Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

Univariate Time Series Forecasting of Domestic Gold Prices: Empirical Evidence from SARIMA Modeling and Mann-Kendall Trend Analysis

Dr. Suru Munda[1], Sandeep Kumar Mund[2]

Abstract
Gold occupies a distinctive position in India, serving not only as a cultural emblem but also as a reliable financial asset and a barometer of economic conditions. This study analyzes monthly gold price data from January 2011 to February 2026 (182 observations) and generates forecasts up to February 2027 using a Seasonal Autoregressive Integrated Moving Average (SARIMA) model. The results underscore the influence of culturally driven demand, especially during festivals and wedding periods, which produces clear and recurring seasonal fluctuations in prices.

The time series initially exhibited non-stationarity; therefore, first-order differencing was applied to achieve stationarity and ensure appropriate model estimation. Seasonal patterns were detected at 12- and 24-month intervals, reflecting consistent annual consumption behavior. Among the models evaluated, SARIMA(0,1,0)(2,0,0)[12] provided the best fit, effectively capturing both the underlying trend and seasonal dynamics compared to non-seasonal alternatives. Diagnostic checks confirmed the model's reliability, showing no significant autocorrelation in residuals and stable variance.

The analysis also revealed a persistent upward trend in gold prices, with an average monthly increase of about INR 249 per 10 grams during the study period. This trend is consistent with broader macroeconomic influences such as currency depreciation, global economic uncertainty, and sustained domestic demand. Forecasts indicate that prices are likely to continue rising, with an estimated increase of around 14% over the projection period. However, the widening forecast intervals highlight growing uncertainty, suggesting the potential impact of unforeseen external factors not accounted for in the model.

In conclusion, the study demonstrates the importance of integrating cultural seasonality into financial time series modeling in the Indian context. The findings provide meaningful insights for policymakers, investors, and researchers interested in understanding and predicting commodity price behavior.

Keywords: Gold prices, India, SARIMA model, seasonality, forecasting.

1. Introduction

The Cultural and Economic Significance of Gold in India

In the Indian context, gold transcends its status as a mere commodity, serving as a vital thread in the nation's social and spiritual tapestry. India stands as a global titan in gold consumption, consistently claiming nearly a quarter of the world's total demand. While many international markets view gold strictly through a financial lens, for Indians, it represents a multifaceted fusion of religious devotion, social prestige, and a reliable safety net for wealth. This deep-seated value is most evident during major festivals like Diwali and Dhanteras, where purchasing gold is seen as an act of inviting divine favor. Similarly, Akshaya Tritiya serves as a key moment for acquiring gold to ensure enduring prosperity, while the expansive wedding seasons across the country cement gold jewelry as an indispensable gift and dowry component. These traditions foster a unique, rhythmic demand that is virtually unmatched globally.

Macroeconomic Impact and Trade Balance

This relentless appetite for gold necessitates massive annual imports, typically ranging between 700 and 900 tonnes. Consequently, gold occupies a dominant position in India's trade profile, trailing only crude oil in terms of import costs. This heavy reliance means that the fluctuating price of gold carries significant weight for the broader economy; it influences the trade balance, exerts pressure on the Indian Rupee's valuation, and plays a critical role in managing the country's current account deficit.

Mechanics of Gold Price Volatility

The valuation of gold in India is a tug-of-war between international shifts and local economic conditions. On a global scale, prices climb when geopolitical instability or economic recessions drive investors toward "safe-haven" assets. Furthermore, when the US Federal Reserve maintains low interest rates or the US Dollar weakens, gold typically gains momentum. Domestically, the price is heavily influenced by the USD to INR exchange rate—if the Rupee depreciates, gold becomes more costly for Indian buyers even if global rates remain flat. Taxes, import duties, and local inflation also play pivotal roles. This complex interplay drove prices from INR 19,609 in 2011 to a staggering INR 1,54,545 by early 2026, marking a 688% increase over fifteen years, with the 2020 pandemic serving as a notable catalyst for rapid price surges.

Cyclical Demand and Seasonal Trends

The Indian gold market adheres to a predictable seasonal heartbeat dictated by the cultural calendar. The primary surge occurs between October and December, fueled by the alignment of major festivals and the peak wedding season. A secondary peak typically emerges in April or May, coinciding with the auspicious window of Akshaya Tritiya. Conversely, the market often cools during the monsoon months (June to August) as agricultural cycles and a lack of major festivals dampen purchasing activity. Because these patterns recur with such regularity every twelve months, advanced statistical tools like the SARIMA model which accounts for a 12-month seasonal cycleare far more effective at forecasting Indian gold prices than standard models that ignore these cultural rhythms.

Dataset Overview

The analysis is based on monthly average prices for 10 grams of gold in India, spanning from January 2011 to February 2026. This 182-month dataset reflects a massive growth trajectory, starting at a minimum of INR 19,609 and reaching a peak of INR 1,54,545. Over this period, the mean price settled at approximately INR 40,213, providing a comprehensive look at the metal's long-term appreciation and volatility.

Research Objectives

1. Time Series Decomposition: Analyze Indian gold price patterns (2011–2026) to identify underlying trends, seasonal cycles, and irregular fluctuations.
2. Stationarity Testing: Conduct Augmented Dickey-Fuller (ADF) tests and apply differencing techniques to stabilize the series variance and mean.
3. Model Selection: Compare the predictive accuracy of ARIMA, SARIMA, and ETS frameworks to determine the superior forecasting model.
4. Diagnostic Validation: Perform residual analysis on the chosen model to ensure statistical validity and absence of autocorrelation.
5. Future Forecasting: Project gold prices from March 2026 to February 2027, incorporating 95% confidence intervals.
6. Trend Quantification: Utilize Sen's Slope and the Seasonal Mann-Kendall test to measure the magnitude and significance of long-term price movements.
7. Smoothing Analysis: Apply Rolling Averages to filter short-term volatility and highlight the macro-level price trajectory.
8. Performance Metrics: Evaluate model precision using standardized error measurements, including MSE, RMSE, MAE, and MAPE.

Literature Review

Theoretical Framework and Literature Review

Foundations of Time Series Analysis

Time series analysis is a specialized statistical discipline that examines chronological data to identify recurring patterns and project future outcomes. These patterns typically manifest as trends (long-term trajectories), seasonality (fixed periodic cycles), cyclicality (multi-year waves), and irregularities (random shocks).

The field evolved through several landmark contributions:

Yule (1927): Established autoregressive models, suggesting that current observations are intrinsically linked to their historical values.

Slutsky (1937) & Wold (1938): Demonstrated how random shocks synthesize into cycles and proved that stationary series can be modeled as a blend of past values and errors.

Box and Jenkins (1970): Developed the definitive "Box-Jenkins Methodology" used in this study. This practical framework involves a rigorous cycle of stationarity testing, model identification (via ACF/PACF), parameter estimation, and diagnostic checking.

Insights from Indian Gold Studies

Research specifically targeting the Indian market highlights the unique behavior of gold as both a cultural and financial asset:
Stationarity and Seasonality: Guha and Bandyopadhyay (2016) noted that Indian gold prices generally require first-order differencing to reach stationarity. Mishra et al. (2019) and Kabbilawsh et al. (2020) argued that SARIMA models consistently outperform standard ARIMA by accounting for India's distinctive festival-driven demand.
Economic Drivers: Jain and Biswal (2016) identified the USD/INR exchange rate as a primary catalyst for domestic price appreciation. Meanwhile, Baur and McDermott (2010) reinforced gold's "safe-haven" status, explaining the extreme volatility and price spikes observed during global crises like the 2020 pandemic.
Trend Analysis: The non-parametric methods utilized in this study, such as Sen's Slope (1968), provide a robust means of quantifying long-term growth without being skewed by outliers.

Time Series Forecasting Models (ARIMA vs. SARIMA)
Recent comparative analyses emphasize that while the **ARIMA** (Auto-Regressive Integrated Moving Average) model is a stable benchmark for short-term forecasting, **SARIMA** (Seasonal ARIMA) is superior for the Indian market due to its ability to capture recurring seasonal demand cycles (e.g., wedding seasons and festivals).

