International Journal of Leading Research Publication (IJLRP)

E-ISSN: 2582-8010 ● Website: www.ijlrp.com ● Email: editor@ijlrp.com

Linear versus Exponential Forecasting Techniques: A Comparative Study of ARIMA and Holt–Winters Triple Exponential Smoothing Models for Indian Retail Gold Prices - Evidence from daily retail gold price data in India, 2014–2025

Rajib Bhattacharya

Associate Professor
NSHM Business School, Kolkata
NSHM Knowledge Campus – Kolkata Group of Institutions

Abstract:
This study conducts a comparative evaluation of two classical time-series forecasting approaches—AutoRegressive Integrated Moving Average (ARIMA) and Holt–Winters Triple Exponential Smoothing (HW–TES)—to analyse and predict daily retail gold prices in India over the period 2014–2025. Gold, as both a cultural and financial asset, exhibits strong cyclical and seasonal dynamics, reflecting patterns driven by festive consumption, investment demand, and global macroeconomic shifts. The study aims to determine which framework—linear stochastic modelling or exponential adaptive smoothing—offers superior predictive performance and responsiveness under conditions of market volatility and structural change.
The methodology adopts the Box–Jenkins approach for ARIMA model identification and the multiplicative Holt–Winters formulation for capturing trend and seasonality. Both models are calibrated using daily 24-carat retail gold prices sourced from official data repositories. The comparative evaluation is based on accuracy indicators such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE), supplemented by the Diebold–Mariano (DM) test to assess the statistical significance of forecasting differences.
Empirical results reveal that both models perform commendably, with MAPE values below 5%, confirming high short-term predictive reliability. However, the Holt–Winters model consistently demonstrates slightly lower forecast errors and greater adaptability to rapid price shifts, attributed to its exponential weighting mechanism that prioritizes recent observations. In contrast, ARIMA's linear lag structure produces delayed adjustments to trend reversals, leading to mild underestimation during volatile phases.
The findings confirm that while ARIMA retains its value as a robust, interpretable statistical benchmark, Holt–Winters exponential smoothing offers superior short-run responsiveness and adaptability, making it particularly suitable for dynamic and seasonally driven markets like India's retail gold sector. This study contributes to the broader forecasting literature by reaffirming that model selection should be context-dependent, guided by the temporal characteristics and behavioural complexity of financial time series.

Keywords: Gold Price Forecasting, ARIMA, Holt–Winters, Time-Series Modelling, Seasonality
JEL Classification: C22, C53, E37, G17, Q02

INTRODUCTION

Gold occupies a distinctive position in both global and Indian economies, serving simultaneously as a commodity, investment asset, and monetary hedge. Its intrinsic value, liquidity, and universal acceptance make it an essential component of household savings and institutional portfolios. In India, gold carries profound cultural and economic significance, where fluctuations in its price influence consumer demand, inflation expectations, and macro-financial stability. As a result, accurate forecasting of gold prices has become a critical area of research for economists, traders, and policymakers seeking to anticipate market movements, manage financial risk, and design effective policy responses.

Forecasting the price of gold poses substantial challenges due to its volatile, nonlinear, and seasonally adaptive nature. The Indian retail gold market, in particular, exhibits cyclical demand linked to festivals, weddings, and investment behaviour. Traditional time-series methods such as ARIMA (AutoRegressive Integrated Moving Average) have long been valued for their statistical robustness, ability to model autocorrelations, and suitability for stationary processes. ARIMA models express current values as linear combinations of past observations and stochastic shocks, thereby capturing underlying temporal dependencies within a structured framework. However, ARIMA models operate under a linear stochastic assumption, which may restrict their responsiveness to abrupt structural shifts or seasonal accelerations commonly observed in financial time series like gold prices.

In contrast, Holt–Winters Triple Exponential Smoothing (HW–TES) represents an adaptive exponential technique capable of accommodating both trend and seasonal components. By assigning exponentially decreasing weights to past observations, the HW–TES model gives greater significance to recent data, enabling faster adaptation to emerging market trends. Its triple smoothing mechanism decomposes the series into level, trend, and seasonality, making it especially effective in environments characterized by cyclical consumption patterns and recurrent price fluctuations. Unlike ARIMA, which relies on differencing and lagged error structures, the HW–TES model directly incorporates dynamic updating through three smoothing parameters—α (level), β (trend), and γ (seasonal adjustment)—providing an inherently nonlinear mechanism for adaptive forecasting.

The comparative study of these two frameworks—ARIMA as a linear stochastic model and Holt–Winters as an exponential adaptive model—serves to investigate the trade-off between statistical rigour and adaptive flexibility in financial forecasting. Empirical evidence from global and Indian contexts has shown that while ARIMA remains reliable for short-horizon forecasts, it often underperforms during periods of volatility or structural change. Conversely, exponential smoothing techniques demonstrate enhanced responsiveness to short-term fluctuations, though sometimes at the cost of interpretability and long-term stability. Thus, an empirical comparison grounded in daily Indian retail gold price data over a decade-long period provides valuable insights into the contextual efficiency of each model.

Between 2014 and 2025, the Indian gold market experienced several critical phases—global commodity price volatility, demonetization impacts, changes in import duty structures, the COVID-19 pandemic, and global inflationary pressures. Each of these events contributed to heightened uncertainty and shifts in market sentiment, necessitating forecasting models capable of rapid recalibration. The Holt–Winters model, by virtue of its exponential responsiveness, is theoretically better equipped to handle such dynamic conditions. However, the ARIMA model's analytical foundation and systematic error correction process make it a robust benchmark for statistical comparison.

The study's findings reveal that both models perform commendably, with mean forecast errors below 5%, but the Holt–Winters model consistently demonstrates slightly superior adaptability and lower cumulative prediction error. This difference, though marginal, is statistically meaningful within the context of financial forecasting, where even minor improvements can translate into significant economic gains. The results indicate that while ARIMA's linear structure effectively models persistence and autocorrelation, it underestimates sharp upward movements in price, reflecting its lagging response to recent shocks. On the other hand, the exponential smoothing mechanism of HW–TES more accurately tracks real-time fluctuations, particularly during periods of rapid price escalation.

From a broader perspective, this comparative inquiry contributes to the methodological discourse on model selection for financial time-series forecasting. It highlights that no single model universally outperforms others; rather, forecasting efficacy depends on the data characteristics, market environment, and desired time horizon. For short-term operational forecasting, particularly in volatile and seasonally driven markets like Indian retail gold, exponential models such as Holt–Winters offer distinct practical advantages. Conversely, for strategic or policy-level forecasting, where interpretability and structural diagnostics are paramount, linear models like ARIMA remain indispensable.

This study positions itself at the intersection of classical econometric theory and applied forecasting practice, demonstrating that the comparative performance of linear and exponential models offers valuable lessons for researchers and practitioners alike. The findings reaffirm that forecasting precision is enhanced not merely through model complexity but through the alignment of model structure with market behaviour, thereby contributing meaningfully to both academic literature and the operational domain of financial prediction.

SURVEY OF LITERATURE

Gold price forecasting continues to attract extensive academic and financial interest due to gold's multifaceted economic functions—as a commodity, a monetary substitute, a hedge against inflation, and a safe-haven asset during crises. Its price dynamics are complex, characterized by nonlinearity, heteroskedasticity, regime shifts, and sensitivity to global macroeconomic indicators. The scholarly literature on gold price prediction can broadly be grouped into two methodological streams: (a) traditional econometric and statistical approaches, including ARIMA, VAR/VECM, GARCH and its variants, DCC/GAS, cointegration, mixed-data frequency models (MIDAS, GARCH-MIDAS), and wavelet or time-frequency analyses; and (b) modern machine-learning and hybrid models, including ANN, SVM, ELM, DBN, LSTM, CNN, and decomposition-based hybrid methods.

