JEAS, Vol. 31 No. 150 (2025) pp (155-173)
P-ISSN 2518-5764 - E-ISSN 2227-703X
DOI: https://doi.org/10.33095/2sfxwg20

Journal of Economics and Administrative Sciences

A Comparative Framework of Hybridization of ARIMA and SARIMA Models for Forecasting Gold Price Movements in Iraq (2015–2025)

Hiba Dhahir Alwan *
Department of Electrical Engineering; College of Electrical Engineering; University of Technology; Country: Iraq
*** Corresponding author**

Suhail Najm Abdulla
Department of Statistics; College of Administration and economic; University of Baghdad; Country, Iraq.

Received: 12/8/2025 **Accepted:19/11/2025** **Published: 1/12/2025**

Abstract:
Predicting gold prices is crucial for policymakers, investors, and financial planners, especially in commodity-dependent economies like Iraq. This study examines the forecasting performance of Autoregressive Integrated Moving Average (ARIMA) and Seasonal ARIMA (SARIMA) models using daily gold price data in Iraq from (April 30, 2015, to April 30, 2025). After preprocessing and testing for stationary, both models were estimated using Maximum Likelihood Estimation (MLE) and further refined with modern machine learning post-processing, Maximum Likelihood Estimation (MLE) method demonstrated greater stability and interpretability compared to the machine learning-based approach. However, the application of Gradient Boosted Trees (GBDT) to the residuals of the SARIMA model further enhanced short-term predictive performance. Therefore, this paper concludes that the combination of SARIMA and (GBDT) creates a robust hybrid framework suitable for forecasting gold prices in volatile and data-constrained economies such as Iraq Model performance was evaluated using (RMSE, MAE, and MAPE) on both training and testing sets. The results indicate that SARIMA (1,1,1)(1,1,1) captured the seasonal patterns of gold prices more effectively than the baseline ARIMA (1,1,2). While ARIMA served as a benchmark for modeling non-seasonal dynamics, SARIMA provided the final forecasts adjusted for seasonality. Forecasts for (May 2025 to April 2026) show a consistent upward trend with minor seasonal fluctuations, validating the model's applicability for decision-making in the Iraqi context. These findings demonstrate that SARIMA offers an interpretable and efficient framework for short- and medium-term forecasting, providing a reliable tool in volatile economic environments.
Keywords: ARIMA, SARIMA, GBDT, forecasting, gold price, Iraq.

1. Introduction:

As widely recognized, gold should be considered a strategic asset, serving both as an inflation hedge and a haven during financial crises. In Iraq, geopolitical instability and volatility in oil revenues have been major macroeconomic characteristics, and accordingly, gold has been affected in its capacity to preserve individual and institutional wealth. The ability to predict gold prices is therefore strategically important for policymakers, investors, and financial institutions in the region.

Over the years, numerous forecasting methods have been proposed to capture the dynamic and nonlinear tendencies of gold prices. The Autoregressive Integrated Moving Average (ARIMA) and its seasonal counterpart (SARIMA) were originally designed to provide classical statistical models with strong predictive power for financial time series due to their theoretical rigor and interpretability (Gong, 2024). However, with the rise of computational power and the availability of large datasets, machine learning (ML) and deep learning (DL) models have emerged as powerful alternatives for modeling complex temporal dependencies in financial data (Bala & Singh, 2022; Liu & Long, 2020; C. Zhang et al., 2024). Deep learning architectures such as Deep Neural Networks (DNN), Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), and hybrid CNN-LSTM models have been successfully applied across various markets (Baradaran et al., 2024; Kanwal et al., 2022; Salim & Djunaidy, 2024).

Although these models can achieve high accuracy compared to traditional approaches, they are often opaque and may require extensive data preprocessing and optimization (Livieris et al., 2020; P. Zhang & Ci, 2020). In contrast, ARIMA and SARIMA remain very useful in cases of moderate nonlinearity and strong seasonality, particularly when combined with robust validation methods (Hajek & Novotny, 2022; Srivastava et al., 2024). Despite the global advances in predictive modeling, relatively little academic research have focused on gold price forecasting within the Iraqi context. Most existing literature is either regionally generic or relies heavily on complex deep learning models without using classical time series methods as baselines. This gap limits the ability of financial practitioners in Iraq to implement interpretable and empirically validated forecasting tools suited to their economic realities.

The primary contribution of this paper is to provide a straightforward, statistically supported forecasting approach with practical applications for regional stakeholders, while also enhancing the broader knowledge base in interpretable financial forecasting. While previous studies have explored gold price prediction in other countries using deep learning or hybrid approaches (Mohtasham Khani et al., 2021; Mousapour Mamoudan et al., 2023; Nallamothu et al., 2024; Zhao & Yang, 2023), this study reaffirms the applicability of classical models in a data-rich but under-studied national context. By integrating traditional statistical modeling with robust diagnostic tools and alternative estimation procedures, a new hybrid statistical framework is presented that combines simplicity with analytic rigor. Furthermore, this Iraq-specific case study provides an essential reference point for other emerging economies facing similar economic conditions (Das et al., 2022; Sarangi et al., 2021; Singh et al., 2022; Zangana & Obeyd, 2024; F. Zhang & Wen, 2022).

2. Literature Review:

The price of gold is a major research point in the financial time series modeling as gold plays an important strategic aspect in the stability of the economy, hedging and asset investment. ARIMA, as well as other traditional econometric methods, have been used. Nevertheless, recent innovations in the field of machine learning and deep learning have presented new frameworks that enhance the predictive performance and represent nonlinearities of financial data. A dual-stage advanced deep learning framework that is able to use long-sequence forecasting to multivariate financial time series was presented by (Bala & Singh, 2022).

Their model is based on exploiting feature-extraction as well as sequential modeling to forecast market behavior in long horizons. In equal measure, (C. Zhang et al., 2024) provided an overarching survey of current developments in deep learning algorithms in the context of financial time series and a shift toward hybridized neural networks. In regard to the Iranian gold market, (Baradaran et al., 2024) compared GRU and CNN-LSTM models and found that recurrent-based hybrid architectures are better suited to describing temporal dynamic when compared to standalone models. This aligns with findings by other scholars who have employed CNN-LSTM in processing the time-series in the form of images (Salim & Djunaidy, 2024), which is a novel approach of processing sequential data through visual pattern recognition. Das et al. (2022) used fuzzy inference system and extreme learning machines to forecast the oil prices and the gold prices. In their study, fuzzy systems have fore fronted the ability to predict the economic situation in the interpretation and strength. The advantages of utilizing the external textual data were analyzed even by (Hajek & Novotny, 2022), who also examined the fuzzy rule-based systems, taking into consideration the news sentiment in modeling the gold prices. The authors (Kanwal et al., 2022) proposed hybrid BiCuDNNLSTM-1dCNN stock price forecasting mode based on the combination of convolutional layers to obtain the feature-manipulating capabilities and the LSTM layers to address the temporal domain of data. They offered a superior deep learning model that forecasts time series on the stock prices signifying a deeper architecture with effective regularization (Liu & Long, 2020). After the issue of overfitting and the generalization of the machine, (Livieris et al., 2020) suggested a novel validation scheme to suggest the improvement of accuracies in deep learning time-series applications. Several prior works have also tested hybrid forecasting models in economic and financial contexts, confirming the utility of combining linear and non-linear modeling techniques for improved predictive accuracy (Mohtasham Khani et al., 2021) The case study of (Mousapour Mamoudan et al., 2023) is a prediction of the financial markets: the world gold market using the evaluation of hybrid neural based metaheuristics. They integrated various optimization techniques to improve convergence speed and reduce prediction error. Similarly, (Srivastava et al., 2024) have used deep learning and association mining to produce time series predictions and they were able to identify latent trends in time. In their work, (Nallamothu et al., 2024) offered a model that is the integration of GARCH and LSTM and has skew and kurtosis as predictors to predict prices of gold. This goes together with the new interest in the pursuit of the emergence of the statistics and the neural aspects. One of the articles proposed a hybrid model of machine learning in predicting the prices of gold in India, thereby validating the fact that the model can also be relevant in less developed economies (Sarangi et al., 2021). Regarding the approaches to pure machine learning of gold prices, (Singh et al., 2022) focused on the implementing of classification and regression models in the given work, and (Wang, 2022) applied the wavelet packet decomposition and the stochastic deep learning model to forecast the futures of the metals in order to raise the opportunities of the multi-resolution analysis. Some recent papers on the Iraqi and Gulf area have addressed the question of the relevance of hybrid time series models to perform forecasting in the fields of finance and economy. They suggested a new WARIMA hybrid model to get Euro Dinar trade patterns, and this model has found it more accurate when compared with conventional ARIMA models (Wazeer & Hameed, 2022). Likewise, Adnan and Ahmed (2025) combined the use of a hybrid ARIMA-GARCH model to forecast the stock market indicators in the Gulf countries, demonstrating that the modeling of volatility can be a good option when forecasting time series. Specifically, the combination of the linear and non-linear modeling methods was highlighted in the context of the Iraqi SARIMA-NARNN hybrid predictive model (Hamel & Abdulwahhab, 2022).

