \hat{\mu} = \frac{1}{n} \sum_{t=1}^n Y_t \quad (2)$$

Where:

- $\bar{Y}$: The population mean of the time series.
- y_t : The observation of the series at time t .
- n : The total number of observations.

2.1.2 Variance Stationarity

$$Var(Y_t) = E(Y_t - \mu)^2 = \sigma_Y^2 = \gamma_0 \quad (3)$$

where γ_0 represents the variance of the time series, which remains constant over time and independent of t .

$$\hat{\gamma}_0 = \frac{1}{n} \sum_{t=1}^n (Y_t - \bar{Y})^2 \quad (4)$$

- Absence of Trend:** There is no systematic increase or decrease in the values.
- Equal relationship between values over time (Autocorrelation):** Meaning that the correlation between values depends only on the time interval between them.

2.2 Autocorrelation Function (ACF)

It is a measure that determines the degree of correlation between observations of the same variable at different time periods. It also serves as a tool to identify the characteristics of time series. The concept of autocorrelation is defined as the correlation of a random variable with its preceding or subsequent variable over a specified time period, i.e., the correlation of the series with itself or shifted by a certain period. This correlation is given by the following formula ^[3]:

$$\rho_Y(k) = \frac{E(Y_t - \mu_Y)(Y_{t+k} - \mu_Y)}{E(Y_t - \mu_Y)^2} = \gamma_Y(k) / \sigma_Y^2 \quad (5)$$

Where:

- ρ_k : Represents the autocorrelation of y_t values at lag k .
- γ_k : Represents the autocovariance at lag k .
- σ_Y^2 : Represents the variance of y_t values.

2.3 Partial Autocorrelation Function (PACF)

It calculates the correlation between y_t and y_{t-k} while holding y constant across the other periods, i.e., calculating the degree of correlation between y_t and y_{t-k} controlling for the effect of other values. From the autoregressive model equation AR of order p , it is obtained as follows ^[5]:

$$\rho_1 = \phi_{11}$$

.

$$\rho_k = \gamma_k / \gamma_0$$

$$\hat{\phi}_{11} = \hat{\rho}_1 = \gamma_1 \quad (6)$$

That is, the first partial autocorrelation is equal to the first autocorrelation, and by varying the number of terms, it will yield partial autocorrelations for different values of k . Its general formula is:

$$\begin{aligned} \hat{\phi}_{kk} &= \frac{\hat{\rho}_k - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} \hat{\rho}_{k-j}}{1 - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} \hat{\rho}_j} \\ &= \frac{\hat{\rho}_k - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} \hat{\rho}_{k-1,j} \hat{\rho}_{k-j}}{1 - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} \hat{\rho}_j} \quad (7) \end{aligned}$$

Where: $\hat{\phi}_{kk}$ is the estimated value of ϕ_{kk} .

$$\hat{\phi}_{kj} = \hat{\phi}_{k-1,j} - \hat{\phi}_{kk} \hat{\phi}_{k-1,k-j} \quad k = 3, 4, \dots, j = 1, 2, \dots, k-1$$

Where: ϕ_{kj} is the partial correlation at period kj .

2.4 Box-Jenkins Models

The Box-Jenkins methodology is a systematic approach for time series analysis and forecasting, primarily focusing on identifying and estimating forecasting models for temporal data, particularly those that can be expressed as Autoregressive Integrated Moving Average (ARIMA) models ^[3]. Box-Jenkins models were developed between the years (1970-1976) and are abbreviated as (B-JM). They are used in analyzing temporal data within the time domain and are divided into two categories: the first is Seasonal Models, which will not be detailed as the study relies on the second category, Non-Seasonal Models, which are used to represent two types of series ^[6]:

2.5 Stationary Time Series Models

These time series models are defined as possessing the property of stationarity, meaning they do not contain a general trend in their data and have a constant mean around which they fluctuate. These models include three types:

1. Autoregressive (AR) Models

Autoregression is defined as the correlation between the current observations of a time series and previous observations of the same series. The model order is denoted by (p) , where (p) represents the order of autoregression as an integer. Autoregressive models of lag order $AR(p)$ can be represented as follows ^[7]:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t \quad (8)$$

For example, the first-order autoregressive model $AR(1)$ is:

$$Y_1 = \phi_1 Y_{t-1} + e_t \quad (9)$$

And if y_t represents the deviation of the observations from the mean μ , the equation takes the following form:

$$(Y_t - \mu) = \phi_1 (Y_{t-1} - \mu) + e_t$$

$$Y_t = (1 - \phi_1) \mu + \phi_1 Y_{t-1} + e_t \quad (10)$$

Where the stationarity condition for the autoregressive parameter ϕ_1 is:

The value of ϕ_1 must be less than one in absolute value for the series to be stationary:

$$-1 < \phi < 1$$

Where the autocorrelation coefficient ρ_k for this model is:

$$\rho_k = \phi_1^k \quad (11)$$

And the partial autocorrelation coefficient is equal to the parameter of the autoregressive model $AR(1)$, while the other partial autocorrelation coefficients are equal to zero ^[10]. As for the second-order autoregressive model $AR(2)$, it is:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t \quad (12)$$

And the autocorrelation function for this model is:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \quad (13)$$

And the stationarity condition for this function is:

$$-1 < \phi < 1$$

While the partial autocorrelation parameters (PACF) for the AR(2) model diminish, as current values depend only on values at time $(t-1)$ and $(t-2)$. The model will be clear at (lag 1, lag 2), and the coefficients begin to decay rapidly toward zero, meaning that the effect of subsequent lags (s) is weak or insignificant.

$$\phi_{11} \neq 0$$

$$\phi_{22} \neq 0$$

Where ϕ_{22} is at the second lag, and each symbol can be expressed as follows:

- y_t : Observations of the time series at time t .
- ϕ_i : Model parameters.
- p : Order of the autoregression.
- μ : Mean of the observations.