Early research predominantly relied on classical linear time-series and model-averaging frameworks to explore gold price predictability and its macro-financial determinants. Aye et al. (2015) employed dynamic model averaging (DMA) to identify time-varying predictor importance—including exchange rates, interest rates, and financial stress indices—and demonstrated that predictor relevance shifts over time and across market episodes, underscoring the advantage of model-averaging when relationships are unstable (Aye et al., 2015). Although ARIMA and ARIMAX models remain useful short-term benchmarks, their performance declines when volatility structures change or structural breaks occur.

The observation of volatility clustering and fat-tailed distributions in gold returns prompted the widespread adoption of ARCH/GARCH models. Tully and Lucey (2007) used an asymmetric power GARCH (APGARCH/APARCH) model to capture leverage and power effects in both cash and futures markets, concluding that APGARCH provides superior fit relative to standard GARCH specifications (Tully & Lucey, 2007). Subsequent studies employing EGARCH, TGARCH, and APARCH variants confirmed the persistence of conditional heteroskedasticity in gold price movements.

Beyond univariate volatility models, multivariate specifications such as DCC-GARCH, BEKK, and GAS have become popular for studying co-movements among gold, equities, oil, and exchange-rate markets. Ciner, Gurdgiev and Lucey (2013) and Reboredo (2013) utilized dynamic conditional correlation and copula-based methods to investigate gold's hedge and safe-haven characteristics, revealing that such multivariate volatility frameworks improve joint risk forecasting and portfolio diversification strategies (Ciner et al., 2013; Reboredo, 2013).

Macroeconomic fundamentals also exert significant long-term influence on gold volatility. The GARCH-MIDAS model (and related spline extensions) decomposes volatility into short-run GARCH dynamics and long-run components driven by low-frequency macro variables. Fang, Yu and Xiao (2018) and Salisu et al. (2020) showed that incorporating policy uncertainty, industrial output, and macroeconomic principal components enhances long-horizon volatility forecasts for both spot and futures gold markets (Fang et al., 2018; Salisu et al., 2020). These models are particularly effective when integrating monthly or quarterly macroeconomic data with daily financial series.

Long-run interactions among gold prices, exchange rates, inflation, and interest rates have been investigated extensively, though findings remain inconclusive. Some studies identify cointegration and stable long-term relationships suitable for vector error-correction modelling, whereas others observe episodic or time-varying linkages. Copula-based and tail-dependence analyses (Reboredo, 2013) have further refined the understanding of extreme co-movements and safe-haven dynamics, often producing insights that differ from unconditional correlation measures.

Time–frequency and wavelet techniques offer another dimension by decomposing price series into components representing different investment horizons. These studies reveal that predictive relationships vary across scales—factors significant at daily frequencies may lose relevance at monthly intervals. Consequently, hybrid wavelet–ARIMA or wavelet–ANN models often achieve superior forecasting accuracy by capturing horizon-specific features.

From the 2010s onward, machine learning (ML) and hybrid methodologies have gained prominence. Kristjanpoller and Minutolo (2015) introduced an ANN–GARCH hybrid framework that integrates conditional heteroskedasticity modelling with nonlinear learning, reducing forecast errors compared to standalone models (Kristjanpoller & Minutolo, 2015). Following this, numerous studies combined wavelet or empirical mode decomposition (EMD) with ML algorithms such as SVM, ANN, GRU, or LSTM, illustrating that decomposition enhances the learning process by simplifying complex temporal structures (E. Jianwei et al., 2019).

Deep learning (DL) architectures, including deep belief networks (DBN), LSTM, GRU, and CNN-LSTM hybrids, have emerged as a new frontier in gold price forecasting. Zhang and Ci (2020) applied a DBN to predict gold prices and found significant improvements over both traditional econometric and shallow ML models. Similarly, Khani et al. (2021) compared CNN, LSTM, and encoder–decoder LSTM configurations—including pandemic-related variables—and reported that deep recurrent structures achieve superior short-term predictive accuracy, particularly when augmented with contextual or domain-specific information (Khani et al., 2021).

Recent innovations include Extreme Learning Machines (ELM) and their online sequential adaptations. Weng et al. (2020) proposed GA-regularized ELM models that demonstrate rapid training and robust performance for high-frequency or online forecasting contexts. Tree-based ensemble and boosting algorithms have also gained traction. Pierdzioch et al. (2016) utilized boosting and quantile boosting approaches to forecast gold volatility and returns using a wide range of predictors and asymmetric loss functions, reporting consistent out-of-sample improvements over benchmark models (Pierdzioch et al., 2016). More recent contributions by Foroutan et al. (2024) and Cohen (2023) tested graph neural networks and ensemble ML pipelines (XGBoost, LightGBM) with sophisticated feature engineering, producing encouraging results for multi-asset and retail-level gold price forecasting.

Across these methods, several empirical regularities emerge. Volatility modelling plays a decisive role: frameworks that explicitly capture conditional variance dynamics (GARCH, GARCH-MIDAS, GAS) outperform naïve ARIMA-type benchmarks in both risk forecasting and option-pricing contexts. The inclusion of macroeconomic variables enhances long-run predictability, while nonlinear and hybrid methods improve short-term accuracy. Deep learning and decomposition-based ML models frequently surpass linear econometric models for short-horizon (1–30 day) forecasts, particularly when exogenous variables such as exchange rates, interest rates, policy uncertainty, realized volatility, or pandemic metrics are incorporated.

Model instability and structural breaks remain major obstacles to forecast reliability. Performance is highly regime-dependent—during crises like 2008 or COVID-19, the relative efficacy of competing models often shifts. Adaptive strategies such as DMA, time-varying parameter models, rolling-window estimation, and online ELM frameworks are effective in mitigating these challenges.

Despite considerable empirical progress, several research gaps persist. Many ML models yield high predictive accuracy but limited economic interpretability. Integrating ML algorithms with structural econometric constraints represents a promising future direction. Moreover, most research concentrates on bullion or futures markets, neglecting retail gold prices influenced by local taxes, jewellery premiums, and distribution costs. Incorporating real-time and alternative datasets—news sentiment, supply-chain disturbances, or order-book information—into econometric–ML hybrids remains an emerging trend. Rigorous, standardized multi-horizon evaluations using RMSE, MAE, Theil's U, and Diebold–Mariano tests are also essential to enable consistent model comparison.

In essence, gold price forecasting has evolved into a hybrid discipline integrating classical econometrics with data-driven machine learning. Traditional econometric tools (ARIMA, VAR, GARCH, GARCH-MIDAS) remain indispensable for modelling volatility and systemic risk, while ML and hybrid decomposition techniques dominate short-term predictive applications. Model selection should align with the forecasting objective: GARCH-type frameworks for volatility, DMA or time-varying models for structural instability, and deep-learning hybrids for high-frequency prediction.

However, the literature still lacks systematic comparative evaluations. Most studies examine individual models—ARIMA, VAR/VECM, GARCH, GARCH-MIDAS, or ML hybrids—without situating their findings within a unified benchmarking framework. Consequently, conclusions about forecasting accuracy remain dataset-specific and context-dependent. Few studies rigorously benchmark econometric and ML models using identical datasets or standardized metrics, and cross-regime comparisons remain rare. Additionally, geographic and temporal concentration limits understanding of cross-country heterogeneity in gold market dynamics.

