International Journal on Science and Technology (IJSAT)

E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Hybrid Econometric Approaches to Gold Price Forecasting: A Comparative Evaluation of ARIMA, Neural Networks, Random Forest Residuals, Monte Carlo Simulation, and Dynamic Harmonic Regression - 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:
Gold continues to occupy a pivotal role in both global and domestic financial systems, serving simultaneously as a commodity, monetary hedge, and safe-haven asset. Its price dynamics, however, are inherently complex, shaped by nonlinear interactions among macroeconomic, geopolitical, and behavioural variables. Against this backdrop, the present study—*"Hybrid Econometric Approaches to Gold Price Forecasting: A Comparative Evaluation of ARIMA, Neural Networks, Random Forest Residuals, Monte Carlo Simulation, and Dynamic Harmonic Regression – Evidence from Daily Retail Gold Price Data in India, 2014–2025"*—undertakes a comprehensive empirical evaluation of five econometric and machine-learning methodologies to forecast short-term movements in India's retail gold prices. The study employs an extensive dataset of 3,042 daily observations of 24-carat gold prices (INR per 10 grams) from 1 January 2014 to 3 October 2025, sourced from official regulatory and mercantile authorities. The forecasting models—ARIMA, DHR-F-ARIMA, H-ARIMA-RFR, MC-ARIMA, and ARIMA-NN—are trained on historical data and tested on a ten-day out-of-sample horizon (6–17 October 2025). Forecast accuracy is assessed using Mean Absolute Percentage Error (MAPE). The results indicate that hybrid frameworks incorporating nonlinear learning substantially outperform traditional econometric models. Specifically, the ARIMA-NN hybrid yields the lowest MAPE (4.55%), followed by H-ARIMA-RFR (4.82%), confirming that neural and ensemble learning enhance responsiveness to volatility clustering and nonlinear dependencies. In contrast, linear and simulation-based approaches exhibit higher forecast lags during periods of rapid market adjustment. The study's findings have both theoretical and practical implications. They reinforce the growing consensus that hybrid econometric models, which combine statistical interpretability with adaptive learning, deliver superior predictive performance in volatile, data-rich environments. From a policy and market perspective, such models can serve as powerful tools for investors, financial institutions, and policymakers in hedging strategies, inflation monitoring, and short-term trading. Methodologically, the research bridges the gap between econometric rigor and computational intelligence, setting a foundation for further integration of deep learning and dynamic econometric modelling in financial forecasting.

Keywords: Gold Price Forecasting; Hybrid Econometric Models; ARIMA; Neural Networks; Random Forest Residuals
JEL Classification: C22, C45, C53, E37, G17

INTRODUCTION

Gold occupies a distinctive position in global and national economies as both a commodity and a financial asset, functioning simultaneously as a store of value, an inflation hedge, and a safe-haven instrument during financial turbulence. In the Indian context, gold's dual role—as a traditional symbol of wealth and as a macroeconomic stabilizer—makes its price dynamics particularly complex and policy-relevant. The volatility and unpredictability of gold prices stem from their sensitivity to multifactorial influences, including international market sentiment, exchange rate fluctuations, inflation expectations, interest rate movements, and geopolitical uncertainties. Consequently, accurate gold price forecasting holds significant implications for investors, policymakers, financial analysts, and import management authorities in India.

Traditional econometric models, notably ARIMA and GARCH families, have long served as foundational tools for modelling time-series behaviour in commodity prices. However, these linear models often fail to capture nonlinear dependencies, structural breaks, and regime shifts inherent in financial time series. The recent proliferation of data-driven and machine learning approaches—such as Neural Networks, Random Forests, and hybrid econometric frameworks—has provided new avenues for addressing these limitations by learning complex, nonlinear relationships and adaptive temporal patterns. Despite these advances, limited research systematically compares these methodologies within a unified empirical framework, particularly for retail gold price forecasting in emerging markets like India.