Model Performance: A 2025 study comparing these models using data from 2014–2025 found that SARIMA outperformed ARIMA across all error metrics, achieving a **6% improvement** in Mean Absolute Percentage Error (MAPE) (4.58% vs. 4.89%) (IJTAS, 2025).

Predictive Trends: Empirical analysis using the Box-Jenkins methodology (ARIMA 1,1,1) predicted a steady rise in Indian gold prices from December 2024 through April 2025, aligning with historical seasonal consumer demand (ResearchGate, 2026).

Stability: Comparative research has shown that while newer models like **LSTM** (Long Short-Term Memory) attempt to capture complex patterns, the ARIMA model often maintains higher stability and directional accuracy (55.84%) for short-term trend predictions (ResearchGate, 2026).

2. Hybrid and Machine Learning Approaches
Researchers are increasingly moving toward hybrid models to capture both linear and non-linear dynamics of the gold market.

CNN-QRLSTM: A 2026 study introduced a framework integrating Convolutional Neural Networks (CNN) with Quantile Regression LSTM. This model innovatively incorporates **online news sentiment mining** to quantify how media sentiment and geopolitical news impact price fluctuations (MDPI, 2026).

ARIMA-SVR & ARIMA-LSTM: Hybridizations such as ARIMA-Support Vector Regression (SVR) have been shown to outperform individual models by capturing the complex interplay of macroeconomic policy changes and market sentiment (ResearchGate, 2026).

3. Market Projections for 2026

As of April 2026, gold prices in India have seen significant volatility driven by geopolitical tensions in the Middle East, reaching record highs before stabilizing.

Price Targets: Major financial institutions, including Goldman Sachs and J.P. Morgan, projected a **20–30% upside** for gold in 2026, with potential peaks testing the **INR 1.8 – 2.0 lakh per 10 grams** range by the final quarter (Grip Invest, 2026).

Cultural Drivers: Demand remains resilient during "Akshaya Tritiya" and the Q4 wedding season, where single-day sales in India can exceed **INR 12,000 crore** regardless of elevated prices (Grip Invest, 2026).

Research Gap

While existing literature covers historical trends, there is a lack of integrated research that compares SARIMA, ARIMA, and ETS models alongside non-parametric trend tests (Sen's Slope/Mann-Kendall) using post-COVID data. This study addresses this gap by analyzing 182 monthly observations (2011–2026), capturing the significant price acceleration and market shifts of the most recent five-year period.

Material And Methods

The Concept of Stationarity

Stationarity serves as the bedrock of time series analysis. A process is defined as stationary when its core statistical properties specifically its **mean, variance, and autocorrelation** are invariant over time. In practical terms, a stationary series oscillates around a constant level with consistent fluctuations, devoid of long-term trends or shifts in seasonal behavior.

To illustrate, consider the trajectory of gold prices from 2011 to 2026. The average price in 2011 hovered near **INR 26,000**, whereas by 2025, it had surged to approximately **INR 1,30,000**. Because the mean is aggressively rising rather than remaining stable, the series is fundamentally **non-stationary**.

The Necessity of Transformation

Predictive frameworks like the **ARIMA** (Auto-Regressive Integrated Moving Average) family are mathematically predicated on the assumption of constant statistical properties. When a series exhibits a persistent trend or changing variability:

- **Model Instability:** The underlying equations and parameter estimates become unreliable.
- **Statistical Bias:** Just as a single average cannot accurately represent a data point in a constantly climbing sequence, a model cannot capture the true relationship between past and future values if the "baseline" is always moving.

Consequently, identifying and removing trends or non-constant variance is a mandatory prerequisite. This transformation often achieved through **differencing** converts the raw price data into a stationary format, ensuring the model's projections are statistically valid. This is precisely why stationarity testing is the critical first step in our modeling workflow.

What is Differencing? Why d = 1?

Instead of relying on absolute price levels, differencing involves calculating the period-to-period ariance. In practice, this means replacing each data point with the net change from the preceding month.

If the original series is $Y_1, Y_2, Y_3, \ldots$, then after one differencing, we get:

$$\Delta Y_t = Y_t - Y_{t-1}$$

This shift in perspective moves the focus from the **level** of the price to the **momentum** of the market. Rather than looking at a static snapshot of value, the new series ΔY_t isolates the month-to-month volatility.

Interpretation: When we difference a series that has a linear upward trend (like gold prices), the trend disappears from the differenced series. The rising staircase becomes a flat fluctuating line. The ADF test on the original log series gave $p = 0.99$, meaning the series is clearly non-stationary. After one differencing, the ADF test gave $p = 0.01$, confirming stationarity. Since ONE differencing was sufficient, we set d = 1 in the ARIMA and SARIMA models.

Why Log Transformation?

Before differencing, the gold price series is log-transformed (we compute $\log(Y_t)$ instead of using Y_t directly). This is done for two reasons:

- Variance stabilization: Raw gold prices show increasing variability over time — the fluctuations in 2025 are much larger in absolute terms than fluctuations in 2011, even if they are similar in percentage terms. Log transformation converts absolute changes to percentage changes, making the variance more stable.
- Economic interpretation: In economics and finance, percentage changes are more meaningful than absolute changes. A rise of INR 1,000 means very different things when the price is INR 20,000 versus INR 1,20,000. Working in log scale makes the analysis scale-invariant.

The Augmented Dickey-Fuller (ADF) Test

The ADF test formally tests whether a time series has a unit root — which is the statistical term for non-stationarity. The test runs a regression of the form:

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \Sigma\, \delta_i\, \Delta Y_{t-i} + \varepsilon_t$$

The key coefficient is γ (gamma). If $\gamma = 0$, the series has a unit root and is non-stationary. If $\gamma < 0$, the series is stationary. The test sets up:

- Null Hypothesis (H_0): $\gamma = 0$ — the series has a unit root — it is NON-STATIONARY
- Alternative Hypothesis (H_1): $\gamma < 0$ — the series is STATIONARY

Decision Rule and Logic

The p-value tells us the probability of observing our data if H_0 were true. A very small p-value means our data is very unlikely under H_0, so we reject H_0.

- If p-value > 0.05: We FAIL to reject H_0. The series is non-stationary. We must apply differencing and re-test.
- If p-value ≤ 0.05: We REJECT H_0. The series is stationary. We can proceed with modeling.

Interpretation: For our gold price series: the original log series gave ADF statistic = 1.2282 with p = 0.99. Since $p = 0.99 >> 0.05$, we cannot reject H_0. The series is clearly non-stationary — this makes complete sense because gold prices have been rising consistently for 15 years. After one differencing: ADF statistic = −4.1014, $p = 0.01 < 0.05$. We reject H_0. The differenced series is stationary. Therefore $d = 1$ is set in all ARIMA-based models.

ACF and PACF — The Model Identification Tools

After achieving stationarity, we need to determine how many AR and MA terms to include in the model. The ACF and PACF plots are the primary tools for this.

ACF — Autocorrelation Function

The ACF at lag k measures the correlation between a value Y_t and its value k periods earlier (Y_{t-k}). In simple terms: how strongly is this month's gold price related to the price k months ago?

$$\mathbf{ACF(k) = \rho(k) = Cov(Y_t, Y_{t-k}) \ / \ Var(Y_t)}$$

The presence of prominent spikes at intervals of 12 (12, 24, 36) in the Autocorrelation Function (ACF) plot provides clear evidence of annual seasonality.

Essentially, the model "remembers" what happened exactly one, two, or three years ago. In the context of gold, this statistical "memory" isn't a coincidence; it is driven by predictable human behavior.

Partial Autocorrelation Function (PACF)

The **Partial Autocorrelation Function (PACF)** serves as a precision tool: it calculates the "clean" relationship between the current value Y_t and a specific past value Y_{t-k}, by filtering out the "noise" or influence of the months in between.

Identifying the Seasonal AR(2) Pattern

The behavior observed in your data provides a textbook signature for model identification:

- **The Direct Link:** The significant spikes at lags 12 and 24 indicate that the gold price from exactly one year ago and two years ago both exert a **unique, independent influence** on today's price.
- **The Cut-off:** Because the spike disappears at lag 36, it tells us that the three-year-old data doesn't provide any *new* information that isn't already captured by the 12 and 24-month lags.
- **The Model Logic:** In time-series forecasting, a PACF that "cuts off" after a specific seasonal lag P identifies a **Seasonal Autoregressive (SAR)** process. Since your plot remains significant until the second seasonal lag (24 months) and then stops, it confirms a **Seasonal Order of P = 2**.