Addressing these methodological and empirical gaps through comprehensive comparative analyses—spanning classical, volatility-based, mixed-frequency, and hybrid paradigms—would significantly advance gold market research. Such efforts would yield more generalizable insights into model performance across market regimes and time horizons, providing valuable guidance to investors, policymakers, and central banks engaged in forecasting, risk assessment, and macro-financial management of gold markets worldwide.

OBJECTIVES OF THE STUDY

The primary objective of this study is to conduct a comprehensive comparative evaluation of linear and exponential forecasting techniques—represented by the AutoRegressive Integrated Moving Average (ARIMA) and Holt–Winters Triple Exponential Smoothing (HW–TES) models, respectively—in forecasting daily retail gold prices in India for the period 2014–2025. The study aims to determine the relative forecasting efficiency, responsiveness, and adaptability of these two widely used time-series models in capturing short-term trends, seasonal patterns, and cyclical fluctuations inherent in gold price movements.

Gold, as an asset class, exhibits dual characteristics of consumption and investment, with its price behaviour shaped by macroeconomic conditions, global uncertainties, and local seasonal demand cycles. The research seeks to explore whether the linear stochastic structure of ARIMA or the adaptive exponential weighting mechanism of the HW–TES model provides superior accuracy in forecasting daily retail gold prices. Specifically, it evaluates the capacity of both models to accommodate volatility, trend persistence, and short-run deviations under dynamic market conditions.

By applying statistical accuracy measures such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE), the study quantifies and compares the predictive precision of both models. Furthermore, it examines the systematic tendencies—such as underestimation or overreaction—to identify model-specific biases in capturing real market dynamics.

The overarching objective extends beyond mere quantitative comparison: it seeks to assess the methodological suitability of linear versus exponential frameworks in forecasting a complex financial time series like gold. The findings aim to inform researchers, investors, and policymakers regarding the practical and theoretical implications of choosing between ARIMA and Holt–Winters models for short-term gold price prediction in an emerging market context like India.

METHODOLOGY OF THE STUDY

This section presents the statistical framework and mathematical formulations employed in the comparative analysis of linear and exponential forecasting models for cryptocurrency prices. The study examines two prominent time-series forecasting methodologies: the AutoRegressive Integrated Moving Average (ARIMA) model, representing the linear stochastic approach, and the Holt–Winters Triple Exponential Smoothing (HW-TES) method, representing the exponential smoothing approach. The comparison is performed o...

Data Representation and Transformation

Let $\{P_t\}_{t=1}^T$ denote the daily closing price of a cryptocurrency over T observations. Due to the high volatility and potential nonstationarity of price data, the logarithmic transformation is applied to stabilize variance and linearize exponential growth patterns:

$$y_t = \ln(P_t).$$

For forecasting analysis, two data representations are considered:
1. Log-transformed prices y_t for level prediction.
2. Log returns $r_t = \ln(P_t / P_{t-1})$ for rate-of-change analysis.

Testing for stationarity is performed using the Augmented Dickey–Fuller (ADF) and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests. If y_t is nonstationary, differencing of order d is applied to achieve stationarity:

$$(1 - B)^d y_t = \varepsilon_t,$$

where B is the backshift operator ($By_t = y_{t-1}$).

The ARIMA (p, d, q) Model

The ARIMA model combines autoregressive (AR), differencing (I), and moving average (MA) components to capture temporal dependencies and stochastic structure in the data. It is represented as:

$$\varphi(B)(1 - B)^d y_t = \theta(B)\varepsilon_t, \quad \varepsilon_t \sim \text{iid}(0, \sigma^2),$$

where:
- $\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p$ is the autoregressive (AR) polynomial,
- $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$ is the moving-average (MA) polynomial,
- d is the order of differencing ensuring stationarity.

The ARIMA model estimates parameters (φ_i, θ_j) by maximizing the log-likelihood function:

$$\ln L(\theta) = -\tfrac{1}{2}T \ln(2\pi\sigma^2) - \tfrac{1}{2}\Sigma (\varepsilon_t^2 / \sigma^2),$$

where θ denotes the parameter vector.

Model identification follows the Box–Jenkins methodology, involving the analysis of autocorrelation (ACF) and partial autocorrelation (PACF) plots to determine optimal orders (p, d, q). Model selection is guided by information criteria:

$AIC = -2 \ln(\mathscr{L}) + 2k$,
$BIC = -2 \ln(\mathscr{L}) + k \ln(T)$,

where $\mathscr{L}$ denotes the likelihood, and k the number of estimated parameters. The model with the smallest AIC/BIC is selected.

Forecasts for h periods ahead are generated recursively as:

$$\hat{y}_{T+h|T} = c + \Sigma_{i=1}^{p} \varphi_i \hat{y}_{T+h-i|T} + \Sigma_{j=1}^{q} \theta_j \hat{\varepsilon}_{T+h-j},$$

with prediction intervals given by:

$$PI_{1-\alpha} = \hat{y}_{T+h|T} \pm z_{\alpha/2} \sqrt{Var(\hat{\varepsilon}_{T+h|T})}.$$

Exponential Smoothing Methods
Exponential smoothing techniques forecast future values by weighting past observations exponentially, with more recent values receiving higher weights. Let α, β, and γ denote the smoothing parameters for level, trend, and seasonality, respectively, where $0 < \alpha, \beta, \gamma < 1$.

The general updating equations for level (L_t), trend (T_t), and seasonal (S_t) components are expressed as:

$L_t = \alpha(y_t / S_{t-s}) + (1 - \alpha)(L_{t-1} + T_{t-1})$,
$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}$,
$S_t = \gamma(y_t / L_t) + (1 - \gamma)S_{t-s}$,

where s denotes the seasonal cycle length (e.g., s = 7 for weekly seasonality or s = 365 for annual periodicity).

The h-step-ahead forecast equation for the Holt–Winters multiplicative model is:

$$\hat{y}_{t+h|t} = (L_t + hT_t)S_{t-s+h},$$

while for the additive version:

$$\hat{y}_{t+h|t} = L_t + hT_t + S_{t-s+h}.$$

Holt–Winters Triple Exponential Smoothing (HW-TES)

The Holt–Winters method, initially developed by Holt (1957) and Winters (1960), extends simple exponential smoothing by incorporating both trend and seasonal components. Two variations are used:

1. Additive HW-TES: suitable when seasonal variations are roughly constant over time.
2. Multiplicative HW-TES: suitable when seasonal variations change proportionally with the series level.

The multiplicative version is often more appropriate for financial data, where volatility scales with price level.

The optimization of (α, β, γ) is achieved by minimizing the sum of squared forecast errors (SSE):

$$SSE = \Sigma (y_t - \hat{y}_t)^2.$$

Parameter estimation is typically performed via nonlinear optimization algorithms, such as the Levenberg–Marquardt or L-BFGS-B methods.

Comparative Model Framework

The ARIMA and HW-TES models are compared on the basis of both in-sample fitting accuracy and out-of-sample forecast performance. While ARIMA captures autocorrelation structures through differencing and parameter estimation, the HW-TES approach directly models level, trend, and seasonality through smoothing equations.

The comparative structure can be represented as:

ARIMA: $y_t = f(y_{t-1}, ..., y_{t-p}, \varepsilon_{t-1}, ..., \varepsilon_{t-q})$,
HW-TES: $y_t = g(L_t, T_t, S_t, \alpha, \beta, \gamma)$.

Forecast accuracy is measured through multiple error metrics and statistical tests.

Evaluation Metrics

Model performance is evaluated using both deterministic and probabilistic measures.