2. Moving Average (MA) Models

The studied phenomenon can be represented by another model other than the autoregressive model, which is the moving average model. The difference between this model and the previous one is that moving averages depend on error terms at different time points. Moving average models of order (q)—where (q) represents the order of the moving average, denoted as MA(q)—can be represented by the following mathematical model ^[12]:

$$Y_t = -e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad (14)$$

When moving average parameters are added to the autocorrelation function (ACF) equation, the equation becomes as follows:

$$\rho_k = \frac{(-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q)}{(1 + \theta_1^2 + \dots + \theta_q^2)} \quad (15)$$

When

$$K \leq q$$

$$\rho_k = 0, K > q$$

And the parameters of the partial autocorrelation function (PACF) never decay absolute to zero (as they decrease exponentially). For example, the first-order moving average model MA(1) is ^[11]:

$$Y_t = e_t - \theta_1 e_{t-1} \quad (16)$$

The behavior of the "Autocorrelation Function and Partial Autocorrelation Function" in the MA(1) model is similar to that of the AR(1) model. The second-order moving average model MA(2) is:

$$Z_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} \quad (17)$$

Where the Autocorrelation Function (ACF) equation for the MA(2) model can be formulated as follows:

$$\rho_k = \frac{-\theta^2}{(1 + \theta_1^2 + \theta_2^2)} \quad (18)$$

When:

$$k \leq 2$$

$$\rho_k = 0, k > 2$$

While the Partial Autocorrelation Function (PACF) equation is as follows:

$$\phi_{kk} = \frac{-\theta_1^k (1 - \theta_1^2)}{(1 + \theta_1^{2(k+1)})} \quad (19)$$

The symbols above can be expressed as:

- y_t : Observations of the series at time t .
- θ_i : Model parameters.
- q : Order of the moving average.

2.6 Non-Stationary Time Series Models

These models apply to represent series in which a general trend is one of their components, resulting in multiple means around which the data fluctuate. These models can be represented using the first type of models, but after removing non-stationarity from the original series by using the differencing method and taking an appropriate number of differences of the original data for time periods representing the length of the season, adding the letter (d) to the notation ^[4]. When introducing the difference operator to the ARMA(p, q) model, the model transforms into an ARIMA(p, d, q) model. These models include the same stationary time series models; however, to distinguish between this type and the first type of models, the word "Integrated" is added to the model name to indicate the application of stationary time series models to non-stationary time series after converting them into stationary time series to represent this type of model. It is preferable to use the mixed model after adding the letter (I), making the model abbreviation (ARIMA), and the model becomes of order (p, d, q). Converting non-stationary time series into stationary ones avoids many problems, as stationary temporal properties are independent of time. Furthermore, it becomes easier to develop stationarity theory to represent stochastic time series, meaning that the new forecasted series will also be stationary. From a practical standpoint, the calculations required to estimate the parameters of stationary models require less work. The model is defined as follows:

$$Z_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad (20)$$

To make the series stationary in this model, and assuming that y_t denotes the original time series data, and w_t denotes the time series data after taking the first difference, the equation for calculating the first difference is as follows:

$$\begin{aligned} \beta Y_t &= Y_{t-1} \\ \nabla Z_t &= \nabla^{(d)} Y_t \\ &= Y_t - Y_{t-1} \\ &= Y_t - \beta Y_t \\ &= (1 - \beta) Y_t \quad (21) \end{aligned}$$

That is, the value of the stationary time series w_t at time t equals the difference between the value of the original series y_t at time t and its value y_{t-1} at the previous time $t - 1$.

That is, where w_t denotes the time series data after taking the first difference, the equation for calculating second-order differences is as follows: ^[16]

$$\begin{aligned} Z_t &= \nabla^2 Y_t = \nabla(\nabla Y_t) \\ &= (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) \\ &= Y_t - 2Y_{t-1} + Y_{t-2} \\ &= Y_t - 2\beta Y_t + \beta^2 Y_t \\ &= (1 - 2\beta + \beta^2) Y_t \quad (22) \end{aligned}$$

If d is the number of differences required to obtain a stationary time series, then the general equation for the stationary series is as follows:

$$Z_t = \nabla^d Y_t = (1 - \beta)^d Y_t \quad (23)$$

Where:

- w_t : Denotes the stationary series after taking d differences.
- y_t : Denotes the original series.
- $\Delta^d y_t$ (or $(1 - B)^d y_t$): Represents the operation of taking d differences of the original series.
- $(1 - B)^d$: Represents the d -th differencing using the Lag Operator (B).

Accordingly, the series is expressed as y_t , and to render it stationary, d differences are taken from it to obtain the series Z_t , where d represents the minimum number of differences required to obtain a time series that is stationary in the mean.

Stages of Applying Box-Jenkins Models in Forecasting

2.4.1 Identification & Diagnostic Phase

In this phase, the orders (p, d, q) of the ARIMA model are identified, which involves monitoring the "Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF)" as a fundamental step in time series analysis ^[15].

- If the Autocorrelation Function (ACF) values fall within the 95% confidence interval bounds, this indicates that the autocorrelation coefficient ρ_k does not significantly differ from zero, demonstrating the stationarity of the series and that it is integrated of order zero ($I(0)$). In this case, analyses can be performed using the original values of the variable y_t without requiring any transformations.

- If the Autocorrelation Function (ACF) does not fall within the 95% confidence interval bounds for several time lags, then ρ_k differs significantly from zero, indicating that the time series is non-stationary and requires differencing to transform it into a stationary series. When taking the first difference, $d = 1$, and when taking the second difference, $d = 2$. In the ARIMA(p, d, q) model, p and q can be identified by inspecting the pattern of the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF):
1. When autocorrelations decay exponentially toward zero, it indicates the presence of an Autoregressive (AR) model, whose order is determined by the number of partial autocorrelations that differ significantly from zero.
 2. If partial autocorrelations decay exponentially toward zero, the model is a Moving Average (MA) model, whose order is determined by the number of statistically significant autocorrelations.
 3. If both autocorrelations and partial autocorrelations decay exponentially toward zero, this indicates an ARMA model.

The order of both AR and MA is determined by observing the behavior of the autocorrelation and partial autocorrelation functions. If the autocorrelation function does not decay rapidly with increasing lag orders, it indicates that the time series is non-stationary and requires differencing. The diagnostic process includes reviewing these two functions to determine the stationarity of the time series and the parameters (p, d, q), as shown in the table below ^[4]:

Table 1: Characteristics of ACF and PACF for Selected Models

Model	ACF (ρ_k)	PACF (ϕ_{kk})
AR(p)	Gradual decay to zero, possibly with oscillatory behavior	Cutoff identifies AR; indicates partial autocorrelations that are non-zero and statistically significant up to lag p
MA(q)	Cutoff identifies MA; indicates autocorrelations that are non-zero and statistically significant up to lag q	Gradual or oscillatory decay
ARMA(p, q)	Gradual decay (direct or oscillatory) starting after lag q	Gradual decay (direct or oscillatory) starting after lag p
ARIMA(p, d, q)	Geometric decay starting from p, q	Zero after ϕ_{kk}

Where:

- ρ_k : Autocorrelation function coefficient.
- ϕ_{kk} : Partial autocorrelation function coefficient.