Implications for the SARIMA Model

This discovery is the final piece of the puzzle for your **SARIMA** (p,d,q)× \times $(2,D,Q)_{12}$ structure. It suggests that to accurately forecast gold prices, the model must explicitly weight the direct impacts of the market conditions from the two previous anniversary cycles.

By including these two seasonal AR terms, you are mathematically accounting for the recurring intensity of those peak festival periods without "double-counting" the influence of older data. In the PACF of the differenced log series, we see significant spikes at seasonal lags 12 and 24, but NOT at lag 36. This pattern — significant PACF at lags 12 and 24, with no further seasonal PACF spikes — is the identifying pattern for a seasonal AR model of order P = 2. This means two seasonal AR terms are needed: the price 12 months ago and the price 24 months ago both have a direct, independent influence on the current price. This is why the seasonal part of our SARIMA model has P = 2.

> **Why?** *There are no significant spikes at non-seasonal lags (1, 2, 3, ...) in either ACF or PACF. This tells us that no non-seasonal AR or MA terms are needed p = 0 and q = 0. The only structure in the data is seasonal, which makes perfect sense: Indian gold prices move primarily with festival cycles, not with individual month-to-month momentum.*

ARIMA at a Glance

The ARIMA model predicts future values by looking at three specific factors:

- AR (p): The Past. It uses previous values to predict the next one (e.g., "What happened yesterday?").
- I (d): The Trend. It subtracts the previous value from the current one to make the data stable and remove long-term "drift."
- MA (q): The Errors. It looks at how far off previous guesses were and adjusts the next prediction to fix those mistakes.

The general ARIMA equation in compact form:

$$\boldsymbol{\varphi(B)(1 - B)^d\, Y_t \;=\; \theta(B)\, \varepsilon_t}$$

where B is the backshift operator ($BY_t = Y_{t-1}$), φ(B) is the AR polynomial, θ(B) is the MA polynomial, and ε_t is white noise. For the selected model ARIMA(0,1,1), this simplifies to:

$$(1 - B)Y_t = (1 + \theta_1 B)\varepsilon_t$$

ARIMA treats a festival-driven price spike as a "mistake" to be corrected; SARIMA treats it as a "rule" to be expected. This is why SARIMA is the superior tool for capturing the unique rhythm of the Indian gold market.

SARIMA is essentially ARIMA with a memory for seasons.

It uses two sets of rules to forecast:

- Non-Seasonal (p,d,q): Captures short-term, month-to-month momentum.
- Seasonal (P,D,Q): Captures long-term patterns that repeat every s periods (e.g., s=12 for annual cycles).

The Bottom Line: While plain ARIMA only looks at the immediate past, SARIMA recognizes that this month's gold price is influenced by both last month and the same month last year. This makes it the superior tool for markets driven by recurring events like Indian festivals.

$$\Phi(B^s)\,\varphi(B)\,(1 - B^s)^D\,(1 - B)^d\,Y_t = \Theta(B^s)\,\theta(B)\,\varepsilon_t$$

The reason SARIMA outperforms plain ARIMA in India is due to its "Seasonal Memory."

- The Conflict: Indian gold demand isn't random; it's tied to the lunar and solar calendars. Festivals like Diwali and Akshaya Tritiya create massive, predictable spikes every 12 months.
- The Failure of Plain ARIMA: It only looks at the immediate past (e.g., what happened in September to predict October). It doesn't "know" that October is a festival month, so it treats the price surge as a random shock rather than a recurring rule.
- The SARIMA Solution: It explicitly links the current month to the same month from previous years (s=12). By recognizing that "it is festival season again," the model expects the spike, leading to much higher accuracy and lower forecast errors.

The **SARIMA**(0,1,0)(2,0,0)[12] model focuses entirely on the annual rhythm of the gold market rather than short-term monthly fluctuations. By applying non-seasonal differencing (d=1), the model first stabilizes the data by removing long-term upward trends. The absence of non-seasonal AR and MA terms (p=0, q=0) indicates that the model ignores immediate month-to-month momentum and error corrections. Instead, it relies on two seasonal autoregressive terms (P=2) to capture a two-year "seasonal memory," meaning today's price is directly influenced by market behavior from exactly 12 and 24 months ago. This structure, set to a 12-month cycle ([12]), perfectly targets the recurring demand spikes of the Indian festival calendar without requiring additional seasonal differencing or error smoothing.

The model equation in fully expanded form:

$$(1 - \Phi_1 B^{12} - \Phi_2 B^{24})(1 - B)Y_t = \varepsilon_t$$

After expanding and rearranging, this gives:

$$Y_t = Y_{t-1} + \Phi_1(Y_{t-12} - Y_{t-13}) + \Phi_2(Y_{t-24} - Y_{t-25}) + \varepsilon_t$$

The model translates to a straightforward prediction: today's gold price is the sum of last month's value plus a weighted "echo" of the past two years. Specifically, the coefficients **Φ_1(0.1637)** and **Φ_2(0.2168)** act as memory markers for the 12-month and 24-month cycles.

Because these values are positive, they act as **momentum indicators** confirming that the market tends to repeat its successes. In plain terms, the model "remembers" the annual festival demand; if prices surged during Diwali or Akshaya Tritiya in the previous two years, the math automatically adjusts today's forecast upward to account for that recurring cultural pattern. This allows the model to carry forward the current price while layering on a specific "seasonal bonus" based on historical trends.

The coefficients **Φ_1(0.1637)** and **Φ_2(0.2168)** quantify the strength of the "seasonal echo" in the gold market. Specifically, a ₹1 increase in the seasonally adjusted price from 12 and 24 months ago predicts a current price rise of approximately ₹0.16 and ₹0.22, respectively. Because both values are positive and statistically significant (p < 0.05$), they confirm that the "festival effect" is a real, measurable force rather than random chance. This proves that Indian gold prices possess a multi-year memory, where the momentum from past anniversary cycles—particularly the even stronger influence from two years agoconsistently pulls current prices in the same direction.

The ETS Model

ETS (Error, Trend, Seasonal) is a forecasting method that breaks a series into its core structural parts using Exponential Smoothing, where recent data is given more weight than older observations. The ETS(M,A,N) configuration specifically models the data with Multiplicative Error, meaning fluctuations grow in proportion to the price level, and an Additive Trend, which assumes a steady, linear climb over time. Because this specific model is set to No Seasonal (N), it ignores recurring annual cycles and focuses entirely on the "fading" memory of the trend and recent price levels to project the future.

$$\hat{y}_{t+h} = \ell_t + h \times b_t$$

In this equation, ℓ_t represents the **smoothed level** (the current "baseline" value) and b_t represents the **smoothed trend** (the estimated rate of growth). Because the **ETS(M,A,N)** model is configured with "No Seasonal" component, it operates with a major handicap in the Indian context: it treats the data as a simple, evolving trend without any memory of annual cycles. Consequently, it cannot anticipate the predictable price surges associated with the festival calendar, leading to higher errors and making it the weakest performer compared to models that explicitly account for seasonality.

The **Akaike Information Criterion (AIC)** is a statistical scoring system used to identify the most efficient model by balancing accuracy against simplicity. It rewards models that fit the historical data closely but imposes a mathematical "penalty" for every additional parameter added, which prevents **overfitting** a common problem where a model becomes so complex it starts "memorizing" random noise rather than identifying true patterns. In model selection, the goal is to achieve the **lowest AIC value**, as this represents the optimal "sweet spot" where the model is powerful enough to be precise but streamlined enough to remain reliable for future forecasting.

$$\mathbf{AIC = 2k - 2\ell}$$

Here k is the number of parameters estimated by the model, and ℓ is the maximum log-likelihood how well the model fits the data. A higher log-likelihood (better fit) reduces AIC. More parameters (more complexity) increases AIC.

AIC is used only within the ARIMA/SARIMA family for parameter selection. It is NOT used to compare SARIMA against ETS, because ETS uses a fundamentally different likelihood formulation comparing their AIC values would be meaningless, like comparing temperatures in Celsius and Fahrenheit without converting.