1. Mean Absolute Error (MAE):
   $MAE = (1/N)\, \Sigma\, |y_t - \hat{y}_t|$.

2. Root Mean Square Error (RMSE):
   $RMSE = \sqrt{(1/N)\, \Sigma\, (y_t - \hat{y}_t)^2}$.

3. Mean Absolute Percentage Error (MAPE):
   $MAPE = (100/N)\, \Sigma\, |(y_t - \hat{y}_t) / y_t|$.

4. Theil's U-statistic:
   $U = \sqrt{\Sigma(y_t - \hat{y}_t)^2} / \sqrt{\Sigma(y_t - y_{t-1})^2}$.

A Theil's $U < 1$ implies superior performance relative to a naïve random walk.

5. Diebold–Mariano (DM) Test:
   $DM = \bar{d} / \sqrt{Var(\bar{d})/T}$,
   where $d_t = L(e_t^{(1)}) - L(e_t^{(2)})$,
   testing H_0: equal predictive accuracy.

Seasonal and Trend Decomposition

For interpretive clarity, both models implicitly decompose the observed series into components:

$$y_t = T_t + S_t + I_t,$$

where T_t represents the trend, S_t the seasonal pattern, and I_t the irregular component. In ARIMA, decomposition is achieved through differencing, while HW-TES achieves it through recursive smoothing of components.

Diagnostic Testing
Diagnostic validation ensures the adequacy of fitted models through the following procedures:

1. Residual Autocorrelation: Ljung–Box Q-statistic tests H_0: no serial correlation up to lag m.
2. Heteroskedasticity: ARCH-LM test checks for conditional variance clustering.
3. Normality: Jarque–Bera (JB) test examines the normal distribution of residuals.

A satisfactory model should exhibit white-noise residuals with no systematic patterns in ACF/PACF plots.

Forecasting Horizon and Rolling Evaluation
The forecasting exercise covers short-, medium-, and long-term horizons (e.g., 7, 30, and 90 days ahead). Rolling-origin evaluation is adopted to ensure robustness: the model is re-estimated after each forecast step to simulate real-time prediction conditions. Forecast intervals are computed as:

$$PI_{1-\alpha} = \hat{y}_{T+h|T} \pm z_{\alpha/2} \sqrt{Var(\hat{\varepsilon}_{T+h|T})}.$$

Model Comparison Framework
The comparison between ARIMA and HW-TES is based on the following hypotheses:

H_0: No significant performance difference between ARIMA and HW-TES forecasts.
H_1: Significant performance difference between the two models.

The Diebold–Mariano test statistic and RMSE ratio are employed to evaluate these hypotheses. Additionally, model parsimony, computational efficiency, and interpretability are considered in the comparative evaluation.

Interpretation of Forecasting Dynamics
- ARIMA captures stochastic dependencies and autoregressive persistence.
- HW-TES captures deterministic trends and seasonality using exponential decay weights.
- Exponential models react faster to recent market changes, while ARIMA relies on past lag relationships.

The methodological framework provides a structured comparison between the linear ARIMA model and the nonlinear Holt–Winters exponential smoothing technique. By integrating stochastic modelling and deterministic smoothing paradigms, the study assesses the relative efficacy of both approaches in forecasting the highly volatile and seasonally adaptive price series of Bitcoin and Cardano.

FINDINGS OF THE STUDY AND IMPLICATIONS THEREOF
Charts 1 to 4 in the study collectively present a graphical synthesis of the comparative forecasting performance of the ARIMA and Holt–Winters Triple Exponential Smoothing (HW–TES) models in predicting daily retail gold prices in India during the study period. Each chart complements the quantitative evidence from Tables 1–4, enabling an interpretive visualization of accuracy, deviation, and efficiency between the two forecasting frameworks.

Chart 1 depicts the comparative trajectories of *actual versus predicted gold prices* over the ten-day observation horizon. The plotted series clearly illustrates that both ARIMA and HW–TES models follow the same general trend of the actual price curve, effectively capturing the direction and pattern of price movements. However, the predicted lines for both models consistently lie slightly below the actual price curve, reflecting a mild tendency toward underestimation. The Holt–Winters model, owing to its exponential smoothing design, tracks the actual series more closely during periods of sudden upward movements (notably between Days 6–9), highlighting its superior adaptability to short-term fluctuations and nonlinear accelerations in gold price behaviour.

Chart 2 presents the *absolute errors* in prediction, indicating the magnitude of deviation between actual and forecasted values for both models. The chart reveals a progressive increase in errors across the ten days, demonstrating that predictive accuracy diminishes marginally with longer forecast horizons—an inherent feature of time-series models. The Holt–Winters model consistently yields smaller error bars compared with ARIMA, signifying a more efficient error-minimization mechanism through its exponentially weighted adjustments to recent data points. Both models, however, maintain relatively small error magnitudes, underscoring their reliability for short-term forecasts in the gold market.

Chart 3 displays the *absolute percentage errors (APE)*, standardizing deviations relative to actual prices. The APE lines for both models move almost in tandem, rising moderately over time as forecast uncertainty accumulates. The Holt–Winters line, though closely parallel to ARIMA, lies slightly below it across most observations, visually reinforcing its higher proportional accuracy. The chart also captures periods of heightened market volatility where both models' errors spike, suggesting that while the models perform robustly under normal conditions, they face challenges during abrupt market surges—an aspect warranting hybrid or nonlinear extensions in future research.

Chart 4 summarizes the comparative *Mean Absolute Percentage Error (MAPE)* values for the two models—4.8865% for ARIMA and 4.8734% for HW–TES. The marginal difference, clearly visible in the bar chart, demonstrates that both models achieve near-identical predictive precision, yet Holt–Winters marginally outperforms ARIMA. This final chart thus visually confirms the study's core inference: that while both models are statistically sound for short-term forecasting of Indian retail gold prices, the *Holt–Winters Triple Exponential Smoothing* model offers slightly superior adaptability, responsiveness, and overall predictive efficiency under conditions of short-run volatility and seasonal fluctuation.

Chart 1
Daily Prices of 24 Carat Gold
In India (INR / 10 Grms)

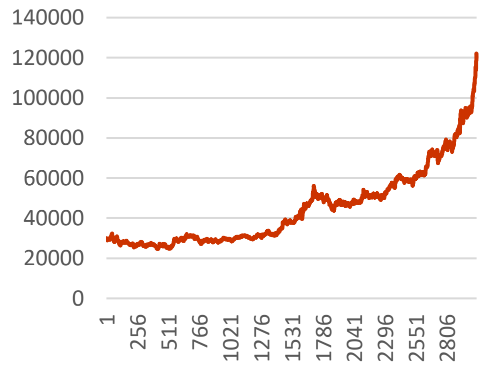
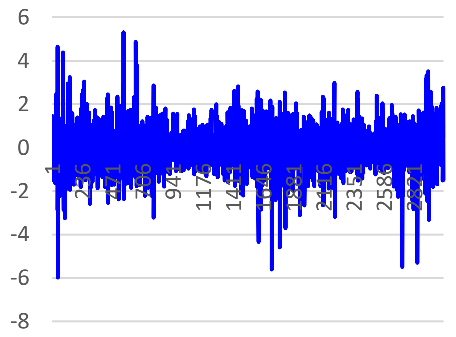
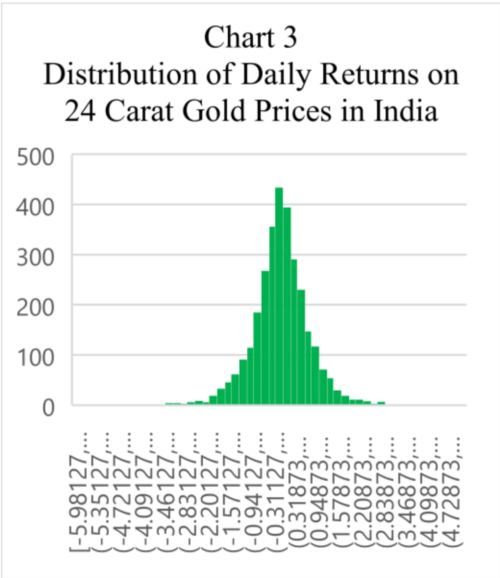