2.4.2 Estimation & Testing Phase

After identifying the model and determining its order, the parameters of this model are estimated, representing the parameters ϕ_i in the AR(p) model, θ_j in the MA(q) model, and (ϕ_i, θ_j) in the ARMA(p, d, q) model.

First: Classical Maximum Likelihood Estimation (MLE)

Maximum Likelihood Estimation is considered the most common standard tool for estimating the parameters of time series models (Berenguer-Rico *et al.*, 2021) ^[5]. This method derives its strength from the assumption that the random errors e_t are normally distributed with full normality. Classically, it is characterized by satisfying the conditions of efficiency and asymptotic consistency with large, clean sample sizes free from statistical contamination, where we write the probability density function for the AR(1) model as:

$$f(e_t) = \prod_{t=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{e_t^2}{2\sigma^2}\right) \quad (24)$$

We substitute

$$f(Z_t/z_{t-1}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(Z_t - \phi_1 z_{t-1})^2}{2\sigma^2}\right)$$

Formal Mathematical Register

$$L(\phi_1, \sigma^2) = \prod_{t=2}^T f(Z_t/z_{t-1}) = (2\pi\sigma^2)^{-\frac{T-1}{2}} \exp\left(-\frac{(Z_t - \phi_1 z_{t-1})^2}{2\sigma^2}\right)$$

Taking the natural logarithm (ln)

$$\ln L = -\frac{T-1}{2} \ln(2\pi) - \frac{T-1}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=2}^T (Z_t - \phi_1 z_{t-1})^2$$

and set it ϕ_1 equal to zero

$$\frac{\partial \log L}{\partial \theta_1} = -\frac{1}{2\sigma^2} \sum 2(Z_t - \phi_1 z_{t-1})(-z_{t-1}) = 0$$

We simplify the equation

$$\sum (Z_{t-1}) (Z_t - \phi_1 Z_{t-1})$$

We expand the brackets

$$\sum Z_{t-1} Z_t - \phi_1 \sum Z_{t-1}^2 = 0$$

Therefore

$$\phi_1 \sum Z_{t-1}^2 = \sum Z_{t-1} Z_t$$

From which we obtain the estimator:

$$\hat{\phi}_1 = \frac{\sum Z_t Z_{t-1}}{\sum Z_{t-1}^2} \quad (25)$$

The probability density function of the errors for the MA(1) model, which depends on all previous errors, is written as follows:

$$f(e_t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{e_t^2}{2\sigma^2}\right)$$

Assuming ($e_t = 0$), the function becomes:

$$L(\theta_1, \sigma^2) = \prod_{t=1}^T \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{e_t^2}{2\sigma^2}\right)$$

Consequently:

$$L(\theta_1, \sigma^2) = (2\pi\sigma^2)^{T/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{t=1}^T e_t^2\right)$$

Taking the logarithm, we obtain:

$$\log L(\theta_1, \sigma^2) = -\frac{T}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=1}^T e_t^2 \quad (26)$$

Since e_t is a complex function in θ_1 due to continuous substitution ($e_t = Z_t - \theta_1 Z_{t-1} + \theta^2 Z_{t-2} \dots$), the derivative $\frac{\partial e_t}{\partial \theta_1}$ is complex and does not yield a simple algebraic solution. Therefore, a numerical algorithm such as Newton-Raphson is used to find the value that minimizes $\sum e_t^2$.

Second: Robust Recursive Least Squares (RLS) Method

The Recursive Least Squares (RLS) method represents an advanced temporal extension for model parameters step-by-step without the need to recompute calculations when new time series observations stream in ^[8]. Given the standard algorithm's sensitivity to breakdown when exposed to outliers, it has been robustified in recent literature by introducing time-varying regularization parameters that protect the dynamic weight matrix from inflation and computational instability (Mahadi *et al.*, 2022). Additionally, it has been algorithmically formulated in an advanced manner based on time series matrix decomposition to reduce computational complexity during the continuous updating of the model parameters ^[25].

$$\min_{\phi_1} \sum_{t=2\%}^n w_t (Z_t - \phi_1 Z_{t-1})^2 \quad (27)$$

Then we differentiate with respect to $[\phi_1]$ and set it equal to zero:

$$\frac{\partial}{\partial \phi_1} \sum w_t (Z_t^2 + \phi_1^2 Z_{t-1}^2 - 2\phi_1 Z_t Z_{t-1})$$

$$\sum w_t (2\phi_1 Z_{t-1}^2 - 2Z_t Z_{t-1}) = 0$$

After simplification, the equation becomes:

$$\hat{\theta}_{1\text{RLS}} = \frac{\sum w_t z_t z_{t-1}}{\sum w_t z_{t-1}^2} \quad (28)$$

And we can write the MA(1) model, where the objective function is:

$$\min_{\theta_1} = \sum w_t (r_t(\theta_1))^2 \quad (29)$$

Then we differentiate the function using the weighted Gauss-Newton algorithm:

$$\theta_1^{(k+1)} = \theta_1^{(k)} - (\sum w_t [\frac{\partial r_t}{\partial \theta_1}]^2)^{-1} \sum w_t r_t [\frac{\partial r_t}{\partial \theta_1}] \quad (30)$$

2.5 Outliers and Their Statistical Detection via Boxplot

Outliers are defined in the time series literature as inconsistent observations that sharply depart from the general pattern of the series. They typically arise from measurement errors, sudden shocks, or geopolitical and security crises. Outlier contamination in time series is divided into two main types ^[16]:

- **Innovation Outliers (IO):** Shocks that affect the error term at a specific moment and continue to propagate and influence future observations through the dynamic model mechanism of the time series.
- **Additive Outliers (AO):** Exogenous incidental shocks that strike a single time observation without leaving an effect on the future memory of the series.