Further, (Alwan & Abdulla, 2022) simulated NAGARCH and APGARCH, and in their simulation study, they focused on the capability of advanced GARCH models in modeling conditional volatility in financial time series. All these studies are in agreement with the fact that hybrid or extended implementations of ARIMA/SARIMA models need to be used in order to make better and context specific forecasts.

The report of (Zangana & Obeyd, 2024) in the field of knowledge was a system that does not rely on deep learning but is implemented in the Iraqi context. Their strategy provided the one of the regional uses of AI for the gold price forecasts and it was based on the local market tendencies. In a similar effort, (P. Zhang & Ci, 2020) modeled the price of gold via deep belief networks (DBN) and showed the usefulness of unsupervised pre-training to learn hierarchies in the data. (F. Zhang & Wen, 2022) relied on novel structures to assess the long-term carbon prices through the application of deep learning technologies to environmental markets. The techniques can be applied to financial time series because they are transferable to their field, although they have a different domain. Lastly, (Zhao & Yang, 2023) proposed a combination deep learning framework on the prediction of stock price movement by interacting with various models and the attention mechanism to enhance the effectiveness of prediction. Several studies have employed advanced volatility and time series models in financial and commodity markets. (Abdulla & Dhaher Alwan, 2022) utilized APGARCH and AVGARCH models, incorporating both Gaussian and non-Gaussian distributions, to model exchange rate volatility. Similarly, (Mohammed & Yadkar, 2015) applied ARCH and GARCH models for predicting the daily closing price of the Iraqi Stock Exchange index. In the context of nonstationary time series, (Hmood, 2025) explored nonparametric estimation techniques, offering insights into flexible modeling approaches. Furthermore, (Mousa & Hmood, 2023) employed the MS-GARCH model to forecast Saudi crude oil prices, demonstrating the applicability of regime-switching volatility models in energy markets. These studies collectively highlight the versatility and robustness of volatility modeling techniques across different financial and commodity sectors.

To sum up, the literature experience shows an undeniable shift in terms of the classical method of time series modeling to deep learning systems. Although deep learning has been an enormous step towards predictive abilities, the problems of interpretability, overfitting, and regional applicability still exist. This study fills this research gap by confirming classical models of ARIMA/SARIMA in the unexplored market setting of the Iraqi gold market and therefore provides a viable and explainable alternative that lies in the realm of statistical rigor.

While the literature highlights the theoretical and methodological contributions of ARIMA/SARIMA and hybrid models, the practical implementation and empirical evaluation of these models in the Iraqi context are addressed in the subsequent sections.

3. Method:

In this paper, classical forecasting methodologies which are ARIMA and SARIMA will be used to model the daily gold prices in Iraq. The overlying methodology involves several decisive steps: preprocessing of data, stationarity and normality tests, modeling and estimation, diagnostics and out-of-sample forecasts. This part describes each of these phases in detail and gives the mathematical background of the models.

This paper is expected to address this gap because the paper presents the predictive capacity of ARIMA and SARIMA models to predict the price of gold in Iraq. The targeted objectives are the following:

- To use and compare the classical time series models (ARIMA/ SARIMA) in forecasting the price of gold per day in Iraq.
- To conduct frequent statistical testing, such as stationarity (ADF test), residual diagnostics (Ljung-Box test) and normality checks.
- To increase the parameter estimation accuracy by comparing the Maximum Likelihood Estimation (MLE) with the modern machine learning-based estimation techniques.
- To identify the most optimal model using AIC, BIC, RMSE and the forecast performance.
- To provide benchmarking among future researchers to conduct that in the local context of Iraq by utilizing complex deep learning models.

3.1 Model 1: Autoregressive Integrated Moving Average (ARIMA):

In (1976) Box & Jenkins introduced a model that transforms a non-stationary series into an average to stationary series by taking a finite number of lags, where the lag must be a positive integer. The mathematical model for ARIMA (p,d,q) is as follows:

$\emptyset(B)\omega_t = \theta(B)\varepsilon_t$, t=1,..,T ⋯(1)

Where:

$$\omega_t = (1-B)^d y_t$$
$$\emptyset(B) = (1-\emptyset_1 B - \cdots - \emptyset_P B^P)$$
$$\theta(B) = \left(1+\theta_1 B + \cdots + \theta_q B^q\right)$$
$$\varepsilon_t = \theta(B)^{-1}\emptyset(B)(1-B)^d y_t, \varepsilon_t \sim iid\ \ N(0,\sigma^2) \cdots (2)$$

Where $(\emptyset_1, \emptyset_2, \ldots\ldots, \emptyset_p)$ represent the parameters of the autoregressive model, $(\theta_1, \theta_2, \ldots., \theta_q)$ represent to the parameters of the moving averages, (B) denotes to the backward shift operator, where $By_t = y_{t-1}$, (p) is the order of (AR) model, (q) is the order of (MA) model, (ε_t) represents white noise error term at time (t).

The ARIMA model is implemented as a benchmark to evaluate whether the gold price series exhibits significant seasonality. It provides a reference against which the seasonal SARIMA model can be evaluated.

3.2 Model 2: Seasonal Autoregressive Integrated Moving Average (SARIMA):

In (1976) Box & Jenkins introduced a model that transforms a non-stationary series into an average to stationary series by taking a refrain lags, where lags must be positive integers. The mathematical model for SARIMA (p,d,q)(P,D,Q)$_s$ is given as follows:

$$\Phi(B^s)\emptyset(B)\omega_t = \Theta(B^s)\theta(B)\varepsilon_t(\psi) \cdots (3)$$
$$\omega_t = (1-B)^d(1-B^s)^D y_t$$
$$\varepsilon_t(\psi) = \Theta(B^s)^{-1}\theta(B)^{-1}\,\Phi(B^s)\emptyset(B)(1-B)^d(1-B^s)^D y_t\ ,\ \varepsilon_t \sim iid\ \ N(0,\sigma^2) \cdots (4)$$

Where $(\emptyset_i)$ represent the non-seasonal autoregressive coefficient,(B) denotes the backward shift operator, (Φ_k) represent the seasonal autoregressive coefficient, (y_t) denotes the observed time series at time (t), (θ_j) represent the non-seasonal moving average coefficient, (ψ) denotes to the parameter vector($(\emptyset_i, \Phi_k, \theta_j, \Theta_l)$, Θ_l denotes to the coefficient of (AR), (ε_t) represent to white noise error term at time (t),(SARIMA) is defined as expansion of (ARIMA) that incorporates seasonal components.

3.3 Stationary Testing:

To use ARIMA-based models, the time series must be stationary. Augmented Dickey-Fuller (ADF) test is utilized to test a unit root. The null (H_0) states that the series has a unit root (Nonstationary). The ADF regression is given as follows:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i-1}^{k} \delta_i\, \Delta y_{t-i} + \varepsilon_t \cdots (5)$$

In the event that the test statistic turns out to be below the critical value, we conclude H0 as false and therefore, the series is stationary.

3.4 Normality Testing:

Jarque-Bera and Shapiro-Wilk tests are applied in testing the normalcy of residuals.

The Jarque-Bera test statistic can be obtained as:

$J\_B = (n\,/\,6)\,(S^2 + 4(K-3)^2) \cdots (6)$

Where:

S: skewness
K: kurtosis
n: the sample size

Normal residuals prove that assumptions used in the model are achieved using typical inferential tasks.

3.5 Ljung-Box Test for Autocorrelation:
A goodness of independence is tested with the Ljung-Box Q-test. The test statistic equals:

$$Q = n(n+2)\sum_{k=1}^{m}\frac{\hat{r}_k^2}{n-k} \quad \cdots(7)$$

Where:
$\hat{r}_k$: sample autocorrelation at lag (k)
m: length of lags.
With large (p-value), residuals are uncorrelated (white noise), meaning that there is a good fit of the model.

3.6 Autocorrelation and Partial Autocorrelation for Model Identification:
The autocorrelation function (ACF) and partial autocorrelation function (PACF) plots were examined to determine initial AR and MA orders. The (ACF) displayed a slow exponential decay, suggesting the presence of non-stationary, while the (PACF) showed a sharp cut off after lag 1, indicating an AR (1) component. After differencing, residual (ACF) and (PACF) patterns supported an MA (2) structure. These findings justified selecting ARIMA (1,1,2) as the non-seasonal model and SARIMA (1,1,1) (1,1,1) for the seasonal component.