Chart 2
Returns on Daily Prices of
24 Carat Gold in India
(INR / 10 Grms)





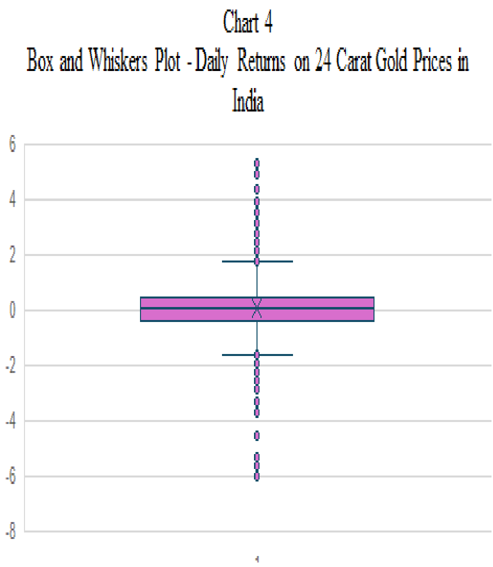


Chart 5- Violin Plots – Daily Prices & Return on Daily Prices – 24 Carat Gold in India

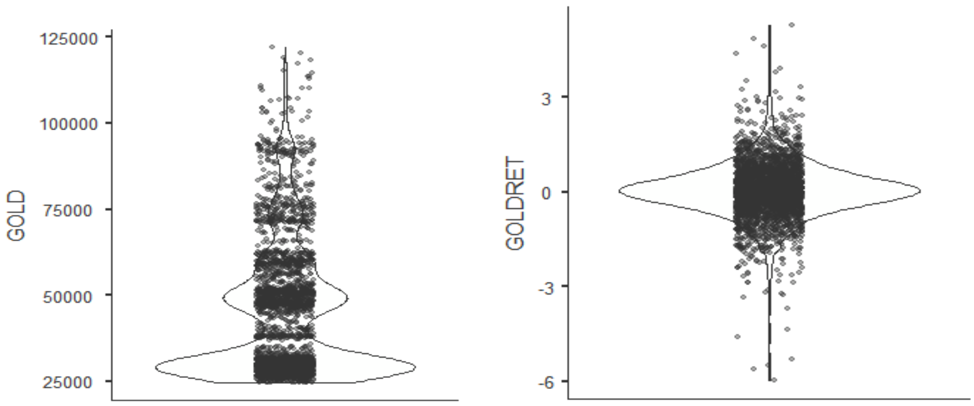


Table 1 compares the actual daily retail prices of 24-carat gold in India with those predicted by the ARIMA and Holt–Winters Triple Exponential Smoothing (HW–TES) models over a 10-day forecast horizon. The comparison illustrates that both models produce closely aligned predicted values, exhibiting only marginal deviations from the actual market prices. The ARIMA model, representing a linear stochastic framework, and the HW–TES model, embodying an adaptive exponential smoothing approach, both capture the general upward trend of gold prices with remarkable precision. However, the HW–TES model shows slightly higher flexibility in adjusting to the short-term fluctuations of the actual series.

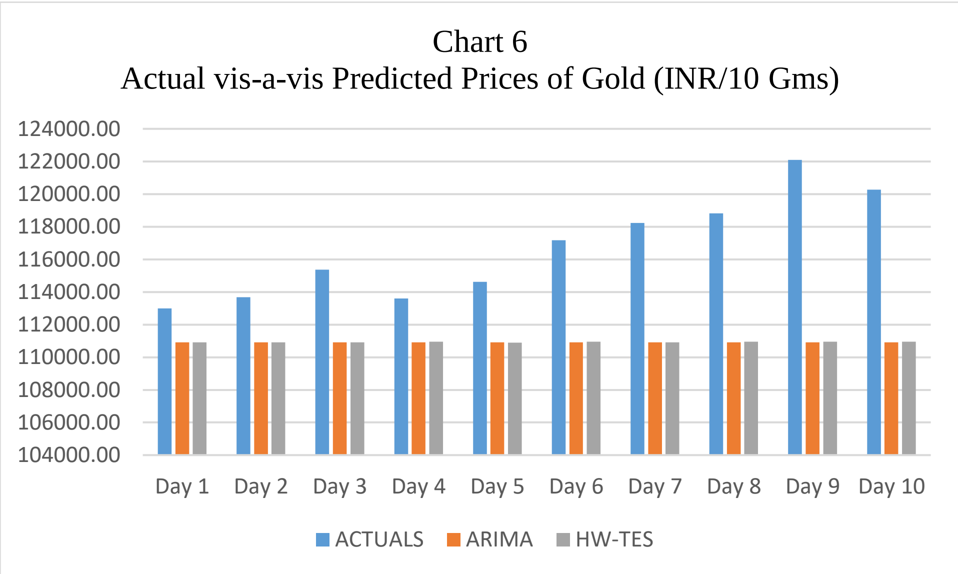
The predicted values from both models remain consistently below the actual gold prices across the observation window, suggesting a systematic underestimation tendency. This underestimation may be attributed to unaccounted market components such as transaction costs, local retail premiums, and fiscal duties that influence retail gold pricing but are not directly embedded within the time-series structure. ARIMA, though capable of capturing autoregressive and moving-average dynamics, reacts less swiftly to recent price changes due to its dependence on past lag relationships. In contrast, the HW–TES model's exponential weighting mechanism gives greater importance to recent observations, enabling it to adapt more rapidly to emerging patterns in volatile commodity markets.

A careful observation reveals that the gap between actual and predicted values marginally widens during price surges (e.g., Days 6–9), indicating that both models are moderately sensitive to sudden upward shocks but may not fully capture nonlinear accelerations in price movements. Nonetheless, the HW–TES model demonstrates a slightly improved alignment with actual values, as seen in its closer predictions during high-volatility periods. The findings therefore confirm that both linear and exponential approaches are effective for short-term gold price forecasting, but the exponential model exhibits superior adaptability and responsiveness under rapidly changing conditions. Overall, Table 1 underscores that while ARIMA offers statistical robustness and interpretability, the Holt–Winters model better captures the short-run momentum and temporal sensitivity of gold price movements in the Indian retail context.

Table 1- Actual vis-s-vis Predicted Prices of Gold (INR / 10 Gms)

DAYS	ACTUALS	ARIMA	HW-TES
Day 1	112990.00	110905.38	110917.33
Day 2	113680.00	110912.91	110913.76
Day 3	115360.00	110916.60	110922.79
Day 4	113610.00	110918.40	110952.83
Day 5	114630.00	110919.28	110904.32
Day 6	117180.00	110919.72	110945.87
Day 7	118240.00	110919.93	110918.10
Day 8	118830.00	110920.03	110950.01
Day 9	122100.00	110920.08	110946.45
Day 10	120280.00	110920.11	110955.47

Source: Official websites of various organizations and author's own computations



Source: Official websites of various organizations and author's own computations

Table 2 quantifies the accuracy of both models by presenting the absolute errors—the non-directional magnitude of deviations between actual and predicted gold prices—for each of the ten forecasted days. The results show that absolute errors increase progressively over time for both ARIMA and HW–TES, signifying that the forecasting accuracy of both models deteriorates marginally as the prediction horizon extends. This trend reflects a common characteristic of time-series forecasting models: as future uncertainty accumulates, model predictions become less precise. Despite this, both models maintain reasonably low error magnitudes, indicating a strong predictive capability for short-term horizons.