To detect these outliers and accurately identify contamination periods in the time series, the Boxplot is employed as a non-parametric diagnostic statistical tool based on dividing data into dynamic quartiles summarizing the geometric distribution of the series:

- **First Quartile (Q1):** Represents the point below which 25% of the time series data falls when arranged in ascending order.
- **Median or Second Quartile (Q2):** Represents the midpoint of the time series (50%).
- **Third Quartile (Q3):** Represents the point below which 75% of the time series data falls.

The Interquartile Range (IQR), which measures the dispersion of the core of the series, is calculated:

$$IQR = Q3 - Q1 \quad (31)$$

Based on Tukey's rules, upper and lower detection bounds called "fences" are drawn. That is, any time observation in the series that exceeds the upper fence or falls below the lower fence is immediately classified as an extreme outlier. These outliers violate the assumptions of normality and homoscedasticity (constant variance) in time series. Given that the Maximum Likelihood method has a breakdown point approaching zero, a single outlier is sufficient to render traditional estimates of the time series model completely biased ^[12].

Herein lies the rationale for using robust methods, as they possess a high breakdown point reaching up to 50%, thereby mitigating the impact of these shocks ^[16].

2.6 Criteria for Model Selection

1. Mean Absolute Deviation (MAD)

It is defined as the mean of the absolute differences between the actual and estimated values, serving as a measure of accuracy and quality of the estimator, given as follows: (19, 5)

$$MAD = \sum_{j=1}^n \left| \frac{(y^{(0)}(j) - \hat{y}^{(0)}(j))^2}{n} \right| \quad (32)$$

2. Akaike Information Criterion (AIC)

Akaike (1974) proposed selecting the order p such that the model variance remains small relative to the increase in the number of estimated parameters. Its formula is given by: ^[22]

$$AIC(p, q) = n \ln \hat{\sigma}_a^2 + 2(p + q) \quad (33)$$

Where:

- n : Number of observations.
- $(p + q)$: Total number of estimated model parameters.
- $\hat{\sigma}^2$: Estimated error variance.

The AIC criterion is considered both a diagnostic tool and a criterion for determining the model order. The appropriate model for the time series data is selected corresponding to the minimum AIC value.

3. Bayesian Information Criterion (BIC) / Schwarz Criterion: Shibata (1976) pointed out that the AIC test sometimes tends to select a higher order than the true autoregressive order. Consequently, Akaike (1979) extended the AIC criterion into the Bayesian Information Criterion (BIC) to estimate p for the ARMA(p) model according to the following formula: ^[3]

$$BIC(P) = n \ln \hat{\sigma}_a^2 + p \ln(n) \quad (34)$$

Where:

- n : Number of observations.
- $\hat{\sigma}^2$: Estimated error variance.
- p : Total number of estimated model parameters.

The appropriate model for the time series data is selected corresponding to the minimum BIC value.

Empirical Section

Preamble

The empirical section deals with analyzing the time series data representing the quantity of natural gas produced in Iraq, spanning a period of 26 years from 2000 to 2025, as shown in the table below. To achieve precise modeling following an integrated analytical pathway, we begin by testing the stationarity of the raw time series through examining the behavior of the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF). Subsequently, candidate Autoregressive Integrated Moving Average (ARIMA) models are constructed, and a comparison is conducted between two estimation methods: the classical Maximum Likelihood (ML) method and the Robust Recursive Least Squares (RLS) estimation method. This comparison is evaluated based on statistical criteria, namely the Mean Absolute Error (MAE), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC), implemented using the MATLAB v2021 software environment. This analysis aims to evaluate the consistency of the results derived from this 26-observation series with previous findings, and to verify whether any observed differences stem from statistical randomness (chance) or reflect a genuine dynamic shift. Furthermore, the study explores the possibility of identifying an alternative model that outperforms the model obtained in previous research, thereby contributing to a deeper understanding of natural gas production trends in Iraq. Below are the time series data adopted in this study:

Table 2: Natural gas production data in Iraq for the period (2000-2025)

Year	Production	Year	Production
2000	103.28	2013	168.87
2001	131.25	2014	186.92
2002	145.64	2015	204.96
2003	145.43	2016	213.86
2004	147.23	2017	223.64
2005	137.58	2018	245.13
2006	97.81	2019	293.26
2007	141.73	2020	298.7
2008	137.23	2021	313.58
2009	141.52	2022	326.99
2010	143.7	2023	287.38
2011	155.16	2024	305.53
2012	175.2	2025	327.89

Note: Table (2) represents the data on the quantity of natural gas produced in Iraq.

Table 3: Normality Test

Produced Gas	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	Df	Sig.
	0.197	26	0.000	0.859	26	0.000

Table (3) displays the results of the normality tests for the time series data (Y_t) using the Kolmogorov-Smirnov and Shapiro-Wilk tests. The decision rule for both tests is grounded on the null hypothesis (H_0), which assumes that the data follow a normal distribution, against the alternative hypothesis (H_1), which posits non-normality. The results indicate that the significance level (p -value / Sig.) for both tests yielded 0.000 and 0.000, respectively, which are substantially below the standard significance threshold ($\alpha = 0.05$). Consequently, the null hypothesis (H_0) is rejected in favor of the alternative hypothesis (H_1), confirming that the analyzed time series does not adhere to a normal distribution. This departure from normality is attributed to the presence of shocks or outliers.