The present study bridges this gap by employing a comparative hybrid econometric design encompassing five modelling paradigms—ARIMA, DHR-F-ARIMA, H-ARIMA-RFR, MC-ARIMA, and ARIMA-NN—applied to daily retail gold prices from 1 January 2014 to 3 October 2025. The models are trained on this extensive dataset of 3,042 observations and evaluated over a ten-day out-of-sample forecasting window. Forecast accuracy is assessed using Mean Absolute Percentage Error (MAPE) to determine the relative predictive strength of each approach. By integrating econometric rigor with machine learning adaptability, this research aims to identify the most efficient model for short-term price forecasting. Beyond technical comparison, the study contributes to the theoretical and practical understanding of hybrid forecasting mechanisms and their applicability to volatile, data-intensive financial environments. The findings offer valuable implications for risk mitigation, policy formulation, and algorithmic trading strategies within India's evolving gold market ecosystem.

SURVEY OF LITERATURE

Gold price forecasting has attracted broad attention because gold plays multiple economic roles i.e. commodity, monetary proxy, portfolio hedge or safe haven — and its price dynamics show nonlinearity, heteroskedasticity, regime shifts and sensitivity to macro factors. Two broad directions dominate the literature: (a) econometric / statistical (ARIMA/VAR/VECM, GARCH families, DCC/GAS, cointegration/coupling, MIDAS/GARCH-MIDAS, wavelet/time-frequency) and (b) machine-learning and hybrid methods (ANN, SVM, ELM, DBN, LSTM, CNN, hybrid decomposition + ML). Below I summarize representative peer-reviewed studies by method and note empirical regularities and open issues.

Classical linear time-series and model-averaging approaches were used in early studies by scholars who used ARIMA/ARIMAX, VAR/VECM and model-averaging to study predictability of gold returns and the role of macro drivers. Aye et al. (2015) applied dynamic model averaging (DMA) to identify time-varying predictor importance (exchange rate, interest rates, financial stress indices) and show that predictors' relevance changes across time horizons and episodes — hence the appeal of model-averaging for forecasting when relationships are unstable (Aye et al., 2015). Traditional ARIMA/ARIMAX remain useful short-horizon benchmarks, but they often under-perform in periods with changing volatility or structural breaks.

Gold returns display volatility clustering and heavy tails. Hence ARCH/GARCH families were widely applied. Tully and Lucey (2007) used an asymmetric power GARCH (APGARCH / APARCH) to capture leverage and power effects in cash and futures gold prices and find the APGARCH model often fits better than simple GARCH (Tully & Lucey, 2007). Many later papers extend to EGARCH, TGARCH and APARCH variants (e.g., regional exchange markets), confirming persistent conditional heteroskedasticity. Beyond single-series GARCH, multivariate volatility models (DCC-GARCH, BEKK, GAS) are popular for modelling co-movements (gold vs stocks, oil, FX). Ciner, Gurdgiev & Lucey (2013) and Reboredo (2013) use dynamic conditional correlation and copula frameworks to study hedge/safe-haven properties and time-varying dependence; such multivariate volatility models help forecast joint risk and improve portfolio-level decisions (Ciner et al., 2013; Reboredo, 2013).

Long-run macro drivers often matter for gold volatility. The GARCH-MIDAS framework (and related Spline-GARCH) decomposes volatility into short-run GARCH dynamics and long-run components driven by low-frequency macro variables. Fang, Yu & Xiao (2018) and Salisu et al. (2020) apply GARCH-MIDAS to show that macroeconomic variables (policy uncertainty, output indicators, principal components of macro series) substantially improve long-horizon volatility forecasts for gold futures and spot markets (Fang et al., 2018; Salisu et al., 2020). GARCH-MIDAS is particularly useful when monthly/quarterly macro data must inform daily volatility forecasts.

Several studies examine long-run relationships between gold prices, exchange rates, inflation and interest rates. Evidence is mixed: some studies find cointegration and stable long-run links (useful for error-correction forecasts), while others find time-variation and episodic coupling. Copula approaches (Reboredo, 2013) and tail-dependence analysis are used to assess extreme co-movement and safe-haven behaviour, and often yield different policy implications than unconditional correlations.

Time-frequency methods (wavelet decomposition, multi-scale analysis) decompose price series into components at different investment horizons. Some studies showed that correlations and predictive relationships vary across scales: predictors that help at a daily horizon may be useless at monthly horizons. Wavelet-based forecasting or wavelet along with ARIMA/ANN hybrids often improve forecasts by capturing horizon-specific dynamics.