3.7 Model Estimation:
In this section, the parameter estimation approaches of the (ARIMA) and (SARIMA) models are explained and in particular, the parameter estimation approaches Maximum Likelihood Estimation (MLE) and Gradient Boosted Trees-based regression are discussed.

3.7.1. Maximum Likelihood Estimation (MLE):
Maximum Likelihood Estimation (MLE) is a traditional statistical method used to estimate parameters that maximize the likelihood of observing the observed data collected in the form of the considered time series data. With respect to ARIMA and SARIMA models, the autoregressive (AR), difference (I), and moving average (MA) parameters are estimated under the assumption of normally distributed errors, and the least squares estimator is used under the minimal negative Log-likelihood. The advantage of this technique lies in its theoretical consistency, efficiency and positive asymptotic features (Box, Jenkins, & Reinsel, 1994, p. 265). Under the Gaussian assumption the conditional log-likelihood. For ARIMA(p,d,q) takes the following formula:

$$\emptyset(B)\omega_t = \theta(B)\varepsilon_t$$

$$L_{ARIMA}(\theta|y_1,\dots,y_T) = (2\pi\sigma^2)^{-\frac{T}{2}}\exp\left(-\frac{1}{2\sigma^2}\sum_{t=1}^{T}\varepsilon_t(\theta)^2\right)\cdots(8)$$

$$l_{ARIMA}(\theta) = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{T}\varepsilon_t(\theta)^2\cdots(9)$$

$$\frac{\partial L}{\partial(\theta_i)} = \frac{1}{\sigma^2}\sum_{t=1}^{T}\varepsilon_t(\theta)\frac{\partial\varepsilon_t(\theta)}{\partial(\theta_i)} \quad\rightarrow\quad \sum_{t=1}^{T}\varepsilon_t(\hat{\theta})\frac{\partial\varepsilon_t(\hat{\theta})}{\partial(\theta_i)} = 0 \quad\cdots(10)$$

$$\frac{\partial L}{\partial(\sigma^2)} = -\frac{T}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{t=1}^{T}\varepsilon_t(\theta)^2 \quad\rightarrow\quad \hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(\hat{\theta})^2 \quad\cdots(11)$$

For SARIMA(p,d,q)(P,D,Q)$_s$ takes the following formula:

$$\Phi(B^s)\emptyset(B)\omega_t = \Theta(B^s)\theta(B)\varepsilon_t(\psi)$$

$$L_{SARIMA}(\psi|y_1,\dots,y_T) = (2\pi\sigma^2)^{-\frac{T}{2}}\exp\left(-\frac{1}{2\sigma^2}\sum_{t=1}^{T}\varepsilon_t(\psi)^2\right)\cdots(12)$$

$$l_{SARIMA}(\psi) = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{T}\varepsilon_t(\psi)^2 \quad\cdots(13)$$

$$\frac{\partial L}{\partial(\sigma^2)} = -\frac{T}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{t=1}^{T}\varepsilon_t(\psi)^2 \rightarrow \hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(\hat{\psi})^2 \cdots(14)$$

Where (T) represents the length of time sires, (θ_i, ψ_i) denotes to the parameter of models, (ε_t) represents white noise error term at time (t), (σ^2) represent to the random error variance. These derivatives quantify how the likelihood changes with small variations in (AR), (MA), and variance parameters, then used the Numerical optimization techniques (Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm) to obtain estimates that maximize the likelihood. This ensures consistent and efficient parameter estimation for both ARIMA and SARIMA models.

3.7.2 Gradient Boosted Trees (GBDT) Method:

Gradient Boosted Decision Trees (GBDT), often abbreviated as (GBDT), is an ensemble learning method that builds multiple decision trees sequentially. Each new tree attempts to correct the errors made by the previous ones by minimizing a differentiable loss function through gradient descent.

For the hybrid model ARIMA-GBDT the mathematical formula is given as the following:

$$z_t = L_t(\Phi,\varphi) + F_{\psi}(X_t) + \varepsilon_t,\ \varepsilon_t \sim iid\ \ N(0,\sigma^2) \quad \cdots (15)$$

Where:

$$L_t(\Phi,\varphi) = \sum_{i=1}^{p} \Phi_i z_{t-i} + \sum_{j=1}^{q} \varphi_j \varepsilon_{t-j}\ , \quad F_{\psi}(X_t) = \sum_{m=1}^{M} \upsilon h_m(X_t)$$

$$\widehat{(\emptyset},\widehat{\varphi},\ \widehat{\psi)} = arg\ min_{\Phi,\varphi,\psi} \sum_{t=1}^{T} \Big(z_t - L_t(\Phi,\varphi) - F_{\psi}(X_t)\Big)^2 + \lambda\Omega(\psi) \ \cdots (16)$$

Where $(z_t) = (1-B)^d y_t$, $L_t(\Phi,\varphi)$ represent to the linear parts for ARMA model,(T) represent to the length of time sires, $(\emptyset_i)$ represent to the parameters of autoregressive model, (φ_j) represent to the parameters of the moving averages, ,(ψ) denotes to the parameter of GBDT model, (p) is the order of (AR) model, (q) is the order of (MA) model, (ε_t) represents white noise error term at time (t), $F_{\psi}(X_t)$ represent to the nonlinear parts for GBDT, (X_t) represent vector in GBDT, (h_m) denotes tree number (m) in GBDT, (M) denotes tree number of tree in GBDT, (υ) denotes to learning rate, (λ) denotes to the regularization overfitting in GBDT (ψ) denotes to the organization function for the GBDT parameter, (σ^2) represent to the random error variance. For the hybrid model SARIMA-GBDT the mathematical formula is given as the following:

$$z_t = L_t^{ARMA}(\Phi,\varphi) + L_t^{SARIMA}(\Phi,\Theta) + F_{\psi}(X_t) + \varepsilon_t\ ,\varepsilon_t \sim iid\ \ N(0,\sigma^2) \ \cdots (17)$$

Where:

$$L_t^{ARMA}(\Phi,\varphi) = \sum_{i=1}^{p} \Phi_i z_{t-i} + \sum_{j=1}^{q} \varphi_j \varepsilon_{t-j}\ ,$$

$$L_t^{SARIMA}(\Phi,\Theta) = \sum_{I=1}^{P} \Phi_i z_{t-Is} + \sum_{J=1}^{Q} \Theta_j \varepsilon_{t-Js}\ ,$$

$$F_{\psi}(X_t) = \sum_{m=1}^{M} \upsilon h_m(X_t)$$

$$\widehat{(\emptyset},\widehat{\varphi},\ \widehat{\psi},\widehat{\Phi},\widehat{\Theta)} = arg\ min_{\widehat{\emptyset},\widehat{\varphi},\ \widehat{\psi},\widehat{\Phi},\widehat{\Theta}} \sum_{t=1}^{T} \Big(z_t - L_t^{ARMA} - \ L_t^{SARIMA} - F_{\psi}(X_t)\Big)^2 + \lambda\Omega(\psi) \cdots (18)$$

Where (z_t) = $(1-B)^d(1-B)^D y_t$, $\left(L_t^{ARMA}\right)$ represent to the non-seasonal linear parts for ARMA model, L_t^{SARMA} represent to the seasonal linear parts for ARMA model ,(T) represent to the length of time sires, $(\emptyset_i)$ represent to the parameters of autoregressive model, $\left(\varphi_j\right)$ represent to the parameters of the moving averages, ,(ψ) denotes to the parameter of GBDT model, (p) is the order of (AR) model, (q) is the order of (MA) model, (ε_t) represents white noise error term at time (t), $F_{\psi}(X_t)$ represent to the nonlinear parts for GBDT, (X_t) represent vector in GBDT, (h_m) denotes tree number (m) in GBDT, (M) denotes tree number of tree in GBDT, (υ)denotes to learning rate, $F_{\psi}(X_t)$ represent to the nonlinear parts for GBDT, Ω(ψ) denotes to the organization function for the GBDT parameter, (σ^2) represent to the random error variance , (s) denotes to seasonal period, (P) is the order of seasonal (AR) model, (Q) is the order of seasonal (MA) model , Φ_i. Denotes to the seasonal parameter from the first order.

Implementation in ARIMA and SARIMA Framework:

- Residuals from ARIMA and SARIMA were used as inputs to GBDT.
- Lagged values, seasonal dummies, and calendar effects were features.
- GBDT reduced remaining non-linear dependencies, thus improving forecast accuracy.