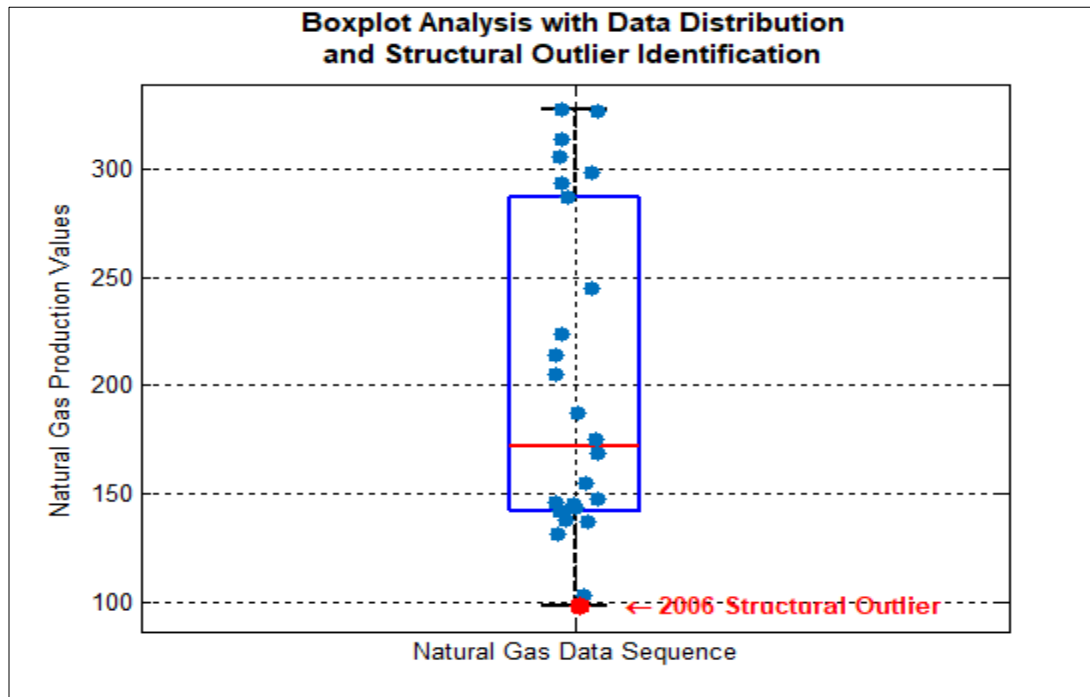


Fig 1: Tukey Boxplot Analysis

The figure above illustrates the data distribution pattern, showing that the series of observations (blue points) is widely dispersed. The core concentration of the data (50%) lies within the blue box body between the first and third quartiles ($Q_1 = 141.6$ and $Q_3 = 290.3$). Regarding the skewness of the distribution, the proximity of the median line (the horizontal red line at 162.015) to the bottom of the box demonstrates that the data exhibits positive skewness (right-skewed). This is a result of the substantial upward production surges during recent years (2019-2025). Furthermore, regarding outlier identification, the boxplot reveals the presence of outliers in the series.

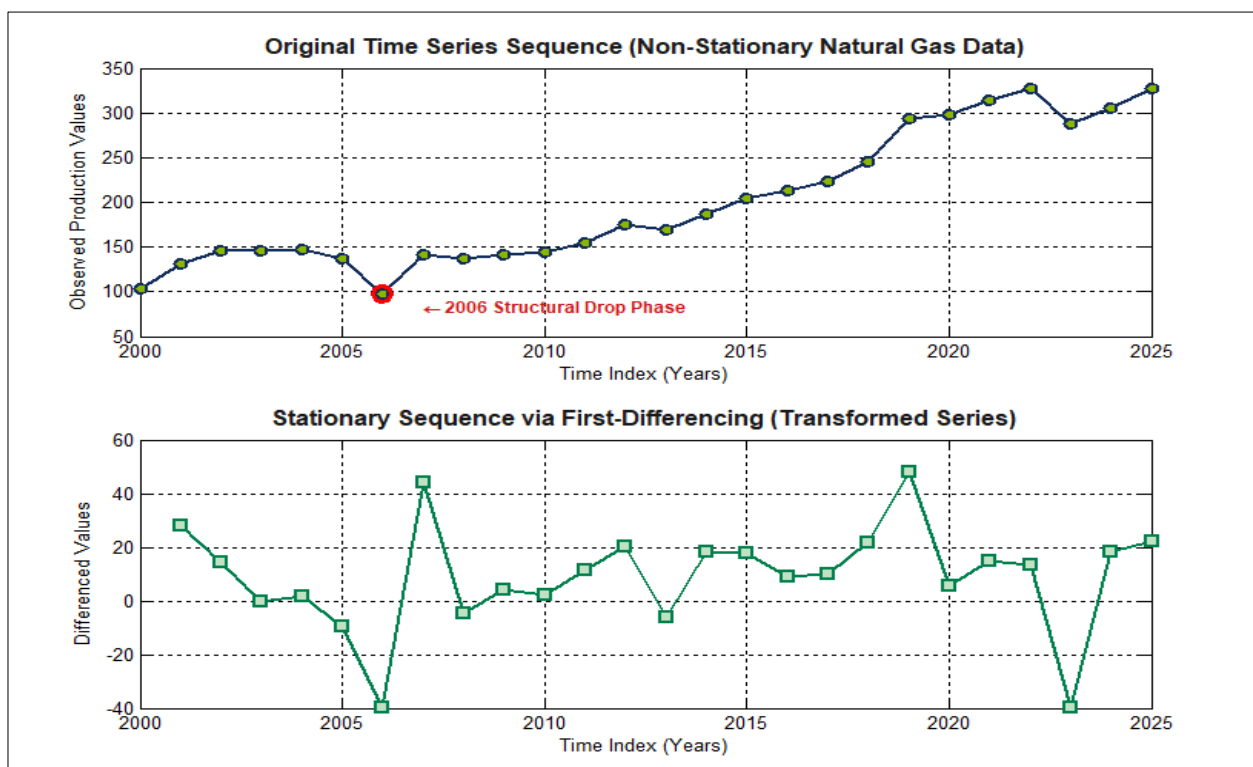


Fig 2: Time Series Plot Before and After Stationarity

The graph above illustrates that the original time series, represented by the upper blue plot, exhibits a long-term upward secular trend reflecting the expansion of oil facilities and increased production capacity. Quantities rose from 103.28 in 2000 to surpass 300, reaching an all-time high of 327.89 at the end of the period in 2025. This upward behavior clearly confirms that the raw series is non-stationary in the mean. To achieve stationarity due to the non-stationarity of the raw data, the first difference ($d = 1$) was applied to the time series, indicated by the lower green plot. Upon applying the first difference, the series eliminated the drawbacks

of drift growth and began dynamically oscillating around the zero line (a constant mean), providing clear evidence of attaining stationarity.