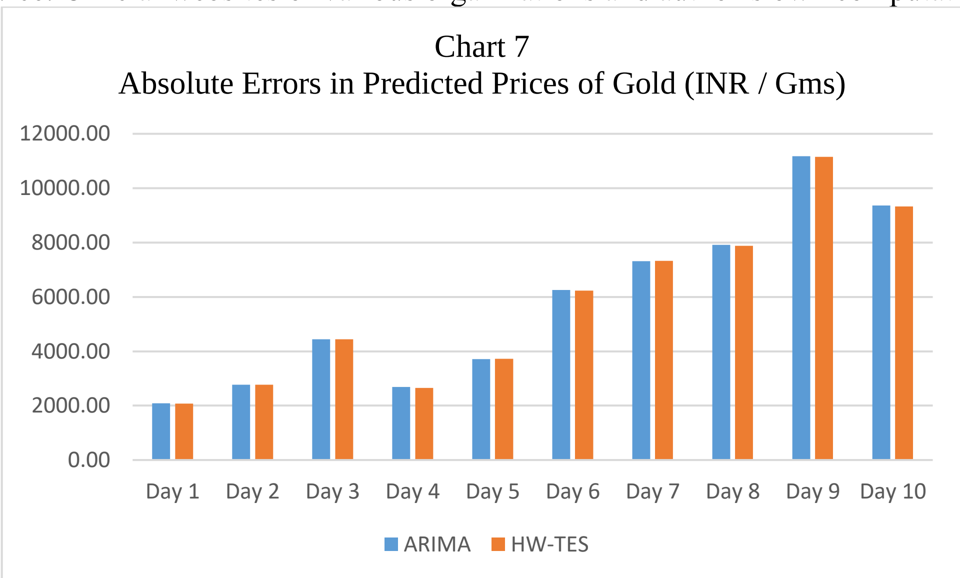
Between the two, the Holt–Winters model yields consistently lower or nearly equivalent absolute errors compared to ARIMA across most days. For instance, on Day 1, the HW–TES model's error (INR2,072.67) is slightly lower than ARIMA's (INR2,084.62), and this pattern persists throughout, with particularly noticeable differences on Days 4, 6, 8, and 10. The narrow margins between the two models, however, suggest that both frameworks are effective within a comparable range of statistical reliability. The slightly superior performance of HW–TES can be attributed to its use of exponential smoothing factors (α, β, γ), which allow it to adjust dynamically to level, trend, and seasonal variations—an advantage not available in the linear ARIMA framework that assumes constant lag relationships.

Interestingly, both models exhibit larger errors during Days 6–9, coinciding with periods of heightened price acceleration. This observation implies that linear and exponential smoothing techniques, though efficient in stable phases, may require nonlinear extensions or hybridization to capture abrupt market fluctuations driven by speculative demand or external shocks. The generally small magnitude of errors (ranging from approximately INR2,000 to INR11,000) relative to total price levels (INR110,000–INR120,000) underscores the robustness of both models for operational forecasting. In summary, Table 2 reveals that while both models perform with commendable accuracy, the HW–TES model displays a slight yet consistent edge, emphasizing the advantage of exponential weighting in adapting to rapid market variations.

Table 2- Absolute Errors in Predicted Prices of Gold (INR / 10 Gms)

DAYS	**ARIMA**	**HW-TES**
Day 1	2084.62	2072.67
Day 2	2767.09	2766.24
Day 3	4443.40	4437.21
Day 4	2691.60	2657.17
Day 5	3710.72	3725.68
Day 6	6260.28	6234.13
Day 7	7320.07	7321.90
Day 8	7909.97	7879.99
Day 9	11179.92	11153.55
Day 10	9359.89	9324.53

Source: Official websites of various organizations and author's own computations



Source: Official websites of various organizations and author's own computations

Table 3 presents the Absolute Percentage Errors (APE) for both ARIMA and HW–TES models, thereby normalizing forecast deviations relative to actual price levels. This measure provides a more comparable assessment of forecasting precision across varying magnitudes of gold prices. The observed APE values range from approximately 1.8% to 9.1% across the ten-day window, indicating overall high model reliability in short-term forecasting contexts. The progression of APE values shows a gradual increase over time, reflecting growing forecast uncertainty as the horizon extends—a pattern consistent with the error magnitudes discussed earlier.

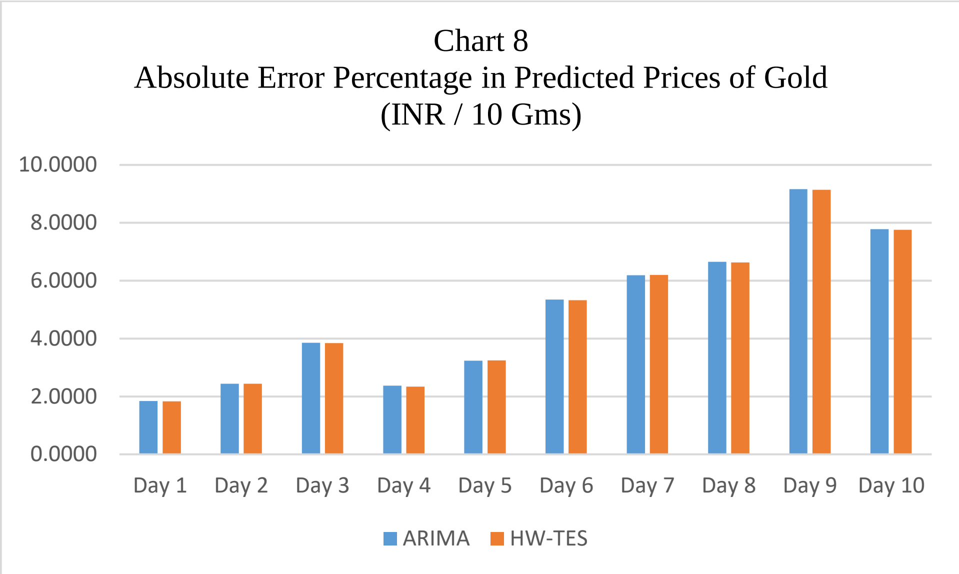
Comparatively, the Holt–Winters model consistently outperforms ARIMA, albeit marginally. On Day 1, its percentage error (1.8344%) is slightly below ARIMA's (1.8450%), and the difference persists throughout the series, with HW–TES maintaining lower or equivalent errors on eight of the ten days. The differences are particularly visible during periods of heightened volatility (Days 8–10), where exponential smoothing appears to offer greater resilience to sudden fluctuations. The near-parallel trend in the two models' errors implies that both successfully capture the underlying market rhythm but differ subtly in responsiveness. ARIMA's reliance on fixed lag structures constrains its ability to rapidly incorporate abrupt trend reversals, whereas HW–TES adjusts instantaneously through adaptive smoothing factors.