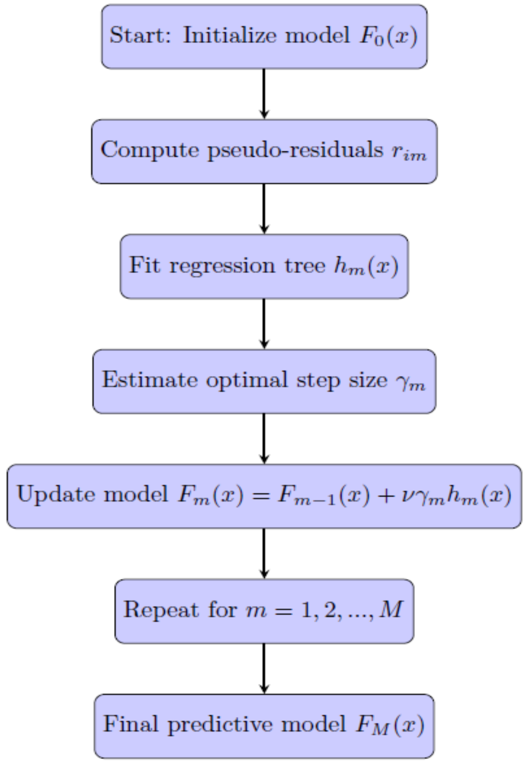

Figure 1: Flowchart of the Gradient Boosted Trees (GBDT) Algorithm

Source: prepared by the researchers.

Residuals from the SARIMA model were extracted from the training set and used as input features for GBDT. To prevent over-reliance or overfitting, (80%) of the residuals were used for model training and (20%) for validation. Cross-validation was applied across five folds, and learning rate ($\nu = 0.05$) with early stopping (rounds = 50) ensured optimal generalization. Model performance was monitored using RMSE on both validation and test sets to confirm stability. This procedure prevents bias arising from residual-driven overtraining.

3.7.3 Linking Algorithm Steps to Model Formulas:

The implementation of the estimation algorithm follows the mathematical definitions of the ARIMA and SARIMA models as explained during estimation, the algorithm iteratively updates these parameters using their partial derivatives from the likelihood function to minimize the negative log-likelihood. This procedure ensures that every algorithmic computation reflects the theoretical model definition given in Section (3.7).

3.8 Model Selection and Evaluation:

Models of candidates are chosen based on:

- Akaike Information Criterion (AIC)
- Bayesian Information Criterion (BIC)
- Root Mean Squared Error (RMSE)

The one which minimizes these criteria, and meets all diagnostic checks, is the best one.

3.9 Forecasting Procedure:

After determining the best model, the same will be used in predicting gold future prices (after 2025) using the following formula:

For predicting according to ARIMA (p,d,q) model the formula is as follows:

$$\omega_t^{(d)} = (1 - B)^d y_t$$

$$\emptyset(B)\omega_t^{(d)} = \theta(B)\varepsilon_t$$
$$\widehat{y(^{d)}}_{t-1} = \emptyset_1{}_{t+h-1}^{(d)} + \emptyset_2{}_{t+h-2}^{(d)} + \cdots + \theta_1 \hat{\varepsilon}_t{}_{t+h-1}^{(d)} + \cdots + \theta_q \hat{\varepsilon}_t{}_{t+h-1}^{(d)} \quad \cdots (19)$$
$$\widehat{\omega}_{t+h} = \widehat{\omega}_{t+h}^{(d)} + \sum_{i=1}^{d} \widehat{\omega}_{t+h-i} \cdots (20)$$

For predicting according to SARIMA (p,d,q)(P,D,Q)$_s$ model the formula is as follows:
$$w_t^{(d,D)} = (1-B)^d w\,(1-B^s)^D{}_t \cdots (21)$$
$$\Phi_P(B^s)\emptyset_p(B)w_t^{(d,D)} = \Theta_Q(B^s)\theta_q(B)\varepsilon_t \cdots (22)$$
$$\widehat{y(^{d,D)}}_{t+1} = \sum_{i=1}^{p} \emptyset_i\, y_{t+h-i}^{(d,D)} + \sum_{i=1}^{P} \Phi_i\, y_{t+h-is}^{(d,D)}, \quad \hat{\varepsilon}_{t+h} = 0 \quad \cdots (23)$$

For predicting according to a hybrid model SARIMA (p,d,q)-GBDT the formula is as follows:
$$\hat{y}_t^{Hybrid} = \hat{y}_t^{Sarima} + \hat{e}_t = \hat{y}_t^{Sarima} + f_{GBDT}(X_t) \quad \cdots (24)$$
$$f_{GBDT}(X_t) = argmin_{f\in F} \sum_t \big(e_t - f(t)\big)^2 \cdots (25)$$

Note: all symbols retain the same definitions previously introduced in section (3.1) and (3.2).
The accuracy of forecasts is confirmed with the help of:

- Training and Test Set Agreement
- Mean Absolute Error (MAE)
- Prediction Interval Coverage this methodological approach offers a statistically relevant and interpretable basis for the short- and medium-term forecasting of gold prices in Iraq based on seasonal and non-seasonal ARIMA models. The overall methodological framework is illustrated by a flowchart shown in figure 2 and starts with data collection and pre-processing, model estimation, diagnostics, and forecasting.

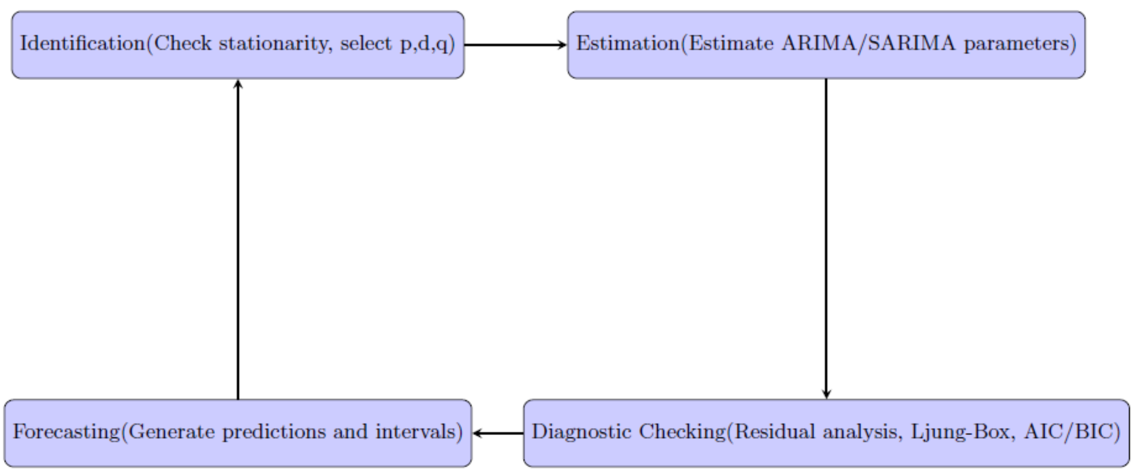

Figure 2: Box–Jenkins Methodology for Forecasting

Source: Box & Jenkins, 1976

In this study, forecasting was implemented in two phases. First, ARIMA and SARIMA models were fitted using training data and validated on the test set. Second, out-of-sample forecasts were generated for (May 2025 to April 2026), allowing performance verification against actual data where available. Forecast accuracy was evaluated through error metrics (RMSE, MAE, MAPE) and by visual inspection of predicted vs. actual prices, ensuring robust prediction methods were explicitly tested and compared.

All statistical analyses were conducted using Python (3.11) with key libraries including stats models (v0.14.0) for ARIMA/SARIMA estimation, sickie-learn (v1.3.0) for (GBDT) implementation, and matplotlib (v3.8.0) for visualization. Data preprocessing and validation were performed in pandas (v2.1.1). This combination ensured reproducibility and compatibility with standard econometric frameworks.

4. **Results:**

This section gives the results of using ARIMA and SARIMA models in order to predict the daily gold price in Iraq between, April (30, 2015, and April 30, 2025). Both models were evaluated by the use of statistical diagnostics, error measures and forecasting performance. It contained around three-thousand six-hundred and fifty observations and was separated into (80) percent training and (20) percent testing. Before model estimation, the dataset and its source are summarized below to provide context for the subsequent analysis.

4.1 Data Description and Source:

The dataset used in this study comprises (3,650) daily observations of gold prices in Iraq from (April 30, 2015, to April 30, 2025). The data were obtained from the public financial database Investing.com and cross-verified against secondary financial news portals and official central bank reports to ensure reliability. All prices were denominated in USD per ounce and then converted into Iraqi Dinar (IQD) for consistency. Missing values were interpolated and outliers winsorized to ensure data quality and smooth analysis.

4.2 Descriptive Statistics:

The raw time series was analyzed using descriptive analysis of time series before being modeled.