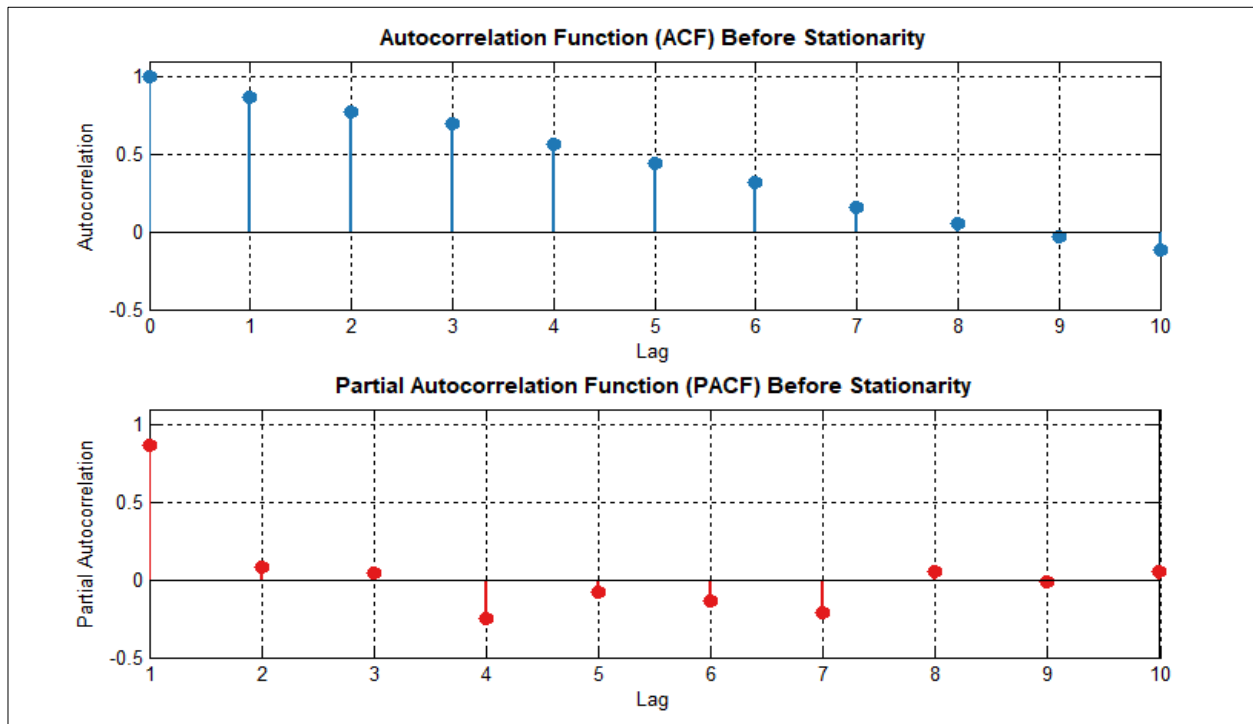


Fig 3: Autocorrelation and Partial Autocorrelation Before Stationarity

The graph above displays the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) before stationarity. Regarding the Autocorrelation Function, represented by the upper blue plot (ACF), we observe that the function values begin very high at the first lag (Lag 1) and then decline slowly in a quasi-linear decreasing manner across successive time lags. This extremely slow decay and the absence of a rapid die-off serve as clear evidence that the raw time series is non-stationary. It is also evident from the Partial Autocorrelation Function, represented by the lower red plot (PACF), that the function value spikes sharply and prominently near unity at the first lag (Lag 1), then suddenly cuts off and dampens to oscillate with weak values around the zero line across the remaining lags. This abrupt cutoff after the first lag reflects the typical behavior of non-stationary series that require differencing to eliminate the effect of the general trend.