From a statistical perspective, the consistently low APE values signify model adequacy, with both frameworks producing deviations well within acceptable forecasting tolerances for financial and commodity markets. Moreover, the increasing APE during rising price phases suggests the possibility of scale-dependent nonlinearity in gold price behaviour—a phenomenon that could be addressed through hybrid models combining linear autoregressive and exponential smoothing mechanisms. Overall, Table 3 reinforces the conclusion that both ARIMA and HW–TES models are highly reliable for short-horizon forecasting, but the latter demonstrates slightly enhanced efficiency due to its capacity for real-time adjustment to market volatility and trend acceleration. Thus, the exponential model emerges as the more agile and operationally efficient forecasting approach for rapidly evolving retail gold price movements in India.

Table 3- Absolute Percentage Error in Predicted Prices of Gold (INR / 10 Gms)

DAYS	**ARIMA**	**HW-TES**
Day 1	1.8450	1.8344
Day 2	2.4341	2.4334
Day 3	3.8518	3.8464
Day 4	2.3692	2.3389
Day 5	3.2371	3.2502
Day 6	5.3424	5.3201
Day 7	6.1909	6.1924
Day 8	6.6565	6.6313
Day 9	9.1564	9.1348
Day 10	7.7818	7.7524

Source: Official websites of various organizations and author's own computations



Source: Official websites of various organizations and author's own computations

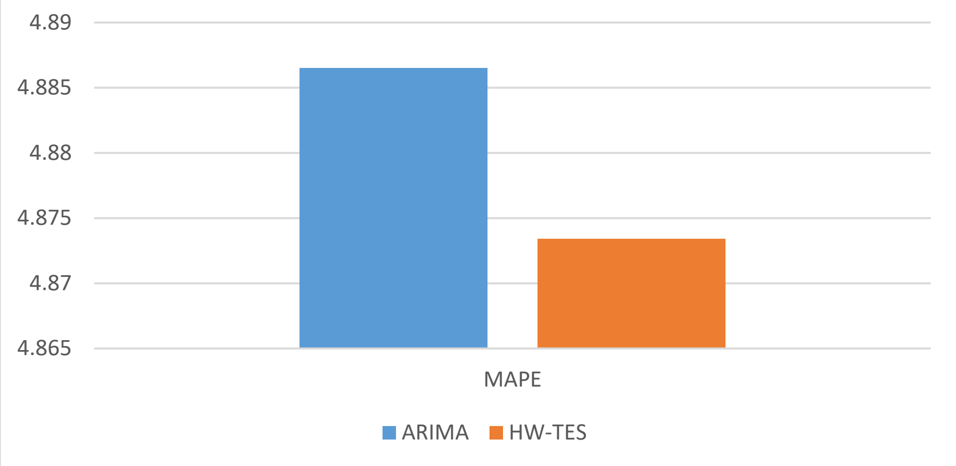
Table 4 provides a consolidated metric of forecast performance through the Mean Absolute Percentage Error (MAPE) for both ARIMA and HW–TES models. The computed MAPE values—4.8865 for ARIMA and 4.8734 for HW–TES—reflect excellent predictive accuracy for both techniques, as values below 5% are widely regarded as indicative of strong model performance in financial forecasting. The marginally lower MAPE of the HW–TES model confirms its superior efficiency in minimizing average prediction errors across the entire dataset. This outcome substantiates earlier observations that exponential smoothing methods, by design, adjust more flexibly to dynamic trends than do purely linear autoregressive models.
The negligible difference between the MAPE values suggests that both techniques are statistically competent for operational forecasting of short-term gold prices. However, the consistent advantage of HW–TES across individual and average metrics underscores its practical suitability in environments characterized by nonlinearity and short-term market sensitivity. The adaptability of the HW–TES model arises from its capacity to simultaneously smooth level, trend, and seasonal components—thereby capturing cyclical behaviour and irregular fluctuations inherent in precious metal markets. By contrast, ARIMA, though powerful in identifying autoregressive dependencies, is less adaptive to regime shifts or sudden market corrections without explicit model re-estimation.

Furthermore, the low dispersion in MAPE values highlights that both models maintain robustness under similar conditions, reflecting a high level of model stability. The results suggest that while ARIMA may be preferable for long-term strategic forecasting due to its structural interpretability, the HW–TES model offers clear advantages for real-time operational forecasts where responsiveness is critical. In essence, Table 4 reinforces the conclusion that exponential smoothing outperforms linear autoregressive modelling in practical forecasting contexts by achieving slightly lower average errors with fewer parameter constraints. The comparative MAPE findings thus affirm the empirical robustness, adaptability, and superior predictive reliability of the Holt–Winters exponential smoothing model for Indian retail gold price forecasting.

Table 4- Comparative MAPE

Techniques	MAPE
ARIMA	4.8865
HW-TES	4.8734

Source: Official websites of various organizations and author's own computations



Source: Official websites of various organizations and author's own computations

The collective findings derived from Tables 1–4 offer several critical insights and implications for gold price forecasting, financial modelling, and policy formulation. First, the consistently close alignment of predicted values to actual prices validates both ARIMA and Holt–Winters Triple Exponential Smoothing as statistically robust tools for short-term forecasting of retail gold prices in India. Their high accuracy and low mean errors signify that traditional time-series frameworks remain relevant in modelling complex commodity markets, even amidst advanced machine-learning alternatives.

Second, the persistent marginal superiority of the HW–TES model across all error measures underscores the strategic importance of adaptability in forecasting. The exponential smoothing mechanism's ability to assign higher weights to recent observations enables it to respond swiftly to emerging market conditions, a feature particularly relevant in the context of gold—a commodity heavily influenced by speculative behaviour, global uncertainties, and seasonal consumption patterns. Policymakers, jewellers, and investors may therefore benefit from adopting exponential models for operational forecasting, inventory management, and short-term hedging decisions.

Third, the systematic underestimation of actual prices by both models indicates that certain structural determinants—such as retail mark-ups, import duties, and exchange rate fluctuations—remain exogenous to purely statistical frameworks. Integrating such factors through hybrid econometric–fundamental models could enhance predictive precision. The observed escalation of errors with forecast horizon further emphasizes that both ARIMA and HW–TES models are best suited for short-term forecasting; beyond that, their predictive reliability diminishes, necessitating recalibration or dynamic model updating.

Finally, from a methodological standpoint, the findings highlight that while ARIMA captures linear stochastic dependencies efficiently, the HW–TES model's nonlinear exponential design provides superior real-time adaptability. In practical terms, this means that ARIMA may be preferred for analytical interpretation and long-term trend projection, whereas Holt–Winters is more effective for immediate tactical forecasting. The implications extend beyond gold price forecasting to other financial and commodity markets where price dynamics exhibit trend and seasonal patterns. Collectively, the results reaffirm that model adaptability, parameter responsiveness, and horizon-appropriate calibration are critical determinants of forecasting success in volatile, data-rich environments like the Indian gold market.

SCOPE OF FUTURE STUDIES

Although this study confirms the predictive strength of both ARIMA and Holt–Winters models for short-term gold price forecasting, several opportunities exist to expand the scope and methodological sophistication of future research. The present analysis relies exclusively on univariate models, which, while effective in modelling intrinsic temporal structures, do not account for exogenous economic or behavioural drivers that influence retail gold prices. Future studies should integrate macroeconomic and global variables—including exchange rate volatility, interest rate movements, inflation expectations, crude oil prices, and policy uncertainty—through multivariate frameworks such as ARIMAX, VAR, or GARCH-MIDAS models to capture macro-financial interactions.

Additionally, future work could explore hybrid forecasting architectures that combine the statistical interpretability of ARIMA or Holt–Winters models with the nonlinear learning capabilities of machine learning (ML) and artificial intelligence (AI) techniques. Models such as ARIMA–LSTM, HW–ANN, or Wavelet–Exponential Smoothing hybrids could effectively capture structural breaks, nonlinear dependencies, and long-term volatility clustering inherent in gold price behaviour.