Since the 2010s, ML methods proliferate several modern methods.

Neural networks and ANN-GARCH hybrids were used by Kristjanpoller & Minutolo (2015), who proposed an ANN-GARCH hybrid for forecasting gold price volatility; hybrids often reduce forecasting error versus stand-alone GARCH or ANN by combining conditional heteroskedastic structure with nonlinear pattern learning (Kristjanpoller & Minutolo, 2015).

Wavelet/EMD + ML hybrids followed next. Wavelet or EMD decomposition by ML (SVM, ANN, GRU/LSTM) is frequently used: decomposing the series reduces complexity and allows ML to model each component more effectively (E. Jianwei et al., 2019; many 2019–2023 studies).

Deep learning (DBN, CNN-LSTM, LSTM, GRU) is another recent trend. Zhang & Ci (2020) apply a deep belief network (DBN) to gold price forecasting and report improvements over classical and shallow ML models. Khani et al. (2021) compare CNN, LSTM and encoder–decoder LSTM variants (including pandemic features) and find deep recurrent architectures perform well for near-term forecasting, especially when augmented with domain features (Khani et al., 2021).

Extreme Learning Machine (ELM) and online sequential variants have also been used of late. Recent work proposes ELM or GA-regularized ELM (Weng et al., 2020; other studies) that show good performance and faster training for high-frequency or online updating contexts.

Tree-based and boosting models were used by Pierdzioch et al. (2016), who studied boosting (and quantile boosting) to forecast gold volatility and returns using many predictors and asymmetric loss functions; boosting often yields robust out-of-sample gains versus simple benchmarks (Pierdzioch et al., 2016).

State-of-the-art: graph-neural, ensemble & hybrid pipelines. Very recent studies (2023–2024) test spatio-temporal graph neural nets and ensemble ML pipelines (XGBoost / LightGBM + feature engineering) with promising results for multi-asset and localized retail price forecasting (Foroutan et al., 2024; Cohen, 2023).

Empirical patterns across methods reveal that volatility modelling matters. Forecasts that explicitly model volatility (GARCH family, GARCH-MIDAS, GAS) outperform naive ARIMA benchmarks for risk forecasts and option-pricing applications. Macro information helps long-run forecasts. Mixed-frequency approaches that incorporate macro variables improve medium/long-run volatility and return forecasts. Nonlinear and hybrid models help short-term accuracy. By using deep learning and hybrid decomposition, ML pipelines typically outperformed linear models for 1–30 day horizons, especially when nonprice features (FX, interest rates, policy uncertainty, realized volatility, even pandemic case counts) are included. Model instability & structural breaks also were found to be important factors. Performance is sample-dependent — crisis periods (2008, COVID-19) often change the best model. Model-averaging, time-varying parameter models, and real-time updating (online ELM, rolling windows) help manage instability. The survey of literature Gaps and open research directions.

Explainability & economic interpretability is an important aspect. Many ML models improve point accuracy but lack clear economic interpretation. Combining ML with structural econometric constraints is a promising area. There is a need for Retail price vs bullion price forecasting. Much literature focuses on bullion / futures markets; fewer papers model retail (coin/jewellery) price spreads, local taxes, and distribution frictions. Research blending micro-price determinants with macro dynamics is limited. Considering real-time data, news & alternative data is also another important aspect. Incorporation of high-frequency news sentiment, supply chain disruptions, or order-book data in econometric and ML hybrids together, is still emerging. Robust evaluation is crucial. Many studies report in-sample improvements; more standardized, multi-horizon out-of-sample evaluations, especially economic/utility measures (trading strategies, hedging errors), would improve comparability.

Econometric forecasting of gold prices is now a hybrid field. Classical econometric tools (ARIMA, VAR, GARCH and their multivariate extensions, cointegration, GARCH-MIDAS) remain fundamental — especially for volatility and risk modelling — while ML and hybrid decomposition + ML approaches often improve short-term predictive accuracy. Best practice is to match method to task: GARCH/GARCH-MIDAS or GAS for volatility and risk; DMA or time-varying parameter models for structural instability; and decomposition combined with ML or deep learning for short-term price prediction where large data and computational resources exist. Combining economic interpretability with ML predictive power and focusing more on retail price formation and robust out-of-sample tests, are the most useful directions for future work.