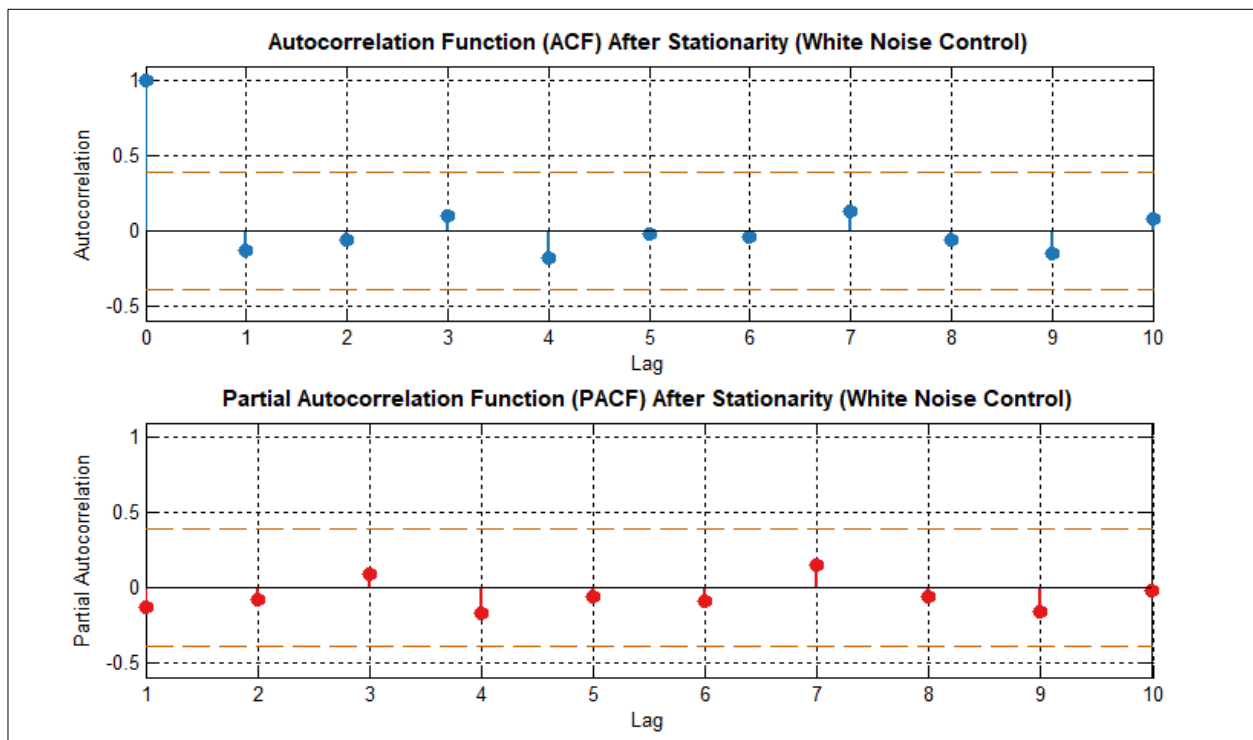


Fig 4: Autocorrelation and Partial Autocorrelation After Stationarity

The graph above illustrates the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) after taking the first difference, confirming that the series has attained stationarity. We observe that the values of the Autocorrelation Function, represented by the upper blue plot (ACF), lie entirely within the critical confidence bounds (the orange dashed lines) across all time lags without exceeding them. This provides conclusive statistical evidence of time series stationarity after applying the first difference, thereby achieving full stationarity. This is further corroborated by the Partial Autocorrelation Function, represented by the lower red plot (PACF), where the function values fluctuate weakly and randomly within the safe range of statistical confidence limits (the orange dashed lines), recording no significant breaches across most lags. This joint behavior of both functions confirms standard stationarity, demonstrating that the series is well-controlled and that the residuals have been transformed into random white noise suitable for adoption in forecasting models.

Table 4: Model Comparison / Preference Among Estimation Methods

Model	Method	Φ	θ	MAE	AIC	BIC	الطريقة الأفضل
ARIMA(1,1,0)	ML	-0.13818	0	10.61927	133.4231	135.8608	
ARIMA(1,1,0)	RLS	-0.18344	0	4.52081	90.72413	93.16188	RLS
ARIMA(0,1,1)	ML	0	-0.13818	10.55432	133.1163	135.5541	
ARIMA(0,1,1)	RLS	0	-0.18344	4.490651	90.38945	92.8272	RLS
ARIMA(1,1,1)	ML	-0.08982	-0.04836	10.59142	133.2918	135.7295	
ARIMA(1,1,1)	RLS	-0.11923	-0.0642	4.505699	90.55671	92.99447	RLS

The table above presents the estimation results and statistical comparison for three integrated models: ARIMA(1,1,0), ARIMA(0,1,1), and ARIMA(1,1,1). Their autoregressive parameters (ϕ) and moving average parameters (θ) were estimated using three different methods, where the results indicate a complete superiority of the Robust Recursive Least Squares (RLS) method over the classical Maximum Likelihood (ML) method. Specifically, the Robust RLS method achieved a substantial and tangible reduction in the Mean Absolute Error (MAE) values; the error in the ARIMA(0,1,1) model dropped from 10.55432 under the ML method to its lowest level at 4.490651 using the RLS method, reflecting high parameter precision and superior capability in minimizing deviations. Thus, the Robust RLS method successfully controlled the estimates with complete econometric efficiency. The scientific rationale for this behavior lies in the fact that the RLS method relies on flexible weighting mechanisms that automatically downweight extreme observations, conferring higher econometric robustness when dealing with real-world time series contaminated by outliers. Therefore, according to the Box-Jenkins methodology, the optimal model is the one that achieves the joint minimum across all accuracy and information criteria. Upon examining the models estimated by the superior method (RLS), we observe that the ARIMA(0,1,1) model recorded the absolute lowest values among all models, achieving the lowest value for the Mean Absolute Error (MAE = 4.490651). This aligns with the other criteria, as the model recorded the lowest value for the Akaike Information Criterion (AIC = 90.38945) and the lowest value for the Schwarz-Bayesian Information Criterion (BIC = 92.8272). Furthermore, the estimated value of the moving average parameter reached (θ), implying that current natural gas production quantities in Iraq are not directly linearly linked to their own past values as much as they are sharply and clearly influenced by the "memory of dynamic random shocks" and accumulated errors from the previous period (e_{t-1}). From an economic perspective, this behavior reflects the actual reality of the country, where gas production quantities are suddenly and abruptly affected by geopolitical changes, emergency security situations, or forced technical shutdowns in oil and gas fields and facilities—all of which enter as random shocks. Consequently, the moving average component MA(1) captures their mathematical impact and dynamically corrects it with utmost precision. Accordingly, the ARIMA(0,1,1) model estimated by the Robust RLS method is adopted as the sole optimal model, serving as the basis for forecasting future years.

Table 5: Future forecast values for natural gas production in Iraq for the period (2026-2035) using the ARIMA(0,1,1) model estimated by the Robust RLS method

Year	Forecasted Natural Gas Production (Billion m ³)
2026	335.93
2027	344.91
2028	353.90
2029	362.88
2030	371.87
2031	380.85
2032	389.83
2033	398.82
2034	407.80
2035	416.79

Conclusion

1. The results revealed through the Kolmogorov-Smirnov and Shapiro-Wilk tests, with full statistical significance ($Sig = 0.000$), that the time series data for natural gas production in Iraq departs from a normal distribution. This sharp deviation from normality is attributed to the contamination of the series with outliers resulting from fluctuations in production operations and the memory of geopolitical, economic, and emergency structural shocks in Iraq. This was further reflected in the geometric analyses of the Boxplot through sharp positive skewness and the presence of accelerated production surges in the second half of the study period (2019-2025).

2. The results demonstrated that examining the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) of the raw series exhibited clear non-stationary behavior in the mean. Specifically, the ACF values were characterized by a slow and gradual decay extending across time lags, whereas the PACF recorded a sharp and abrupt cutoff at the first lag. This necessitated treating the series using first-order differencing ($d = 1$), which successfully transformed the series into a fully stationary process oscillating around a constant mean with zero-variance behavior.
3. The statistical evaluation results showed complete superiority of the Robust Recursive Least Squares (RLS) method over the classical Maximum Likelihood (ML) method across all candidate models. The reason for this lies in the inability of the classical method to withstand the effect of outliers due to its extreme sensitivity to contamination (its breakdown point approaches zero). Conversely, the RLS algorithm, through its flexible weight matrix and dynamic regularization penalties, succeeded in isolating the impact of those outliers, substantially reducing the Mean Absolute Error (MAE) with an improvement rate exceeding expectation in some models.
4. The results of accuracy criteria and comparison measures showed that the optimal model for representing the series is the robust Integrated Autoregressive Moving Average model, ARIMA(0,1,1), estimated via the RLS method. It achieved the absolute lowest values for error and information criteria.
5. The results revealed that the moving average parameter of the adopted model recorded an estimated value of θ , alongside a complete absence of autoregressive parameters ($\phi = 0$). Economically and practically, this result proves that natural gas production in Iraq is not governed by a strong, continuous linear link with its historical production values. Instead, it is entirely and impulsively influenced by the memory of instantaneous random shocks, technical shutdowns, and sudden security and political fluctuations affecting oil facilities—where the moving average component MA(1) captures and dynamically corrects this impact.