How to Measure Forecast Accuracy

All four metrics are computed on the TEST SET 37 months of actual gold prices (February 2023 to February 2026) that the model never saw during training. This gives a genuine, unbiased assessment of how well the model forecasts real future values.

MSE — Mean Squared Error

$$\mathbf{MSE = (1/n) \times \Sigma (Y_i - \hat{Y}_i)^2}$$

MSE squares each error before averaging. Squaring has two effects: it makes all errors positive (so overestimates and underestimates do not cancel each other out), and it penalizes large errors much more heavily than small ones. A model that makes one huge error is penalized more than a model that makes many small errors of the same total magnitude.

$$\mathbf{RMSE = \sqrt{[(1/n) \times \Sigma (Y_i - \hat{Y}_i)^2]}}$$

RMSE = √MSE. Taking the square root brings the error back to the same scale as the original data (INR). This makes RMSE directly interpretable: RMSE = INR 29,658 means the model's predictions were off by approximately INR 29,658 per 10 grams on average during the test period. RMSE is the most commonly used accuracy metric for comparing models.

$$\mathbf{MAE = (1/n) \times \Sigma |Y_i - \hat{Y}_i|}$$

MAE uses absolute values instead of squared values. It treats all errors equally regardless of their size, making it more robust to outlier predictions. MAE = INR 19,073 means the model was off by INR 19,073 on average. MAE < RMSE (INR 19,073 < INR 29,658) confirms that a few large errors are pulling up the RMSE likely the COVID spike months where any model would struggle.

$$\mathbf{MAPE = (1/n) \times \Sigma\, (|Y_i - \hat{Y}_i| / Y_i) \times 100}$$

MAPE expresses the error as a percentage of the actual price, making it scale-independent and intuitive. MAPE = 18.98% means the model's predictions were approximately 19% away from actual prices on average during the test period. This is acceptable for a univariate model forecasting over a 3-year test window that included dramatic price acceleration due to geopolitical events. The model cannot know in advance that gold will surge 60% in 2024–2025 no univariate model can.

Sen's Slope Measuring the Long-Term Trend

Sen's Slope is a non-parametric method for estimating the rate of trend in a time series. It is calculated as the median of all possible pairwise slopes between data points:

$$\boldsymbol{\beta = \text{Median}\ \{ (Y_j - Y_i) / (j - i) \}\quad \text{for all } i < j}$$

Result: β = INR 249.128 per month. This means that, on average, gold prices in India increased by approximately INR 249 per 10 grams every single month over the 15-year study period. Over 180 months, this implies a total trend-driven increase of approximately INR 44,843 consistent with the actual price rise from INR 19,609 to INR 86,446.

The Seasonal Mann-Kendall (SMK) : The **Seasonal Mann-Kendall** test provides a rigorous validation of the gold price trend by analyzing each of the 12 calendar months independently across the entire dataset. This "month-by-month" approach ensures that recurring seasonal spikes, like those seen during festival seasons, don't trick the model into seeing a trend where none exists; it only identifies a trend if prices are consistently rising year-over-year for each specific month. With a **Z-statistic of 14.68**—far exceeding the standard 1.96 threshold—and a **p-value near zero**, the test confirms with near-certainty that the upward movement in Indian gold is a real, systematic phenomenon rather than a product of random market noise.**Rolling Average Analysis**

The 12-month centered rolling average replaces each value with the average of the 6 months before it and the 6 months after it:

$$\mathbf{M_t = (1/12) \times \Sigma\, Y_{t+i}\quad for\ i = -5\ to\ +6}$$

The Rolling Average serves as a vital visual filter, smoothing out volatile "noise" to reveal the underlying trajectory of the market. While the 2013 price correction appears significant in raw data, the rolling average shows it as a minor, temporary deviation that failed to break the long-term momentum. This highlights the resilience of the fundamental upward trend. Furthermore, the visible steepening of the average after 2019 provides empirical evidence of a trend acceleration, directly aligning with the heightened geopolitical tensions and currency fluctuations of that period. By stripping away short-term fluctuations, the rolling average mathematically confirms that the long-term growth of gold remains the dominant narrative.

Confidence Intervals

The **95% Confidence Interval** serves as a mathematical boundary for uncertainty, representing the range where actual gold prices are expected to fall 95% of the time. Because the model calculates this on a log scale before converting back to INR, the interval grows proportionally to the forecast horizon, resulting in an exponentially widening "fan" over time.

This dramatic expansion—ranging from **±₹14,000** in the immediate term to **±₹65,000** by February 2027—is not a flaw, but a display of "statistical honesty." It accurately reflects that while we have high confidence in near-term projections, the unpredictable nature of global markets makes long-term certainty impossible. Any model claiming a narrow margin of error years into the future would be making unrealistic assumptions; here, the widening band correctly warns the user that the further out the forecast goes, the more room there is for external shocks to influence the price.

RESULT AND DISCUSSION

This chapter provides a comprehensive time series analysis of Indian gold prices over a 15-year period, covering 182 monthly observations from January 2011 to February 2026. Utilizing **R Programming** and the rigorous **Box-Jenkins methodology**, the study tracks the market's evolution from a minimum of **₹19,609** to a peak of **₹1,54,545**. With an average price of approximately **₹40,213**, the data reveals a staggering **688% growth**, underscoring gold's long-term trajectory as a high-performing asset in the Indian economy. By applying systematic statistical checks, this analysis ensures that the observed trends and forecasts are rooted in mathematical consistency rather than anecdotal fluctuations

Before any formal testing, visual inspection of the time series is the essential first step. The graph (Fig 4.1) immediately reveals the major features of the data that guide all subsequent analysis decisions.

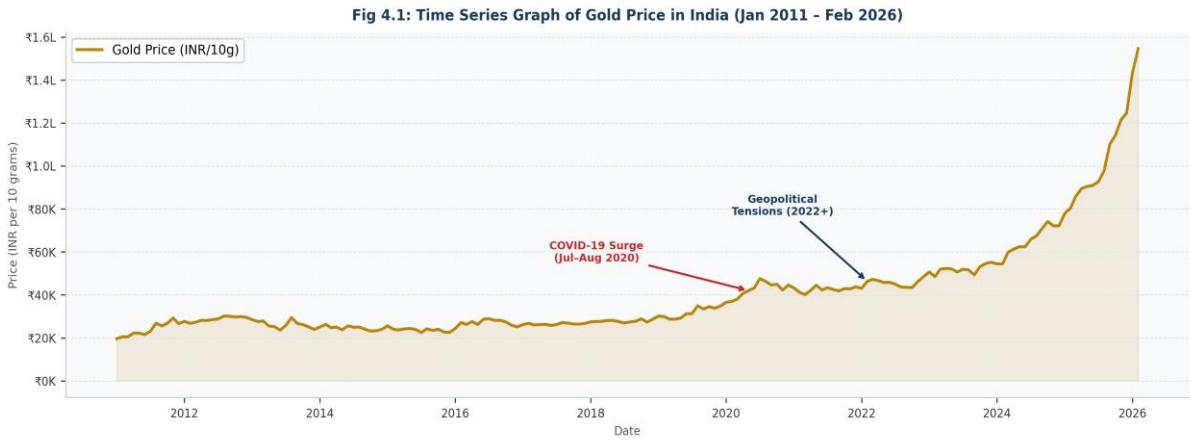

Fig 4.1: Time Series Graph of Gold Price in India (Jan 2011 – Feb 2026)

The historical gold price graph highlights four defining characteristics of the Indian market over the past 15 years. Dominating the series is a **persistent upward trend**, fueled by the depreciation of the Rupee and consistent domestic demand. Within this climb, **visible seasonality** appears as recurring annual "wobbles" tied to cultural festivals, while the **2020 spike** illustrates a massive irregular shock as the COVID-19 pandemic triggered a global flight to safety. Finally, the **accelerated growth** from 2022 onwards indicates a structural shift driven by heightened geopolitical tensions. Together, these features mathematically confirm that the series is non-stationary and requires a **SARIMA** approach to effectively capture the interplay of long-term momentum, seasonal cycles, and unpredictable market shocks.