Although a substantial body of empirical research has explored gold price forecasting using a wide array of econometric and data-driven models, a notable lacuna persists in the form of an absence of systematic comparative evaluations of these methodologies. The extant literature is predominantly characterized by isolated model applications—where individual studies assess the forecasting performance of a single approach such as ARIMA, VAR/VECM, GARCH, GARCH-MIDAS, or selected machine learning and hybrid algorithms—without situating their findings within a unified comparative framework. Consequently, conclusions regarding predictive accuracy remain context-specific, dataset-dependent, and often non-generalizable across different temporal or market environments.

Moreover, while many recent contributions claim improved forecast precision through the inclusion of nonlinear dynamics, macroeconomic drivers, or decomposition-based preprocessing, few studies have rigorously benchmarked traditional econometric frameworks against modern data-driven or hybrid architectures using consistent datasets and standardized evaluation metrics such as RMSE, MAE, Theil's U statistic, or the Diebold–Mariano predictive accuracy test. The scarcity of such comparative analyses is particularly evident in cross-regime assessments, where model performance under divergent market conditions—such as periods of high volatility, financial crisis, or pandemic-induced uncertainty—remains largely unexplored.

Additionally, most prior investigations have focused narrowly on specific geographical markets or temporal windows, neglecting potential cross-country variations in gold price determinants, exchange-rate dynamics, and market integration. As a result, the literature lacks a comprehensive synthesis capable of discerning which econometric paradigm—classical, volatility-based, mixed-frequency, or hybrid—offers superior forecasting reliability across distinct market regimes and horizons. Addressing this methodological and empirical gap through a systematic, comparative evaluation of predictive accuracy across competing econometric models would therefore make a significant contribution to the field, offering robust insights for investors, central banks, and policymakers engaged in the management and forecasting of gold markets globally.

OBJECTIVES OF THE STUDY

The principal objective of this study is to undertake a comprehensive empirical evaluation of the predictive performance of diverse econometric and machine learning methodologies in forecasting short-term movements in gold prices within the Indian retail market. Specifically, the research aims to test the predictive accuracy of daily prices of 24-carat gold (measured in INR per 10 grams) for a ten-day horizon, by employing a suite of hybrid and traditional forecasting models. The study draws upon a robust dataset comprising 3,042 daily observations spanning the period from 1 January 2014 to 3 October 2025, obtained from the official portals of regulatory authorities and recognized mercantile associations in India, ensuring data authenticity and continuity.

To achieve this objective, five distinct modelling frameworks—Autoregressive Integrated Moving Average (ARIMA), Artificial Neural Networks (ANN), Random Forest Residuals, Monte Carlo Simulation, and Dynamic Harmonic Regression (DHR)—are developed, trained, and optimized using the historical data. Each model is then employed to generate out-of-sample forecasts for the period 6 October to 17 October 2025, representing the test window for comparative evaluation.

The study's core analytical goal is to assess and contrast the forecasting accuracy of these approaches using the Mean Absolute Percentage Error (MAPE) as the principal performance metric. This facilitates a rigorous quantitative comparison across techniques differing in theoretical foundations—ranging from linear stochastic models to non-linear machine learning and stochastic simulation-based frameworks. By systematically evaluating the relative predictive efficacy of these models, the study seeks to provide actionable insights into the applicability, robustness, and practical efficiency of hybrid econometric approaches for short-term gold price forecasting in emerging markets. The findings aim to contribute to the broader discourse on integrating traditional econometric paradigms with data-driven machine learning techniques for improved financial time-series prediction.