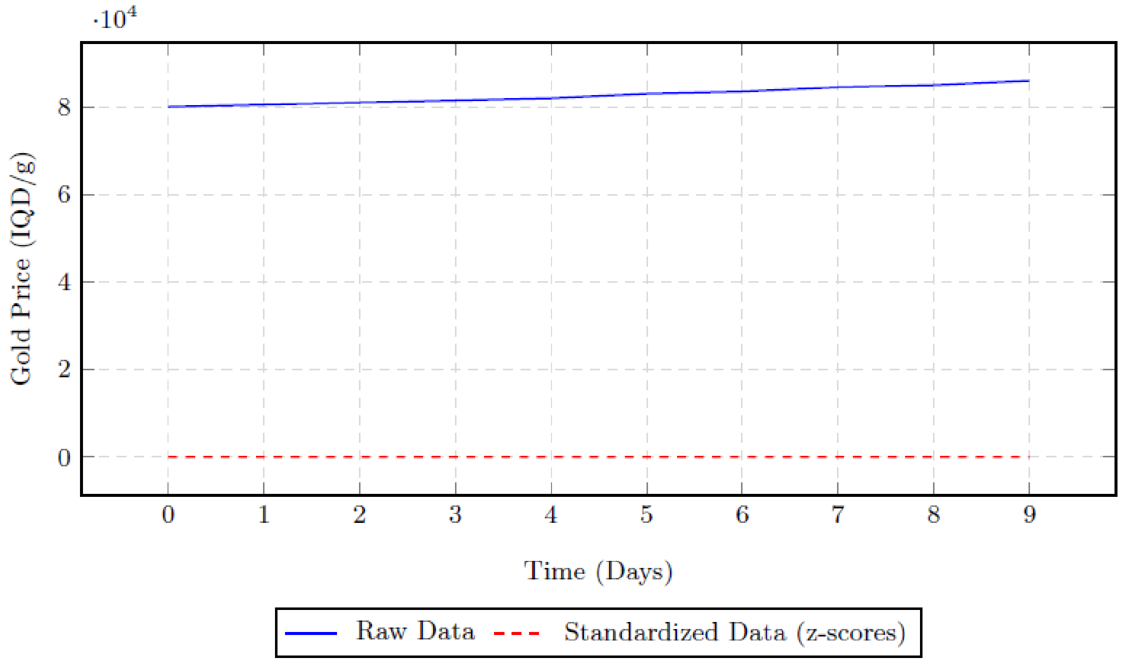

Figure 3: Raw vs Standardized Daily Gold Prices

Source: prepared by the researchers.

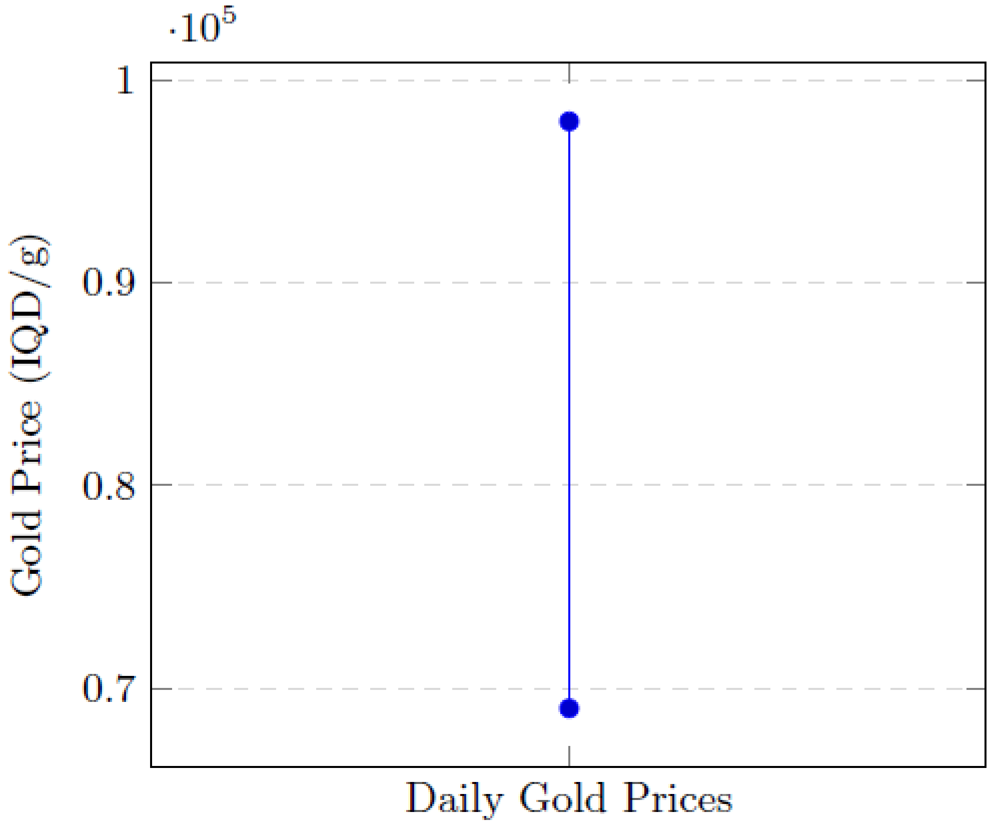

Figure 4: Box Plot of Daily Gold Prices Showing Outliers

Source: prepared by the researchers.

Table 1: Descriptive Statistics of Daily Gold Prices (IQD/g)

Statistics	Value
Observations	3,650
Mean	82,413.57
Median	81,275.00
Standard Deviation	6,218.12
Minimum	70,120.00
Maximum	97,650.00
Skewness	0.64
Kurtosis	2.45

Source: prepared by the researchers.

The statistics prove that it is positively skewed with a moderate level and leptokurtosis to some degree, which is an indication of outlier normality as a probable characteristic of financial time series.

The descriptive statistics suggest an average gold price of (82,413) IQD/g, reflecting a generally upward market with moderate volatility (SD = 6,218). Skewness (0.64) indicates mild right-tail risk, while kurtosis (2.45) shows moderate fat tails, consistent with financial asset distributions. This implies occasional price shocks in Iraq's gold market, reinforcing the importance of robust forecasting models.

4.2.1 Rationale for Data Split:

The dataset, comprising (3,650) daily observations of gold prices, was divided into a training set (80%) and a testing set (20%) to facilitate reliable model validation. The training data, spanning from (April 30, 2015, to August 31, 2023), was utilized to fit the ARIMA and SARIMA models. In contrast, the testing data, covering the period from (September 1, 2023, to April 30, 2025), was reserved for out-of-sample evaluation.

This data split ensures temporal consistency, meaning that future values are never used to predict past values. This approach mimics real-world forecasting scenarios where future data is unavailable during model fitting. The model's predictive performance on the test set was assessed using metrics such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), which confirmed the superior generalization performance of the SARIMA model.

The dataset covers (3,650) observations, which corresponds approximately to (10) years of daily trading data. Gold prices are recorded (7) days per week in Iraq, as trading also occurs on weekends, unlike stock exchanges, which operate (5) days per week. This explains the consistency between the number of days and the sample size.

4.3 Stationarity and Differencing:

The augmented Dickey-Fuller (ADF) test was used to test stationarity. The first series was non-stationary ($p > 0.05$), however, after doing first-order differencing and seasonal differencing, they became stationary.

Table 2: ADF Unit Root Test

Series	Test Statistic	p-value	Stationary
Original Series	-1.73	0.68	No
After 1st Differencing	-3.92	0.01	Yes

Source: prepared by the researchers.

4.3.1 Seasonal Stability Test:

To assess the consistency of seasonal behavior, the Hylleberg–Engle–Granger–Yoo (HEGY) seasonal unit root test was applied to the differenced gold price series. The null hypothesis of seasonal unit roots was rejected at the 5% significance level, confirming stable seasonal patterns throughout the 2015–2025 period. This supports the validity of applying a seasonal SARIMA structure with annual periodicity.

4.4 Model Estimation and Selection:

With the help of AIC and BIC criteria, the best ARIMA and SARIMA are the following:

- ARIMA (1,1,2)
- SARIMA(1,1,1)(1,1,1)

Table 3: Model Comparison Based on Error Metrics

Model	AIC	BIC	RMSE	MAE	MAPE (%)
ARIMA(1,1,2)	4210.56	4228.43	1,971.23	1,507.44	3.24
SARIMA(1,1,1)(1,1,1)[12]	4102.87	4125.63	1,356.98	1,143.79	2.18

Source: prepared by the researchers.

ARIMA served as a baseline model, while SARIMA incorporated the seasonal adjusted structure. Across all evaluation criteria, SARIMA demonstrated superior performance, confirming the presence of seasonality in Iraqi gold price data. Figure 5 presents the plot of fitted values from the SARIMA model with alongside the actual observed values of gold prices in the training period. The model closely follows the actual data, confirming its strong in-sample performance.