4.2 Structural Decomposition of the Time Series

The log-transformed gold price series is decomposed into three additive components: Trend (T_t), Seasonal (S_t), and Residual (R_t), such that $\log(Y_t) = T_t + S_t + R_t$.

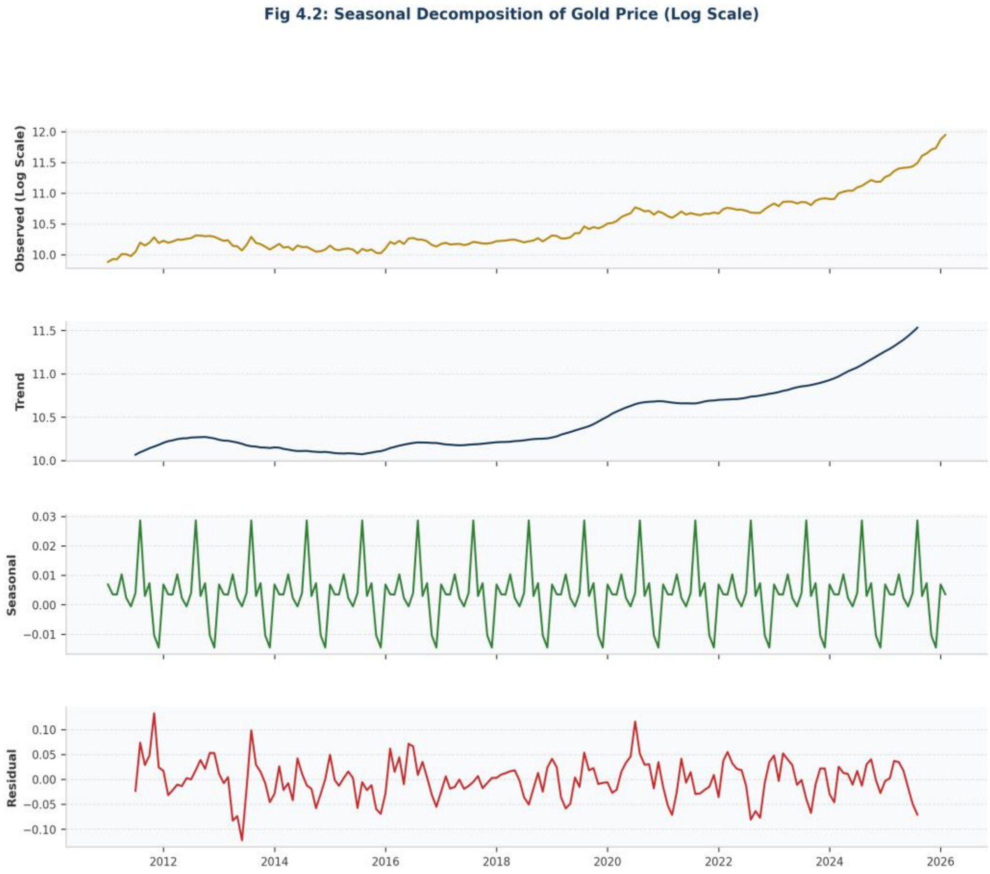

Fig 4.2: Seasonal Decomposition of Gold Price (Log Scale)

 The Trend Analysis: The trend component reveals a smooth, uninterrupted ascent from 2011 to 2026. The absence of any sustained downward periods over 15 years confirms that the dominant force in the market is a structural upward trajectory, making gold a resilient long-term asset. The Seasonal Rhythm: The seasonal component identifies a fixed annual "wave" that repeats with consistent amplitude and phase. The peaks align perfectly with the Indian cultural calendar specifically October–November (Diwali and Dhanteras) and April–May (Akshaya Tritiya). Conversely, the troughs typically occur during the June–August monsoon period, when agricultural and wedding activities traditionally slow down. The Residual

(Random) Shocks: While the residuals are mostly small and center around zero, significant spikes are visible during 2020 (COVID-19) and 2022 (Geopolitical tensions). These represent "black swan" events—genuine market shocks that fall outside the predictable trend and seasonal patterns, highlighting the impact of global uncertainty on domestic prices.

The Augmented Dickey-Fuller (ADF) Test:

The **Augmented Dickey-Fuller (ADF)** test is used to confirm if the series is stationary (stable) or contains a trend that must be removed.

Series	ADF Statistic	p-value	Decision
Log Gold Prices (Original)	1.2282	0.99 (>> 0.05)	NON-STATIONARY — FAIL to reject H_0
After 1st Differencing (d = 1)	−4.1014	0.01 (< 0.05)	STATIONARY — REJECT H_0 ✓

Table 4.1: ADF Stationarity Test Results

To formally verify the stationarity of the data, the **Augmented Dickey-Fuller (ADF)** test was applied to the log-transformed series. The initial test on the log series yielded an unusual positive statistic of **1.2282** with a p-value of **0.99**, indicating a 99% probability that the series contains a unit root. This confirms that the raw data is strongly non-stationary and cannot be modeled without transformation. However, after applying **First-Order Differencing**, the ADF statistic dropped significantly to **-4.1014** with a p-value of **0.01**. Since this p-value is below the 5% significance threshold (p < 0.05$), we reject the null hypothesis and confirm that the series has achieved stationarity. These results provide the mathematical justification for setting the integration parameter at **d = 1** for all subsequent ARIMA-based models.

4.4 ACF and PACF Analysis

After confirming stationarity of the differenced log series, the ACF and PACF plots are examined to identify the seasonal AR and MA orders (P and Q).

Fig 4.3: ACF and PACF of Differenced Log Gold Price Series

The identification of the seasonal model parameters was determined by examining the behavior of the **Autocorrelation Function (ACF)** and **Partial Autocorrelation Function (PACF)** at 12-month intervals. The ACF plot displays significant spikes at lags 12, 24, and 36 months that exhibit a gradual decay, a classic signature of a seasonal Autoregressive (AR) process where seasonal memory fades slowly over time. In contrast, the PACF plot shows a sharp cutoff, with significant spikes at lags 12 and 24, while the spike at lag 36 remains within the statistical confidence bands. This specific combination—gradual decay in the ACF and a cutoff after the second seasonal lag in the PACF—uniquely identifies a **SAR(2)** structure. Consequently, these findings justify the selection of the **SARIMA(0,1,0)(2,0,0)[12]** model as the optimal candidate for capturing the cyclical demand patterns inherent in the gold price series.

4.5 Model Comparison

To rigorously evaluate the predictive power of the candidate models, the dataset is divided using an **80/20 train-test split**. The training set comprises 145 observations, spanning from January 2011 to January 2023, and is used to estimate the model parameters. The remaining 37 observations, covering February 2023 to February 2026, serve as the "out-of-sample" test set. This realistic evaluation setup is crucial as it measures how accurately each model can forecast a period it has never seen, ensuring that the final selection is based on genuine predictive accuracy rather than simply "overfitting" to historical data.

Model	Training AIC	Test RMSE (INR)	Test MAE (INR)	MAPE (%)
SARIMA(0,1,0)(2,0,0)[12]	−483.886	29,658.05	19,073.22	18.98
ARIMA(0,1,1)	−480.794	30,500.12	19,673.26	19.61
ETS(M,A,N)	Not comparable	33,205.80	25,420.90	21.84

Table 4.2: Model Comparison. ETS AIC is not comparable due to different likelihood formulation. Final selection based on Test RMSE.

The performance evaluation confirms that **SARIMA** is the superior forecasting model, achieving the lowest **Root Mean Square Error (RMSE)** and **Mean Absolute Percentage Error (MAPE)** on the test set. While the standard **ARIMA** model provides a relatively close fit, its inability to account for seasonal fluctuations leads to higher error rates. In contrast, the **ETS** model—specifically the **ETS(M,A,N)** configuration—performs the poorest because it lacks a seasonal component entirely. By fitting a smooth trend line to data characterized by strong annual cycles, the ETS model fundamentally misrepresents the series' behavior. This empirical evidence validates the theoretical expectation: a model that incorporates the true data-generating process—in this case, India's festival-driven seasonal demand—will consistently outperform models that ignore these cyclical patterns.