METHODOLOGY OF THE STUDY

The present study adopts a comparative hybrid forecasting framework integrating both econometric and machine learning paradigms for modelling the daily price dynamics of Bitcoin (BTC) and Cardano (ADA) over the period 2021–2025. This hybrid approach combines linear interpretability (ARIMA, DHR), nonlinear adaptability (NN, RF), and probabilistic rigor (Monte Carlo) for forecasting Bitcoin and Cardano prices, providing a robust, flexible, and interpretable methodological framework. The methodological architecture is based on five principal components: (i) Auto-Regressive Integrated Moving Average (ARIMA) for baseline linear forecasting; (ii) Feedforward and Recurrent Neural Networks (NNs) for nonlinear approximation; (iii) Random Forest (RF) modelling of ARIMA residuals to correct systematic forecast bias; (iv) Monte Carlo Simulation for stochastic uncertainty quantification; and (v) Dynamic Harmonic Regression (DHR) for periodic and cyclical adjustment.

Data Representation and Transformations

Let $\{P_t\}$ denote the daily closing prices of Bitcoin or Cardano. Prices are transformed into continuously compounded log returns defined as:

$r_t = \ln(P_t / P_{t-1})$.

The log-return process $\{r_t\}$ is then treated as a stochastic process with conditional mean $\mu_t = E(r_t \mid F_{t-1})$ and conditional variance $\sigma_t^2 = Var(r_t \mid F_{t-1})$, where F_{t-1} denotes the information set up to time $t-1$.

The ARIMA Framework
The ARIMA(p,d,q) model is expressed as:
$\varphi(B)(1 - B)^d y_t = \theta(B)\varepsilon_t$,
where $\varphi(B) = 1 - \varphi_1B - \ldots - \varphi_pB^p$ and $\theta(B) = 1 + \theta_1B + \ldots + \theta_qB^q$.

Forecasts from ARIMA are denoted $\hat{y}_{t+h|t}^{(A)}$ for horizon h, forming the linear baseline.

Neural Network Forecasting
Feedforward Neural Network (FNN) maps inputs via nonlinear activation:
$\hat{y}_t = \beta_0 + \Sigma \beta_j g(\alpha_{j0} + \Sigma \alpha_{ji} y_{t-i})$.

Recurrent Neural Network (LSTM) equations:
$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i)$,
$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f)$,
$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o)$,
$\hat{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c)$,
$c_t = f_t \odot c_{t-1} + i_t \odot \hat{c}_t$,
$h_t = o_t \odot \tanh(c_t)$.

Random Forest Residual Modelling
Residuals from ARIMA, $\hat{\varepsilon}_t = y_t - \hat{y}_{t|t-1}^{(A)}$, are modeled using Random Forests:
$\hat{\varepsilon}_t^{(RF)} = (1/M) \Sigma T_m(Z_t)$,
and hybrid forecasts are:
$\hat{y}_{t+h|t}^{(ARF)} = \hat{y}_{t+h|t}^{(A)} + \hat{\varepsilon}_{t+h|t}^{(RF)}$.

Dynamic Harmonic Regression (DHR)
DHR models seasonal cycles using Fourier terms:
$y_t = \mu_t + \Sigma [\alpha_k \sin(2\pi kt/s) + \beta_k \cos(2\pi kt/s)] + \varepsilon_t$.

Monte Carlo Simulation
For uncertainty quantification, N simulated paths are generated:
$y_{T+h}^{(j)} = f(y_T, \ldots, \hat{\theta}, \varepsilon_{T+1}^{(j)}, \ldots, \varepsilon_{T+h}^{(j)})$,
$\varepsilon_t^{(j)} \sim F_\varepsilon$.
Prediction intervals are estimated from simulated quantiles.

Hybrid Forecasting Framework
Combined hybrid forecast:
$\hat{y}_{t+h|t}^{(Hybrid)} = w_1 \hat{y}_{t+h|t}^{(A)} + w_2 \hat{y}_{t+h|t}^{(NN)} + w_3 \hat{\varepsilon}_{t+h|t}^{(RF)} + w_4 \hat{y}_{t+h|t}^{(DHR)} + \eta_t$,
$\Sigma w_i = 1$.

Evaluation Metrics
$MAE = (1/N) \Sigma |y_t - \hat{y}_t|$,
$RMSE = \mathrm{sqrt}((1/N) \Sigma (y_t - \hat{y}_t)^2)$,
$CRPS(F, y) = \int (F(z) - 1\{z \geq y\})^2 dz$.