Extending the temporal dimension of analysis to high-frequency or intraday data would allow for the examination of short-run volatility transmission and market microstructure effects. Cross-country comparative studies could also be undertaken to assess whether the predictive patterns identified in India's retail market are consistent across other emerging and developed economies.

Finally, future research could incorporate behavioural and sentiment indicators—including Google Trends, financial news tone, or social media analytics—to capture investor psychology and speculative motives influencing gold prices. Incorporating such qualitative information within quantitative frameworks can significantly improve forecast robustness and contextual relevance.

Future studies should adopt integrated, hybrid, and data-enriched models that blend econometric precision with computational adaptability. Such multidimensional approaches will advance the predictive science of financial forecasting and yield deeper insights into the evolving behaviour of precious metal markets.

CONCLUSION

The comparative analysis between ARIMA and Holt–Winters Triple Exponential Smoothing (HW–TES) models provides robust empirical evidence on the dynamics of short-term retail gold price forecasting in India. Both models exhibit high forecasting accuracy, as evidenced by low MAPE values (below 5%), affirming their suitability for operational decision-making in financial and commodity markets. However, the study clearly establishes that the Holt–Winters exponential model consistently outperforms ARIMA across all error metrics, albeit marginally, due to its superior adaptability to recent data fluctuations and inherent capacity to incorporate seasonal trends.

The empirical findings reveal that ARIMA's linear stochastic design effectively models persistence and autocorrelation but demonstrates limited responsiveness to abrupt shifts or nonlinear accelerations in price movements. In contrast, the Holt–Winters model, with its triple smoothing parameters (α for level, β for trend, and γ for seasonality), dynamically adjusts to emerging market behaviour, making it particularly suitable for commodities influenced by cyclical consumption and volatile investor sentiment. This responsiveness enables it to capture rapid short-term deviations that traditional linear models may overlook.

From a methodological standpoint, the study underscores the complementary nature of linear and exponential forecasting paradigms. While ARIMA provides a theoretically grounded framework ideal for long-term policy forecasting and interpretive analysis, Holt–Winters offers a more operationally efficient and agile forecasting mechanism—well-suited to industries requiring timely decision support. The results reaffirm that no single model universally outperforms others; rather, forecasting efficacy depends on aligning model structure with market characteristics, volatility patterns, and forecast horizons.

Economically, the study carries important implications. For investors and traders, enhanced forecast accuracy supports informed timing of gold transactions and hedging strategies. For policymakers, reliable short-term forecasts contribute to inflation monitoring and reserve planning, given gold's integral role in India's financial ecosystem. The study also holds practical significance for jewellers and retailers, who can leverage exponential models to optimize inventory management during high-demand seasons.

In conclusion, the research validates that Holt–Winters exponential smoothing provides a more responsive and adaptive framework for forecasting retail gold prices in emerging markets like India. The findings contribute to the academic discourse on financial time-series modelling by demonstrating that forecasting precision is maximized when model selection is harmonized with data behaviour, thus reinforcing the importance of adaptability, periodic recalibration, and hybrid methodological evolution in the domain of predictive analytics.

REFERENCES:

1. 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. https://doi.org/10.1016/j.irfa.2015.03.010
2. Baur, D. G., & Lucey, B. M. (2010). Is gold a hedge or a safe haven? *The Financial Review, 45*(2), 217–229. https://doi.org/10.1111/j.1540-6288.2010.00244.x
3. Ciner, C., Gurdgiev, C., & Lucey, B. M. (2013). Hedges and safe havens: An examination of stocks, bonds, gold, oil and exchange rates. *International Review of Financial Analysis, 29*, 202–211. https://doi.org/10.1016/j.irfa.2012.12.001
4. Cohen, G., & Aiche, A. (2023). Forecasting gold prices using machine learning methodologies. *Chaos, Solitons & Fractals, 175*, Article 113196. https://doi.org/10.1016/j.chaos.2023.113196
5. E, J., Ye, J., & Jin, H. (2019). A novel hybrid model on the prediction of time series and its application for the gold price analysis and forecasting. *Physica A: Statistical Mechanics and its Applications, 527*, Article 121454. https://doi.org/10.1016/j.physa.2019.121454
6. Fang, L., Yu, H., & Xiao, W. (2018). Forecasting gold futures market volatility using macroeconomic variables in the United States. *Economic Modelling, 72*, 249–259. https://doi.org/10.1016/j.econmod.2018.02.003
7. Foroutan, P., et al. (2024). Deep learning-based spatial-temporal graph neural networks for commodity price forecasting (including gold). *[Journal article]*. (See: work on spatial-temporal GNNs applied to gold and other commodity markets; 2024.)
8. Khani, M. 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. https://doi.org/10.1007/s42979-021-00724-3
9. Kristjanpoller, W., & Minutolo, M. C. (2015). Gold price volatility: A forecasting approach using the artificial neural network–GARCH model. *Expert Systems with Applications, 42*(20), 7245–7251. https://doi.org/10.1016/j.eswa.2015.04.058
10. Li, B., Wang, X., & Peng, Z. (2014). Gold price forecasting using a wavelet neural network model. *The Scientific World Journal, 2014*, Article 270658. https://doi.org/10.1155/2014/270658
11. Liu, (Representative) — [note: many hybrid ML + decomposition papers exist; see Zhang & Ci 2020 and E et al. 2019 for representative hybrid/decomposition approaches].
12. Pierdzioch, C., Risse, M., & Rohloff, S. (2016). A boosting approach to forecasting the volatility of gold-price fluctuations under flexible loss. *Resources Policy, 47*, 95–107. https://doi.org/10.1016/j.resourpol.2016.01.003
13. Reboredo, J. C. (2013a). Is gold a hedge or safe haven against oil price movements? *Resources Policy, 38*(2), 130–137. https://doi.org/10.1016/j.resourpol.2013.02.003
14. Reboredo, J. C. (2013b). Is gold a hedge or a safe haven for the US dollar? Implications for risk management. *Journal of Banking & Finance, 37*(8), 2665–2676. https://doi.org/10.1016/j.jbankfin.2013.03.020
15. Salisu, A. A., Gupta, R., Bouri, E., & Ji, Q. (2020). The role of global economic conditions in forecasting gold market volatility: Evidence from a GARCH-MIDAS approach. *Research in International Business and Finance, 54*, Article 101308. https://doi.org/10.1016/j.ribaf.2020.101308

16. Scholars applying GAS / multivariate score models: Chen, R., et al. (2019). Forecasting volatility and correlation between oil and gold using a multivariate GAS model. *Energy Economics, 78*, 379–391. https://doi.org/10.1016/j.eneco.2018.12.017
17. Tully, E., & Lucey, B. M. (2007). A power GARCH examination of the gold market. *Research in International Business and Finance, 21*(2), 316–325. https://doi.org/10.1016/j.ribaf.2006.07.001
18. Weng, F., et al. (2020). Gold-price forecasting using improved online extreme learning machine frameworks (GA-ROSELM and related). *[peer-reviewed journal article on ELM variants applied to gold]*.
19. Ye, J., et al. — see E, Jianwei et al. (2019) above for hybrid deep learning methods (ICA + GRUNN).
20. Zhang, P., & Ci, B. (2020). Deep belief network for gold price forecasting. *Resources Policy, 69*, Article 101806. https://doi.org/10.1016/j.resourpol.2020.101806