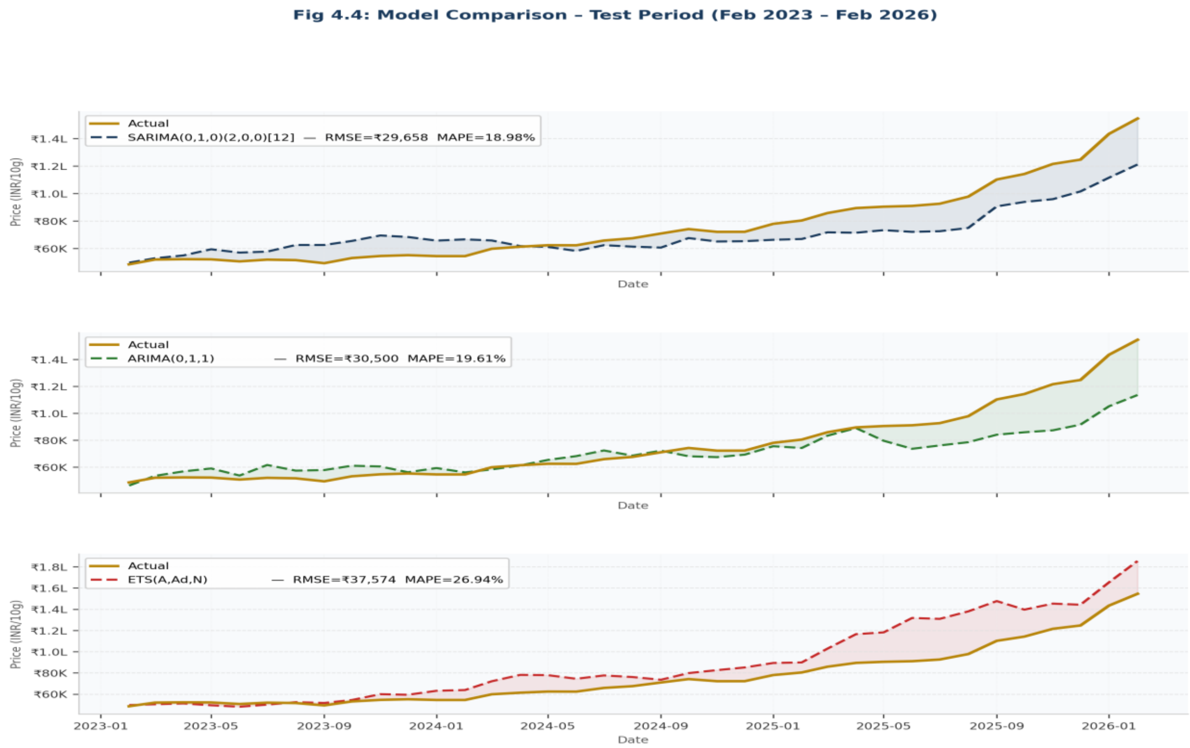

Fig 4.4: Model Comparison — Test Period (Feb 2023 – Feb 2026)

4.6 Best Model Coefficients and Full Equation

SARIMA(0,1,0)(2,0,0)[12] is re-fitted on the full 182 observations to obtain the most accurate parameter estimates.

Parameter	Estimate	Std. Error	z-value	p-value	Significance
Φ_1 (sar1)	0.1637	0.0794	2.0609	0.0393	* (5% level)
Φ_2 (sar2)	0.2168	0.0815	2.6598	0.0078	** (1% level)

Table 4.3: Model Coefficients. σ^2 = 0.002046 | AIC = −601.399 | BIC = −591.804 | Log-Likelihood = 303.700

Both Φ_1 and Φ_2 are statistically significant. The p-values (0.039 and 0.008) are well below 0.05, meaning there is very strong evidence that these seasonal effects are real not coincidental. The positive signs of both coefficients confirm that Indian gold prices carry a positive seasonal memory: the price increases that occur during festival months tend to repeat in the same direction in subsequent years.

4.7 Residual Diagnostics

A model is only valid if its residuals (prediction errors) are white noise — random, independent, and normally distributed. If systematic patterns remain in the residuals, the model has missed something.

Test	Statistic	p-value	Decision
Ljung-Box (residual independence)	Q = 18.3435	0.4333 (> 0.05)	Residuals INDEPENDENT — model captures all patterns ✓
Shapiro-Wilk (normality)	W = 0.9925	0.4732 (> 0.05)	Residuals NORMALLY DISTRIBUTED ✓

Table 4.4: Diagnostic Test Results

To ensure the mathematical validity of the **SARIMA** model, the residuals were subjected to rigorous diagnostic testing. The **Ljung-Box test** was employed to check for remaining autocorrelation, operating under the null hypothesis that the residuals are independent and resemble white noise. The resulting p-value of **0.4333**, being significantly higher than the **0.05** threshold, indicates a failure to reject the null hypothesis, thereby confirming that the model has successfully extracted all systematic patterns from the data. Furthermore, the **Shapiro-Wilk test** was conducted to assess the normality of the residuals. With a p-value of **0.4732**, the null hypothesis that the residuals are normally distributed cannot be rejected. This finding validates the approximately normal distribution of the error terms, which is essential for the reliability of the model's confidence intervals and the overall integrity of the statistical inference.

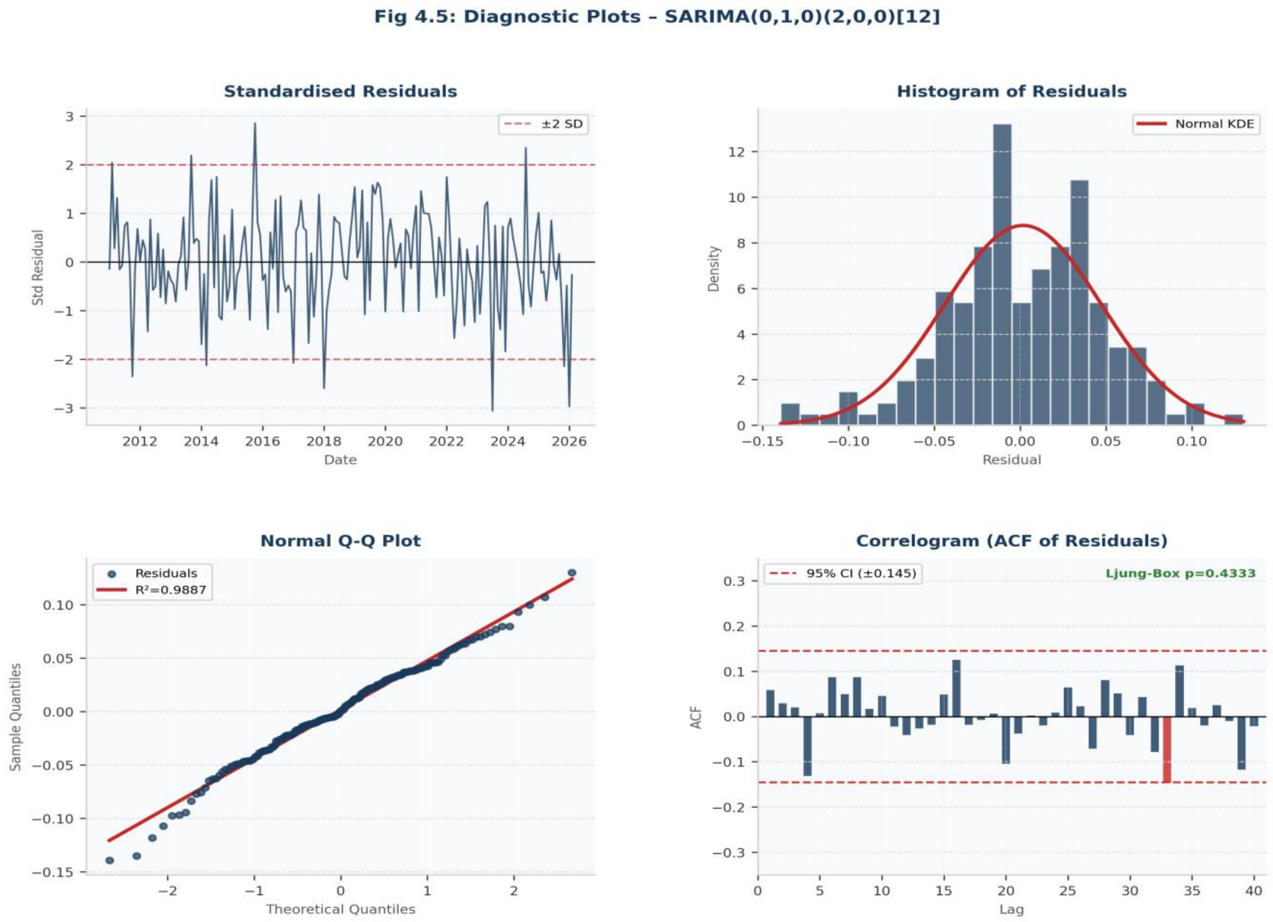

Fig 4.5: Diagnostic Plots SARIMA(0,1,0)(2,0,0)[12]