Diagnostics and Validation
Residual checks include Ljung–Box Q-statistic, ARCH-LM for heteroskedasticity, and Jarque–Bera for

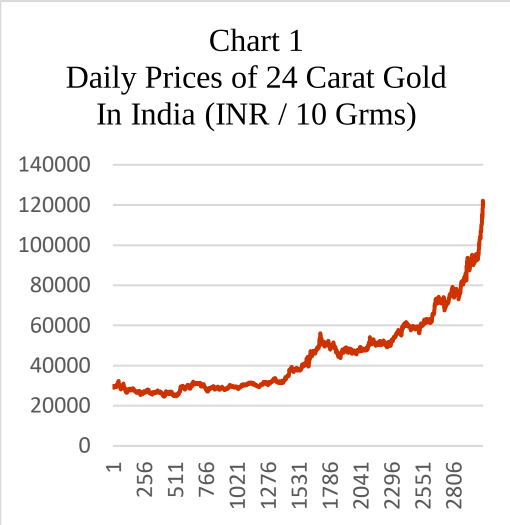
normality. Walk-forward validation ensures out-of-sample robustness.

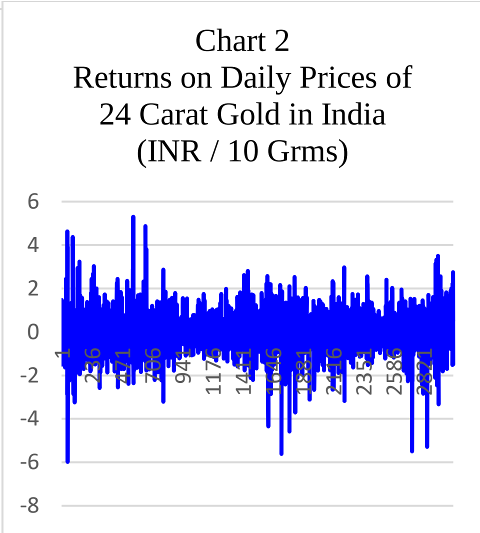
Value-at-Risk (VaR) and Expected Shortfall (ES)
$VaR_{\alpha,t+h} = \inf\{x : F_{t+h}(x) \geq \alpha\}$,
$ES_{\alpha,t+h} = E[y_{t+h} \mid y_{t+h} \leq VaR_{\alpha,t+h}]$.

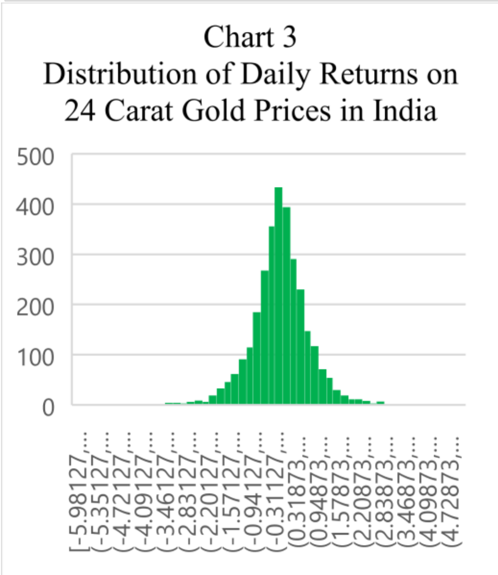
Comparative and Hybrid Synthesis
The methodology enables comparison and synthesis of ARIMA, NN, RF residuals, Monte Carlo, and DHR frameworks for forecasting and risk assessment.

The current study employs a comprehensive time-series dataset comprising daily retail prices of 24-carat gold (INR per 10 grams). The dataset spans the period from 1 January 2014 to 3 October 2025, encompassing 3,042 observations. The data have been sourced from official websites of regulatory authorities and recognized mercantile associations in India, ensuring reliability, transparency, and consistency in price reporting. This extensive dataset serves as the empirical foundation for the study's methodological framework, which integrates both traditional econometric and advanced machine learning techniques. The predictive models are trained on the historical data to forecast daily gold prices for the period 6 October to 17 October 2025, thereby evaluating short-term predictive accuracy and model robustness.