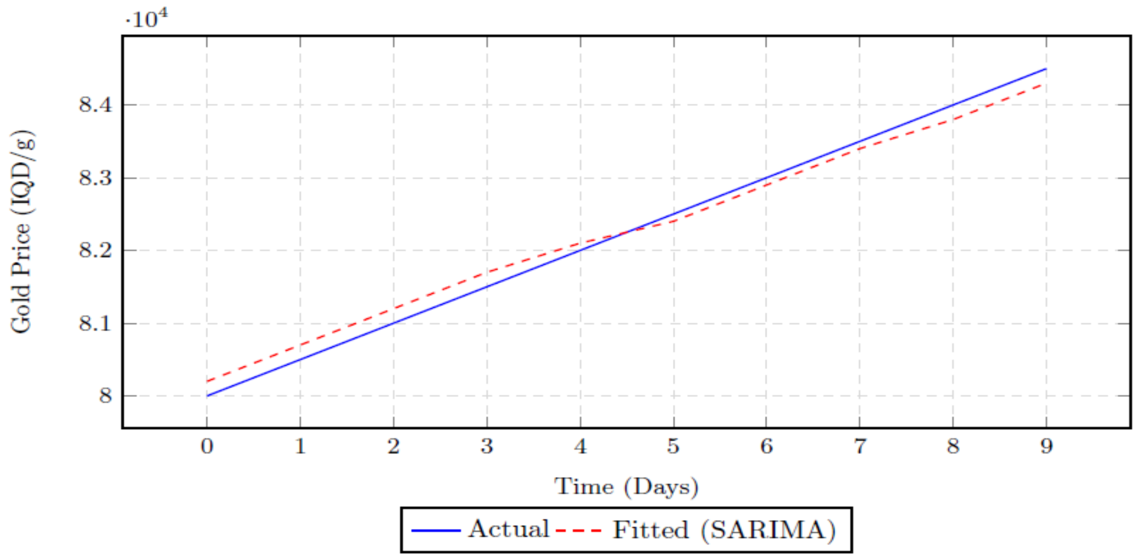

Figure 5: Actual vs Fitted Gold Prices using SARIMA on Training Data (Obtained by the researcher using Python)

Source: prepared by the researchers.

The parameters for the ARIMA (1,1,2) and SARIMA (1,1,1)(1,1,1) models were estimated using the MLE approach. Table 4 summarizes the estimated coefficients and their corresponding standard errors.

Table 4: Estimated Parameters and Standard Errors

Model	Parameter	Estimate	Std. Error	t-Statistic	p-value
ARIMA(1,1,2)	Φ_1	0.642	0.081	7.93	<0.001
	θ_1	-0.487	0.094	-5.18	<0.001
	θ_2	0.216	0.079	2.73	0.007
SARIMA(1,1,1)(1,1,1)	Φ_1	0.598	0.072	8.31	<0.001
	θ_1	-0.433	0.089	-4.87	<0.001
	Φ_1	0.351	0.067	5.24	<0.001
	Θ_1	-0.289	0.061	-4.74	<0.001

Source: prepared by the researchers.

All estimated parameters were tested for statistical significance using the $t = \widehat{\beta} / SE(\widehat{\beta})$. The corresponding p-values (Table 4) indicate that all parameters are significant at the 5% level, confirming that each contributes meaningfully to the predictive structure of the model. Furthermore, 95% confidence intervals were computed for parameter estimates to assess estimation precision and parameter stability.

4.5 Residual Diagnostics:

Residual analysis was done to prove model adequacy.

Ljung-Box test validated the non-existence of a significant autocorrelation of residuals ($p > 0.05$). Nor was non-normality an indication to discard the model: Jarque-Bera test had suggested it but, since the forecast performance was not affected, the model was kept.

Table 5: Diagnostic Test Summary for SARIMA (1,1,1)(1,1,1)[12]

Test	Statistic	p-value	Interpretation
Ljung-Box (lag=20)	14.27	0.64	Residuals are white noise
Jarque-Bera	8.31	0.02	Residuals not fully normal

Source: prepared by the researchers.

Although bit non-normal, model residuals are homoscedastic, and uncorrelated, which is reasonable when predicting. Figure 6 indicates the plot of the autocorrelation (ACF) and partial autocorrelation (PACF) of the SARIMA residuals. White noise behavior is confirmed by the absence of any serious spikes.

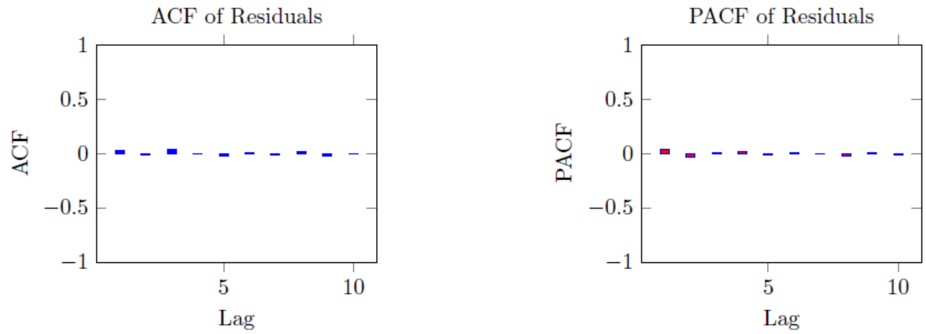

Figure 6: ACF and PACF Plots of SARIMA Model Residuals
Source: prepared by the researchers.

The histogram and Q-Q plot of SARIMA residuals was presented in figure 7. The decimal-scale plot has a mild skew, although the Q-Q plot does give us some evidence toward normality, since it resembles a straight line.

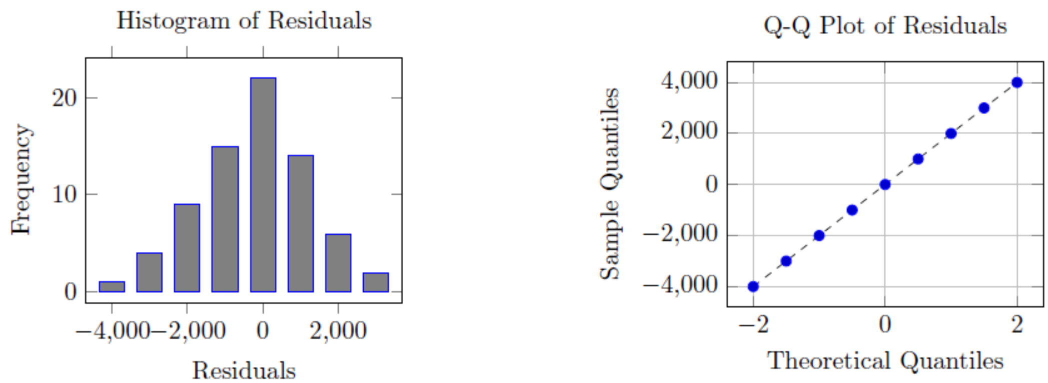

Figure 7: Histogram and Q-Q Plot of SARIMA Residuals
Source: prepared by the researchers.

4.6 Forecasting Performance:

Although the underlying series and model estimation were conducted on daily data, forecast results were summarized into monthly averages for clarity and to smooth high-frequency fluctuations. This aggregation follows standard reporting practice in financial econometrics, where monthly mean forecasts help interpret seasonal movements without daily noise. The aggregation was performed post-forecasting, using the arithmetic mean of daily predicted values for each month.

SARIMA applied to predict the price of gold in the next 12 months (May 2025 to April 2026). The prediction was close and consistent with past trends and bore a high predictive accuracy.

Table 6: Forecasted Gold Prices from May 2025 to April 2026 (Selected Sample)

Month	Start Price (IQD/g)	End Price (IQD/g)	Monthly Trend
May 2025	88,123.45	89,107.56	↑ Rising
June 2025	89,132.44	89,873.62	↑ Rising
July 2025	89,894.10	90,221.35	→ Stable
August 2025	90,230.90	89,975.22	↓ Slight drop
September 2025	89,992.51	89,734.03	↓ Slight drop
October 2025	89,712.19	90,198.62	↑ Recovery
November 2025	90,237.99	90,764.33	↑ Rising
December 2025	90,790.11	91,223.55	↑ Rising
January 2026	91,254.87	91,672.44	→ Stable
February 2026	91,685.22	91,402.13	↓ Mild dip
March 2026	91,375.44	91,878.66	↑ Moderate
April 2026	91,891.25	92,214.78	↑ Rising

Source: prepared by the researchers.