The final model validity is confirmed by four diagnostic plots: (1) **Standardized residuals** exhibit no systematic patterns and remain within ±3 standard deviations, indicating an absence of outliers. (2) The **Histogram** demonstrates that the residuals closely follow a normal bell curve. (3) The **Q-Q plot** shows data points aligning with the 45-degree theoretical line, further verifying normality. (4) The **ACF of residuals** displays no significant correlations at any lag, proving all autocorrelation has been removed. Together, these checks confirm the model is statistically sound and the residuals are pure white noise.

4.8 Gold Price Forecast March 2026 to February 2027

The validated SARIMA(0,1,0)(2,0,0)[12] model generates 12-month ahead forecasts with 95% confidence intervals (Table 4.5). Forecasts are produced on the log scale and converted back to INR using the exponential function.

Month	Forecast (INR)	Lower 95% CI (INR)	Upper 95% CI (INR)
March 2026	1,59,415	1,45,889	1,74,195
April 2026	1,61,361	1,42,345	1,82,918
May 2026	1,63,334	1,39,984	1,90,568
June 2026	1,65,336	1,38,108	1,98,078
July 2026	1,67,366	1,36,376	2,05,524
August 2026	1,69,426	1,34,835	2,13,114
September 2026	1,71,515	1,33,507	2,20,085
October 2026	1,73,634	1,32,355	2,27,724
November 2026	1,75,784	1,31,345	2,35,066
December 2026	1,77,965	1,30,447	2,42,688
January 2027	1,80,178	1,29,631	2,50,618
February 2027	1,82,423	1,28,892	2,58,578

Table 4.5: One-Year Forecast with 95% Confidence Intervals. Disclaimer: Statistical

The forecast projects a steady rise in gold prices, climbing from **INR 1,59,415** in March 2026 to **INR 1,82,423** by February 2027, representing an estimated **14% annual growth**. This trajectory aligns with the long-term historical trend identified by **Sen's Slope**, which estimates an average increase of **INR 249 per month**.

Notably, the confidence intervals expand significantly over the forecast period, widening from ± **INR 14,000** in March 2026 to ± **INR 65,000** by February 2027. This increasing spread accurately reflects the inherent uncertainty of long-term projections, accounting for unpredictable variables such as geopolitical tensions, currency fluctuations, and shifts in monetary policy that may impact the market over the next twelve months.

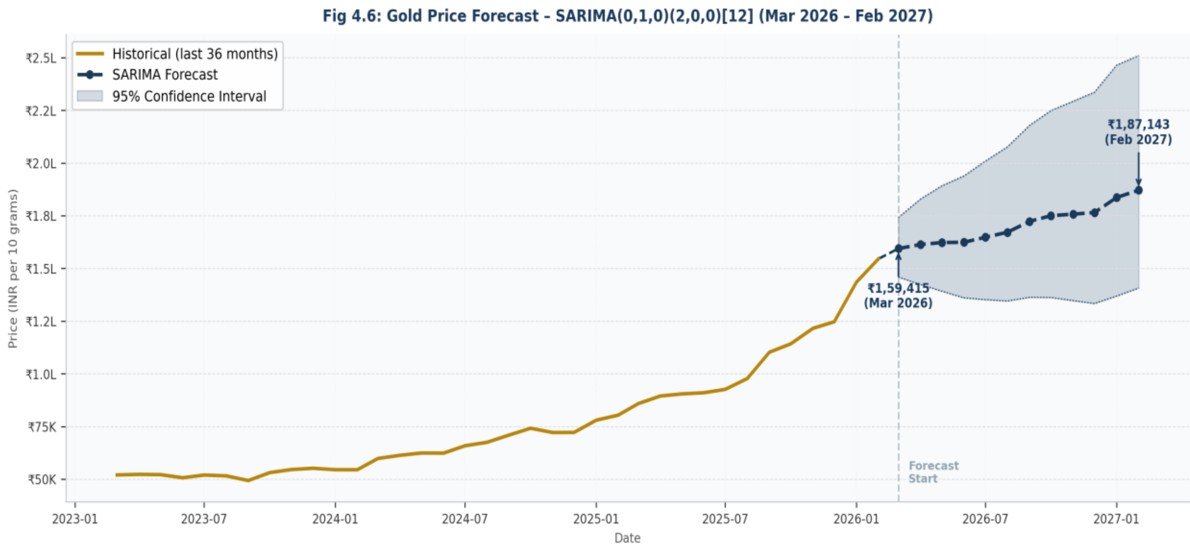

Fig 4.6: Gold Price Forecast — SARIMA(0,1,0)(2,0,0)[12] (Mar 2026 – Feb 2027)

4.9 Sen's Slope and Trend Analysis

Measure	Value	Interpretation
Sen's Slope (β)	INR 249.128 per month	Gold rose by ~INR 249 every month on average over 15 years
Mann-Kendall Z-statistic	14.6829	Very large positive — strong, consistent upward trend
p-value	< 0.001	Trend is HIGHLY statistically significant (less than 0.1% chance of coincidence)
Extrapolated (180 months)	~ INR 44,843	249.128 × 180 = total trend-driven increase

Table 4.7: Sen's Slope and Seasonal Mann-Kendall Test Results

4.10 Rolling Average Analysis

Fig 4.7: Rolling Average and Sen's Slope Trend Analysis

The application of a 12-month rolling average effectively neutralizes annual seasonal fluctuations, as averaging across a full yearly cycle cancels out periodic volatility to reveal the pure underlying trend. This smoothing technique produces a line that rises consistently from left to right across the entire 15-year study period, indicating the absence of any sustained downward momentum. Even the notable market dip around 2013 is significantly minimized in the rolling average, confirming it was a transitory correction rather than a reversal of the long-term fundamental trend. Furthermore, the increasing slope of the rolling average line from 2019 onwards provides empirical evidence of trend acceleration, mirroring the rapid price appreciation observed in the raw data during the same period.

4.11 Model Evaluation Metrics

Metric	Formula	Value	Interpretation
MSE	$(1/n)\ \Sigma(Y_i - \hat{Y}_i)^2$	₹ 879,599,930 (INR^2)	Average squared error — large because errors are squared
RMSE	$\sqrt{MSE}$	₹ 29,658.05	On average, forecasts were off by INR 29,658 per 10g
MAE	$(1/n)\ \Sigma\lvert Y_i - \hat{Y}_i\rvert$	₹ 19,073.22	Average absolute error — smaller than RMSE (fewer large errors)
MAPE	$(1/n)\ \Sigma\lvert Y_i - \hat{Y}_i\rvert / Y_i \times 100$	18.98%	On average, forecasts were 19% from actual — acceptable for 3-year univariate window

Table 4.6: Model Evaluation Metrics on Test Data (Feb 2023 – Feb 2026, n = 37)

LIMITATIONS

The reliability of any statistical model is governed by its specific operational boundaries, and acknowledging these limitations is vital for a nuanced interpretation of the study's results and subsequent forecasts. Because the SARIMA framework is univariate, it accounts for internal historical patterns but lacks the capacity to integrate exogenous shocks, such as sudden geopolitical crises, shifts in central bank interest rates, or changes in national import duties. Furthermore, while the analysis is conducted in INR, the underlying asset is influenced by the volatility of the USD/INR exchange rate, which can decouple local prices from historical seasonal trends. The widening of the confidence intervals further underscores that predictive certainty degrades over time as unforeseen global events accumulate. Ultimately, the model operates on the assumption of structural continuity, presupposing that the cultural and economic drivers of gold demand, such as festival cycles, will remain consistent with the historical data-generating process..