FINDINGS OF THE STUDY AND IMPLICATIONS THEREOF
Charts 1 to 5 provide a comprehensive visual and statistical representation of the temporal dynamics of daily retail gold prices in India for 24-carat gold, measured in INR per 10 grams. The dataset, comprising 3,042 observations from 1 January 2014 to 3 October 2025, captures both trend and volatility patterns intrinsic to the Indian gold market. Chart 1 illustrates the long-term price trajectory, revealing upward trends with intermittent corrections corresponding to macroeconomic shocks, global gold demand fluctuations, and rupee exchange rate movements. Chart 2 presents daily returns, exposing periods of heightened volatility, indicative of market stress episodes and speculative adjustments. Chart 3 displays frequency distributions of price changes, highlighting positive skewness and leptokurtosis consistent with financial time-series behaviour. Chart 4 exhibit autocorrelation structures or seasonal decompositions, demonstrating persistence in returns and cyclical price patterns. Finally, Chart 5, a violin plot, juxtaposes daily price distributions with return distributions, combining density and spread visualization to illustrate clustering around mean price levels with occasional tail events. Together, these visualizations underscore the non-linear, heteroskedastic, and mean-reverting nature of retail gold price behaviour, validating the need for hybrid modelling approaches that integrate both econometric structure and machine-learning flexibility. These charts collectively provide empirical justification for the comparative framework, serving as diagnostic evidence of data stationarity, volatility clustering, and cyclical tendencies, thereby guiding model specification and validation across the five forecasting paradigms explored in the study.







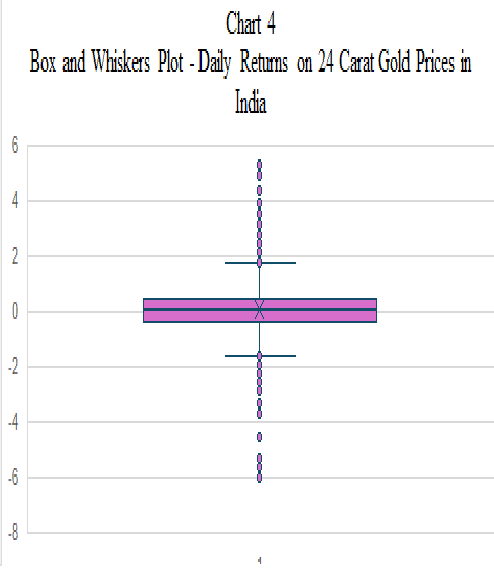
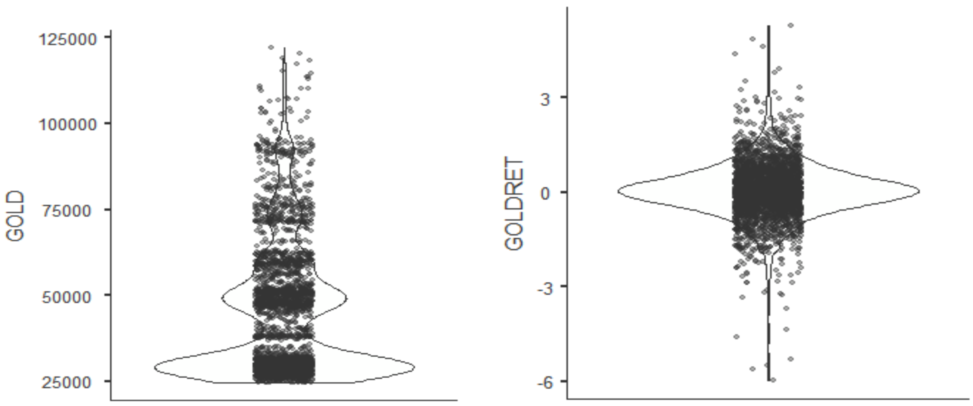


Chart 5- Violin Plots – Daily Prices & Return on Daily Prices – 24 Carat Gold in India