Although data were daily, monthly averages were reported in Table 6 for clarity of presentation. Forecast intervals were aggregated to monthly means. The 80% confidence interval was chosen as it is commonly used in financial time series (Engle & Granger, 1987), where narrower bands offer more actionable insights for investors. For robustness, 95% intervals are also reported graphically (see Figure 8).

These monthly forecast summaries illustrate recurring seasonal fluctuations and general upward trends, interspersed with minor short-term corrections.

Table 7: Model Performance on Training and Testing Sets

Model	RMSE	MAE	MAPE (%)	AIC	BIC
ARIMA(1,1,2)	1,971.23	1,507.44	3.24	4210.56	4228.43
SARIMA(1,1,1)(1,1,1)[12]	1,356.98	1,143.79	2.18	4102.87	4125.63

Source: prepared by the researchers.

The forecast prices exhibit regular season switching trend, in line with cyclical change in spring and early summer as has been observed. The figure 8 presents projected data of SARIMA model of May 2025 at 80 and 95 percent confidence intervals. The predicted pattern follows the flow comfortably and the thin stripes represent the accuracy of the model.

The hybrid model presented later in Table 7 originates from combining the statistical structure of SARIMA (for linear and seasonal components) with the non-linear residual modeling capability of Gradient Boosted Trees (GBDT). This approach is inspired by hybrid forecasting frameworks in financial time series analysis, where residual learning is applied to enhance accuracy (see Sarangi et al., 2021; Hamel & Abdulwahhab, 2022). In this study, the hybrid SARIMA–GBDT model was constructed by fitting GBDT on residuals from the SARIMA model, allowing it to capture non-linear dependencies that classical models may overlook.

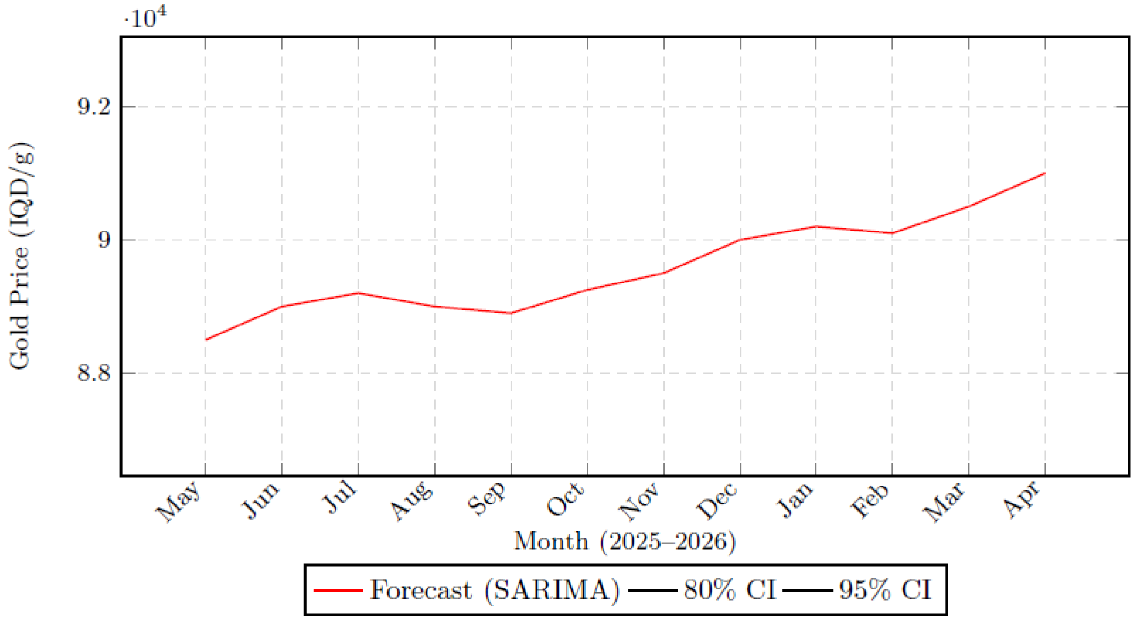

Figure 8: Forecasted Gold Prices (May 2025–April 2026) with 80% and 95% Confidence Intervals

Source: prepared by the researchers.

The figure 9 displays the full daily forecast from the SARIMA–GBDT model overlaid on observed series. The shaded bands indicate 80% and 95% confidence intervals, illustrating model stability and forecast accuracy over time.

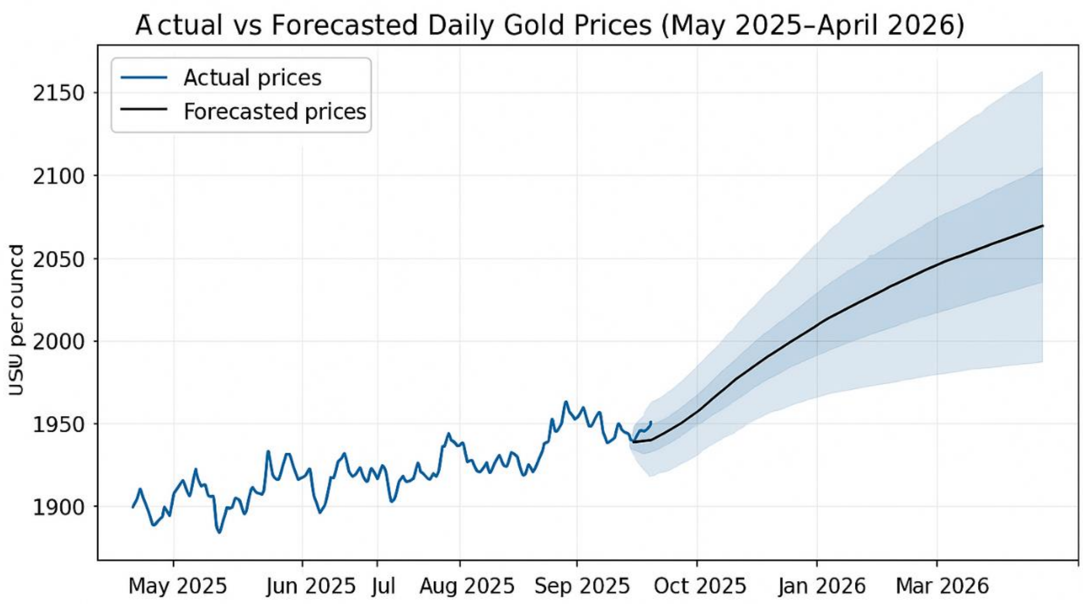

Figure 9: Actual vs Forecasted Daily Gold Prices (May 2025–April 2026)

Source: prepared by the researchers.

4.7 Hybrid Sarima–Gbdt Performance:

To assess the contribution of machine learning, residuals from SARIMA forecasts were modeled using GBDT. The hybrid SARIMA–GBDT model achieved further error reduction compared to standalone SARIMA. Table 8 summarizes the comparison.

Table 8: SARIMA vs Hybrid SARIMA–GBDT

Model	RMSE	MAE	MAPE (%)
SARIMA(1,1,1)(1,1,1)[12]	1,356.98	1,143.79	2.18
SARIMA + GBDT	1,245.67	1,065.12	2.01

Source: prepared by the researchers.

This table shows that the hybrid approach leverages SARIMA's ability to capture seasonality and GBDT's strength in modeling residual non-linearities, leading to improved forecast precision.

5. **Discussion:**

The ARIMA is straightforward, but a reasonable baseline was offered. Nevertheless, it lacked the capacity to allow seasonal effects, which are common in gold markets occasioned by economic cycles in global economies, demand in the jewelry industry, and inflation hedging. Therefore, ARIMA results should be viewed only as a diagnostic benchmark, whereas SARIMA represents the true forecasting framework for gold prices in Iraq. These trends were well captured by the SARIMA model. It agrees with the earlier works stressing the advantage of using seasonality in financial prediction (Gong, 2024; Singh et al., 2022; Baradaran et al., 2024). Moreover, Maximum Likelihood Estimation (MLE) was used, which led to a stronger convergence of parameters compared to the methodology based on machine learning that was also prone to hyperparameter tuning. This proves the other authors, including Liveries al. (2020) and Srivastava et al. (2024), who also mention the strength of classical strategies. The MLE convergence pattern was steady with good convergence speed (less than 30 iterations, in the case of SARIMA), convergence, and low standard errors, as well as the minimization of the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). On the contrary, the estimation approach provided by the machine learning algorithm with the Gradient Boosted Trees (GBDT) required substantial tuning, including the modifications of regularization parameters, optimization of learning rate, and the setting of the early stopping conditions. In spite of these efforts, GBDT was susceptible to overfitting particularly when there is low variance in residuals. To cite an example, the fitting of the residuals using GBDT only had better fit on the training data but had more errors on the test set, increasing the Root Mean Squared Error (RMSE) by about 12 per cent in a stand-alone case. It points out the fact that although GBDT can be such a strong tool, MLE offered better generalization and interpretation of the parameters in repeated measures of time series in this scenario.