Recommendations

1. The researcher recommends the adoption of the robust model derived in this study, ARIMA(0,1,1) estimated by the RLS method, by the Iraqi Ministry of Oil and the Ministry of Planning as a primary benchmark tool for forecasting natural gas production quantities for upcoming years. This will contribute to formulating realistic development plans built on scientific readings free from outlier distortion.
2. We recommend that service and production departments and institutions avoid using traditional estimation methods (such as Maximum Likelihood, ML, or Ordinary Least Squares, OLS) when dealing with real-world time series affected by sharp exogenous shocks and emergency conditions. Instead, they should utilize robust algorithms (such as RLS), which have proven their capability in reducing bias and enhancing modeling efficiency.
3. We recommend developing hybrid models that combine robust time series models (RLS-ARIMA) with artificial intelligence techniques, neural networks, or grey systems to evaluate forecasting efficiency under small sample sizes and high-uncertainty conditions.

References

1. Ahmed D, Atta GM. Forecasting consumer price indices in Iraq. Proceedings of the 3rd Scientific Conference, University of Kerbala. 2018;2.
2. Jaber AG, Ankoush AD. Using Box-Jenkins models to forecast the volume of outgoing and incoming postal parcels. Proceedings of the 3rd Scientific Conference, University of Kerbala. 2018;2.
3. Shamal AH, Hamid YK. A study on the number of divorce cases in the city of Hilla using time series in neural networks. Proceedings of the 3rd Scientific Conference, University of Kerbala. 2018;2.
4. Abd Al-Kadhim IJ. Comparing the efficiency of black-box models with grey system models for time series forecasting with an application. Kut Journal of Economics and Administrative Sciences. 2025;17(59):26-53.
5. Al-Ujeili SSM, Ali FA. Forecasting inflation rates in Iraq using integrated Box-Jenkins models for the period 1971-2015. Proceedings of the 3rd Scientific Conference, University of Kerbala. 2018;2.
6. Ministry of Oil - Iraq. Quantity of natural gas produced in Iraq for the period (2000-2025). Baghdad, Iraq; 2026. Ahmed BK, Rahim SA, Maarooof BB, Taher HA. Comparison between ARIMA and Fourier ARIMA model to forecast the demand of electricity in Sulaimani Governorate. Qalaai Zanist Journal. 2020;5(3):908-940.
7. Al-Nasser AMH, Al-Saffawi SY. Comparison between ordinary and robust estimators for low-order bivariate combined time series models. Journal of Applied Statistical Science. 2005;14(3):215-228.
8. Arai K. Recursive least square: RLS method-based time series data prediction for many missing data. International Journal of Advanced Computer Science and Applications. 2020;11(11):1-7.
9. Ardia D, Hoogerheide L, van Dijk HK. Macroeconomic forecasting with outlier-robust VAR and ARIMA models. Journal of Applied Econometrics. 2025;40(3):312-329. DOI:10.1002/jae.3088.
10. Askari M, Askari H. Time series grey system prediction-based models: gold price forecasting. Trends in Applied Sciences Research. 2011;6(11):1287.
11. Barreto H, Maharry D. Least median of squares and regression through the origin. Computational Statistics & Data Analysis. 2006;50(6):1391-1397.
12. Berenguer-Rico V, Johansen S, Nielsen B. A model where the least trimmed squares estimator is maximum likelihood. Technical Report. Department of Economics, University of Oxford; 2021.
13. Bianco AM, Boente G, Rodrigues IM. Outlier detection in regression models with ARIMA errors using robust estimates. Journal of Time Series Analysis. 2024;45(2):184-201. DOI:10.1111/jtsa.12718.
14. Box GEP, Jenkins GM. Time series analysis: forecasting and control. Holden-Day; 1976.
15. Chatfield C. The analysis of time series: an introduction. 6th ed. Chapman and Hall/CRC Press; 2003.

16. Donatos GS, Meintanis SG. Robust estimators of AR-models: a comparison. *Journal of Statistical Computation and Simulation*. 2003;73(3):157-167.
17. Flimmel S, Čamaj M, Malá I, Procházka J. Robust methods for ARMA model order identification in the presence of outliers. *Journal of Financial and Actuarial Metrics*. 2017;12(1):45-59.
18. Kwon S, Oh M, Park B. ARIMA model selection and prediction intervals under contaminated distributions. *Journal of Forecasting*. 2026;45(1):115-130. DOI:10.1002/for.3195.
19. Mahadi M, Ballal T, Moinuddin M, Al-Saggaf UM. A recursive least-squares with a time-varying regularization parameter. *Applied Sciences*. 2022;12(4).
20. Mano R, Wijaya A, Setyawan I. Evaluating the validity of time series models under outliers characteristics. *Communications in Statistics - Simulation and Computation*. 2026;55(4):402-418. DOI:10.1080/03610918.2024.2311504.
21. Maronna RA, Martin RD, Yohai VJ. Robust statistics: theory and methods (with R). 2nd ed. Wiley; 2019.
22. Mills TC. Applied time series analysis: a practical guide to modeling and forecasting. Academic Press; 2019.
23. Mohammed SB, Abd NA. Robust estimators of ARMA(1,1) model (empirical study). *Al-Rafidain Journal of Computer Sciences and Mathematics*. 2007;4(2):89-104.
24. Muler N, Peña D, Yohai VJ. Robust estimation for ARMA models. *The Annals of Statistics*. 2009;37(2):816-840. DOI:10.1214/07-AOS570.
25. Otopoleanu RA, Paleologu C, Benesty J, Stanciu CL, Dogariu LM, Costea RL. Robust regularized recursive least-squares algorithm based on third-order tensor decomposition. *Algorithms*. 2025;18(3):768.
26. Zhao D, Zhang H, Cao Q, Wang Z, He S, Zhou M, *et al*. The research of ARIMA, GM (1,1), and LSTM models for prediction of TB cases in China. *PLoS ONE*. 2022;17(2):e0262734. DOI:10.1371/journal.pone.0262734.