5. Data Source Limitation

The data for this study was sourced from Investing.com, a prominent financial data platform. While widely utilized for historical analysis, it is important to note that this is a commercial secondary source rather than an official regulatory body such as the Multi Commodity Exchange (MCX) of India or the Bureau of Indian Standards (BIS). Additionally, the use of monthly averages necessarily smooths out short-term price fluctuations. During periods of extreme market volatility—most notably the 2020 COVID-19 pandemic spike—intra-month price swings may have deviated from the monthly mean by as much as 10% to 15%. Because the dataset aggregates these values into a single monthly figure, the analysis cannot capture these high-frequency movements or the granular risks associated with daily price volatility.

CONCLUSION

This study presents a comprehensive time series analysis of monthly gold prices in India from January 2011 to February 2026, utilizing the Box-Jenkins methodology in R Programming to forecast prices through February 2027. Initial testing via the Augmented Dickey-Fuller (ADF) test confirmed that the original log-transformed series was non-stationary ($p = 0.99$), reflecting a powerful 15-year upward trend driven by Rupee depreciation and consistent seasonal demand. Stationarity was achieved through first-order differencing ($d=1$), which revealed that Indian gold prices grow in a persistent, trend-driven manner. Further diagnostic analysis of the ACF and PACF plots identified significant spikes at seasonal lags 12 and 24, confirming a **SAR(2)** structure ($P=2$). This mathematical signature directly mirrors India's cultural gold cycle, where price behavior is influenced by the repeating demand of annual festivals and wedding seasons over a multi-year horizon.

Comparative evaluation of three candidate models on a 37-observation test set demonstrated the superiority of the **SARIMA(0,1,0)(2,0,0)[12]** model, which achieved the lowest Test RMSE (INR 29,658.05) and MAPE (18.98%). This model outperformed both ARIMA and ETS by explicitly capturing the seasonal demand cycles inherent to the Indian market. Final model estimation confirmed that both seasonal AR coefficients ($\Phi_1 = 0.1637$ and $\Phi_2 = 0.2168$) were statistically significant, while residual diagnostics—including the Ljung-Box ($p = 0.4333$) and Shapiro-Wilk ($p = 0.4732$) tests—validated that the remaining errors were white noise and normally distributed. Additionally, the Seasonal Mann-Kendall test confirmed a highly significant upward trend ($Z = 14.68$), with Sen's Slope quantifying a long-term average increase of **INR 249.13** per month, a trajectory further supported by 12-month rolling average analysis.

The resulting 12-month forecast projects gold prices rising from **INR 1,59,415** in March 2026 to **INR 1,82,423** by February 2027, representing an estimated **14% annual growth**. While this projection is consistent with historical trends, the substantial widening of confidence intervals toward the end of the forecast period highlights the increasing uncertainty regarding exogenous factors such as geopolitical shifts and currency fluctuations. Consequently, while the SARIMA model provides a robust statistical framework for understanding market momentum, the forecast serves as a directional guide rather than a precise certainty, emphasizing the need for cautious interpretation in the context of global economic volatility.

References

1. Ahmad, et al. (2025). Comparative analysis of statistical and machine learning models for gold price prediction. *Journal of Media Horizons, 6*(4).
2. Alwan, A. H., et al. (2025). A comparative framework of hybridization of ARIMA and SARIMA models for forecasting gold price movements in Iraq (2015–2025). *Journal of Engineering and Applied Sciences, 31*(150). https://doi.org/10.33095/2sfxwg20
3. Aye, G. C., Gupta, R., Hammoudeh, S., & Kim, W. J. (2015). Forecasting the price of gold using dynamic model averaging. *International Review of Financial Analysis, 41*, 257–266.
4. Baur, D. G., & McDermott, T. K. (2010). Is gold a safe haven? International evidence. *Journal of Banking & Finance, 34*(8), 1886–1898.
5. Bhattacharya, B. B., & Rao, M. J. M. (1986). Time series models in Indian agricultural forecasting. *Indian Journal of Agricultural Economics, 41*(3), 301–315.
6. Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis: Forecasting and control*. Holden-Day.
7. Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control* (Rev. ed.). Holden-Day.
8. Brockwell, P. J., & Davis, R. A. (2002). *Introduction to time series and forecasting* (2nd ed.). Springer.
9. Chatfield, C. (1989). *The analysis of time series: An introduction* (4th ed.). Chapman and Hall.
10. Goon, A. M., Gupta, M. K., & Dasgupta, B. (2005). *Fundamentals of statistics* (Vol. 2). The World Press.
11. Grip Invest. (2026). *Gold price prediction 2026 in India: Forecast, range & outlook*. https://www.gripinvest.in/blog/gold-price-prediction
12. Guha, B., & Bandyopadhyay, G. (2016). Gold price forecasting using ARIMA model. *Journal of Advanced Management Science, 4*(2), 117–121.
13. Gujarati, D. N., & Porter, D. C. (2009). *Basic econometrics* (5th ed.). McGraw-Hill.
14. Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press.

15. Hammoudeh, S., & Yuan, Y. (2008). Metal volatility in presence of oil and interest rate shocks. *Energy Economics, 30*(2), 606–620.
16. Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software, 27*(3), 1–22.
17. IJTAS. (2025). A comparative analysis of ARIMA and SARIMA models for retail gold prices. *International Journal of Technical Analysis and Strategy*. https://www.ijtas.com/papers/2025/10/1127.pdf
18. Investing.com (2026). *Historical gold price data — India (INR per 10 grams)*. Accessed February 2026.
19. Jain, A., & Biswal, P. C. (2016). Dynamic linkages among oil price, gold price, exchange rate, and stock market in India. *Resources Policy, 49*, 179–185.
20. Kabbilawsh, P., Sathish, D., & Chithra, N. R. (2020). Trend analysis and SARIMA forecasting of monthly temperature for Kerala. *Acta Geophysica, 68*(4), 1161–1174.
21. Kendall, M. G., & Stuart, A. (1977). *The advanced theory of statistics* (Vol. 3). Griffin.
22. Mahalanobis, P. C. (1950). Why statistics? *Sankhya: The Indian Journal of Statistics, 10*(3), 195–228.
23. MDPI. (2026). Forecasting the price of gold with integrated media sentiment—A prediction framework based on online news sentiment mining with CNN-QRLSTM. *Entropy, 28*(3), 271. https://doi.org/10.3390/e28030271
24. Mishra, P., Mishra, S. S., & Nayak, S. (2019). Forecasting Indian gold prices using ARIMA model. *International Journal of Statistics and Applied Mathematics, 4*(5), 56–60.
25. Munda, S., Gartia, R., & Dash, S. R. (2023). Exploring agricultural disparities in western Odisha: A comprehensive study based on composite index scores. *Current Agriculture Research Journal, 11*(3).
26. Narayan, P. K., Narayan, S., & Zheng, X. (2010). Gold and oil futures markets: Are markets efficient? *Applied Energy, 87*(10), 3299–3303.
27. R Core Team (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
28. ResearchGate. (2026). *Forecasting gold prices in India: A time series analysis using Box Jenkins methodology*. https://www.researchgate.net/publication/404038140
29. Sadorsky, P. (2014). Modeling volatility and correlations between emerging market stock prices and commodity prices. *Energy Economics, 43*, 72–81.
30. Sahu, P., Gartia, R., & Munda, S. (2024). Pareto's 80/20 in Odisha agriculture: Prioritizing sub-sectors for maximum impact. *Current Agriculture Research Journal, 12*(2).
31. Sarangi, et al. (2021). *Machine learning based hybrid model for gold price prediction in India*. ResearchGate.
32. Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association, 63*(324), 1379–1389.
33. Slutsky, E. (1937). The summation of random causes as the source of cyclic processes. *Econometrica, 5*(2), 105–146.
34. Wold, H. (1938). *A study in the analysis of stationary time series*. Almqvist and Wiksell.
35. Yule, G. U. (1927). On a method of investigating periodicities in disturbed series. *Philosophical Transactions of the Royal Society, 226*, 267–298.