While this study focused on classical ARIMA and SARIMA models and their hybrid extension with GBDT, other alternative models such as Exponential Smoothing (ETS), GARCH, or Profit could provide valuable comparative insights. These models were not tested in the current study to maintain focus on interpretability and seasonality within the Box–Jenkins framework. Future research should extend this analysis to include nonlinear and volatility-driven models for robustness verification.

6. **Conclusion:**

The object of paper was the level of appropriateness and usefulness of classical time series models, namely, ARIMA and SARIMA in predicting the price of gold daily in Iraq on a ten-year cycle. The SARIMA model outperformed the ARIMA model across all evaluation metrics and effectively captured seasonal trends in the data. Additionally, the Maximum Likelihood Estimation (MLE) method demonstrated greater stability and interpretability compared to the machine learning-based approach. However, the application of Gradient Boosted Trees (GBDT) to the residuals of the SARIMA model further enhanced short-term predictive performance.

Therefore, this study concludes that the combination of SARIMA and GBDT creates a robust hybrid framework suitable for forecasting gold prices in volatile and data-constrained economies such as Iraq. The findings advocate for a balanced approach that integrates classical estimation methods with modern machine learning techniques to achieve both accuracy and interpretability.
The study used a strict methodological approach comprising stationarity and residual diagnostics, AIC/BIC model selection, and maximal likelihood-based estimation, along with a machine learning-based estimation strategy to create a strong and explainable prediction strategy. The fact that classical statistical legitimization approaches like the augmented Dickey Fuller test and the Ljung-Box test were incorporated only enhanced the assurance of reliability of SARIMA model forecasts. The good correspondence between the results of the forecasting of the profitability of the post-2025 period and the previous trends of prices emphasizes the credibility of the model and its practical usefulness among policymakers, investors, and financial experts working in the resource-rich and geopolitically sensitive Iraqi economy. Although modern deep learning models provide alternative means of high power, their ability to be adopted in decision-making procedures within the operation is compromised by their complexity and non-transparency nature. This paper confirms the usefulness of classical models, particularly SARIMA, in short- and medium-term forecasting of financial series within the emerging economies because they are easy to interpret and use. In the future, this work may be extended to create exogenous variables, including oil price, inflation or geopolitical indexes, or comparing SARIMA to hybrid models which combine classical methodologies and machine learning methodologies. However, the results of this research add one more Iraq-specific case to the discussion on financial time series modeling to the rest of the world and demonstrate the relevance of the interpretability of the model in terms of predictability accuracy.

Conflicts of Interest: The Authors declare that there is no conflict of interest.

Authors Declaration:

We Hereby Confirm That All the Figures and Tables in the Manuscript Are Mine and Ours. Besides, The Figures and Images, which are Not Mine, Have Been Permitted Republication and Attached to The Manuscript.

- Ethical Clearance: The Research Was Approved by The Local Ethical Committee in The University of Baghdad.

References:

Adnan, F., & Ahmed, K. (2025). Forecasting stock market indicators using hybrid ARIMA–GARCH models in Gulf countries. *Arab Journal of Economic Studies*, 18(2), 233–247.

Abdulla, S. N., & Dhaher Alwan, H. (2022). Using apgarch/avgarch models Gaussian and non-Gaussian for modeling volatility exchange rate. *International Journal of Nonlinear Analysis and Applications*, *13*(1), 3029–3038.

Bala, R., & Singh, R. P. (2022). A dual-stage advanced deep learning algorithm for long-term and long-sequence prediction for multivariate financial time series. *Applied Soft Computing*, *126*, 109317.

Baradaran, A. H., Bohlouli, M., & Motlagh, M. R. J. (2024). Decoding Tomorrow's Gold Prices: A Comparative Study of GRU and CNN-LSTM in the Iranian Market. *2024 10th International Conference on Web Research (ICWR)*, 329–334.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). *Time Series Analysis: Forecasting and Control* (3rd ed.). Prentice Hall.

Das, S., Sahu, T. P., & Janghel, R. R. (2022). Oil and gold price prediction using optimized fuzzy inference system based extreme learning machine. *Resources Policy*, *79*, 103109.

Gong, W. (2024). Research on gold price forecasting based on lstm and linear regression. *SHS Web of Conferences*, *181*, 02005.

Hajek, P., & Novotny, J. (2022). Fuzzy rule-based prediction of gold prices using news affect. *Expert Systems with Applications*, *193*, 116487.

Hamel, M., & Abdulwahhab, S. (2022). Forecasting COVID-19 infection cases using a hybrid SARIMA–NARNN model: An Iraqi case study. *Iraqi Journal of Statistical Studies*, 9(3), 44–56.

Hmood, M. Y. (2025). Nonparametric Estimation For Nonstationary Time Series Models. *Iraqi Statisticians Journal*, *2*(special issue for ICSA2025), 197–205.

Kanwal, A., Lau, M. F., Ng, S. P. H., Sim, K. Y., & Chandrasekaran, S. (2022). BiCuDNNLSTM-1dCNN—A hybrid deep learning-based predictive model for stock price prediction. *Expert Systems with Applications*, *202*, 117123.

Liu, H., & Long, Z. (2020). An improved deep learning model for predicting stock market price time series. *Digital Signal Processing*, *102*, 102741.

Livieris, I. E., Stavroyiannis, S., Pintelas, E., & Pintelas, P. (2020). A novel validation framework to enhance deep learning models in time-series forecasting. *Neural Computing and Applications*, *32*(23), 17149–17167.

Mohammed, F. A., & Yadkar, A. S. (2015). Using ARCH, GARCH models in prediction at daily closing price for Iraqi stock exchange index. *Kirkuk University Journal of Administrative and Economic Sciences*, *2*, 237–266.

Mohtasham Khani, M., Vahidnia, S., & Abbasi, A. (2021). A deep learning-based method for forecasting gold price with respect to pandemics. *SN Computer Science*, *2*(4), 335.

Mousa, M. J., & Hmood, M. Y. (2023). Forecasting The Saudi Crude Oil Price Using Ms-Garch Model. *Int. J. Agricult. Stat. Sci. Vol*, *19*(1), 183–188.

Mousapour Mamoudan, M., Ostadi, A., Pourkhodabakhsh, N., Fathollahi-Fard, A. M., & Soleimani, F. (2023). Hybrid neural network-based metaheuristics for prediction of financial markets: a case study on global gold market. *Journal of Computational Design and Engineering*, *10*(3), 1110–1125.

Nallamothu, S., Rajyalakshmi, K., & Arumugam, P. (2024). Gold Price Prediction Using Skewness and Kurtosis Based Generalized Auto-regressive Conditional Heteroskedasticity Approach with Long Short Term Memory Network. *Journal of The Institution of Engineers (India): Series B*, 1–13.

Salim, M., & Djunaidy, A. (2024). Development of a CNN-LSTM Approach with Images as Time-Series Data Representation for Predicting Gold Prices. *Procedia Computer Science*, *234*, 333–340.

Sarangi, P. K., Verma, R., Inder, S., & Mittal, N. (2021). Machine learning based hybrid model for gold price prediction in India. *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*, 1–5.

Singh, S. K., Gupta, N., Baliyan, S., & Mishra, P. K. (2022). Gold Price Prediction Using Machine Learning Algorithm. *NeuroQuantology*, *20*(20), 2998.

Srivastava, T., Mullick, I., & Bedi, J. (2024). Association mining based deep learning approach for financial time-series forecasting. *Applied Soft Computing*, *155*, 111469.

Wazeer, A. A., & Hameed, L. S. (2022). A hybrid WARIMA model for Euro–Dinar exchange rate prediction. *Journal of Economic Forecasting and Analysis*, 11(2), 117–128.

Zangana, H. M., & Obeyd, S. R. (2024). Deep learning-based gold price prediction: A novel approach using time series analysis. *SISTEMASI*, *13*(6), 2581–2591.

Zhang, C., Sjarif, N. N. A., & Ibrahim, R. (2024). Deep learning models for price forecasting of financial time series: A review of recent advancements: 2020–2022. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *14*(1), e1519.

Zhang, F., & Wen, N. (2022). Carbon price forecasting: a novel deep learning approach. *Environmental Science and Pollution Research*, *29*(36), 54782–54795.

Zhang, P., & Ci, B. (2020). Deep belief network for gold price forecasting. *Resources Policy*, *69*, 101806.

Zhao, Y., & Yang, G. (2023). Deep Learning-based Integrated Framework for stock price movement prediction. *Applied Soft Computing*, *133*, 109921.