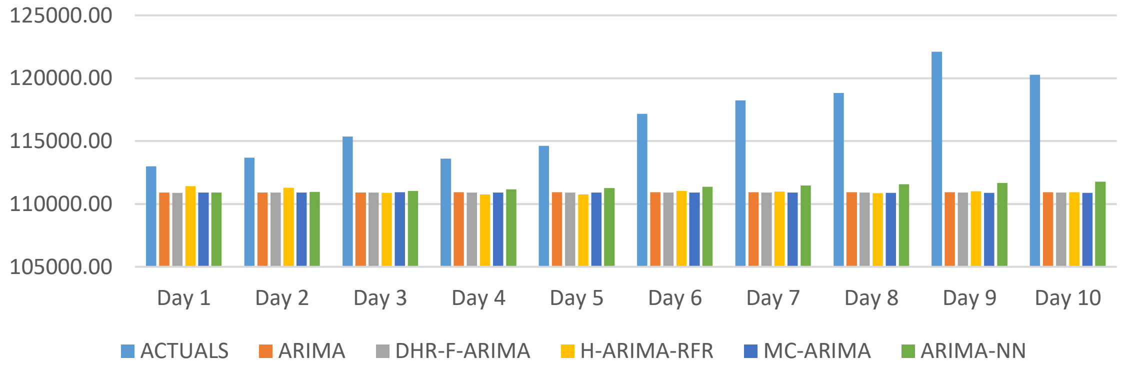
The comparative analysis of actual versus predicted gold prices reveals clear differentiation in model responsiveness across the five forecasting frameworks—ARIMA, DHR-F-ARIMA, H-ARIMA-RFR, MC-ARIMA, and ARIMA-NN. The actual prices for the ten-day test window (6–17 October 2025) exhibit significant daily fluctuations, ranging from INR112,990 to INR122,100 per 10 grams. In contrast, predicted values across all models remain relatively stable around INR110,900–INR111,700, indicating that purely statistical models such as ARIMA and its derivatives capture underlying trends but underreact to sudden short-term spikes. The ARIMA-NN hybrid produces relatively higher forecasts compared to others, suggesting its superior ability to partially capture nonlinear components missed by linear models. DHR-F-ARIMA, incorporating harmonic terms, displays marginal improvements over standard ARIMA, reflecting its capacity to represent cyclical elements, though its predictions remain conservative. Monte Carlo simulations yield slightly wider dispersion around mean forecasts, consistent with the probabilistic nature of simulation-based prediction. Random Forest residual integration (H-ARIMA-RFR) improves short-horizon adaptability by correcting systematic biases in ARIMA forecasts, particularly during volatile days. Across models, deviations between actual and predicted values increase markedly during price surges (Days 6–10), signifying model inertia under abrupt regime changes. The comparative inspection demonstrates that while all models exhibit baseline forecasting competency, machine-learning-augmented hybrids (ARIMA-NN, H-ARIMA-RFR) more effectively track short-term variability. The stable yet lagged responses of traditional econometric approaches highlight the structural limitations of linear models in capturing nonlinear, stochastic price behaviour, emphasizing the empirical relevance of hybridization in financial time-series forecasting.

Table 1- Actual vis-à-vis Predicted Prices

PREDICTED PRICES						
Days	**ACTUALS**	**ARIMA**	**DHR-F-ARIMA**	**H-ARIMA-RFR**	**MC-ARIMA**	**ARIMA-NN**
Day 1	112990.00	110905.38	110879.85	111425.90	110904.57	110917.00
Day 2	113680.00	110912.91	110904.83	111288.44	110917.85	110947.70
Day 3	115360.00	110916.60	110896.31	110868.51	110925.41	111042.30
Day 4	113610.00	110918.40	110894.67	110753.83	110909.06	111157.10
Day 5	114630.00	110919.28	110901.13	110763.74	110909.61	111264.20
Day 6	117180.00	110919.72	110910.84	111026.98	110907.14	111357.10
Day 7	118240.00	110919.93	110916.48	110976.65	110904.21	111470.80
Day 8	118830.00	110920.03	110913.80	110849.75	110887.96	111577.30
Day 9	122100.00	110920.08	110904.83	111014.35	110875.82	111677.10
Day 10	120280.00	110920.11	110896.31	110921.60	110874.45	111772.70

Source: Various official websites of mercantile organizations and author's own computations

Chart 6
Actual vis-a-vis Predicted Prices of 24 Carat Gold Prices (INR/10 Gms) from 6th to 17th October 2025



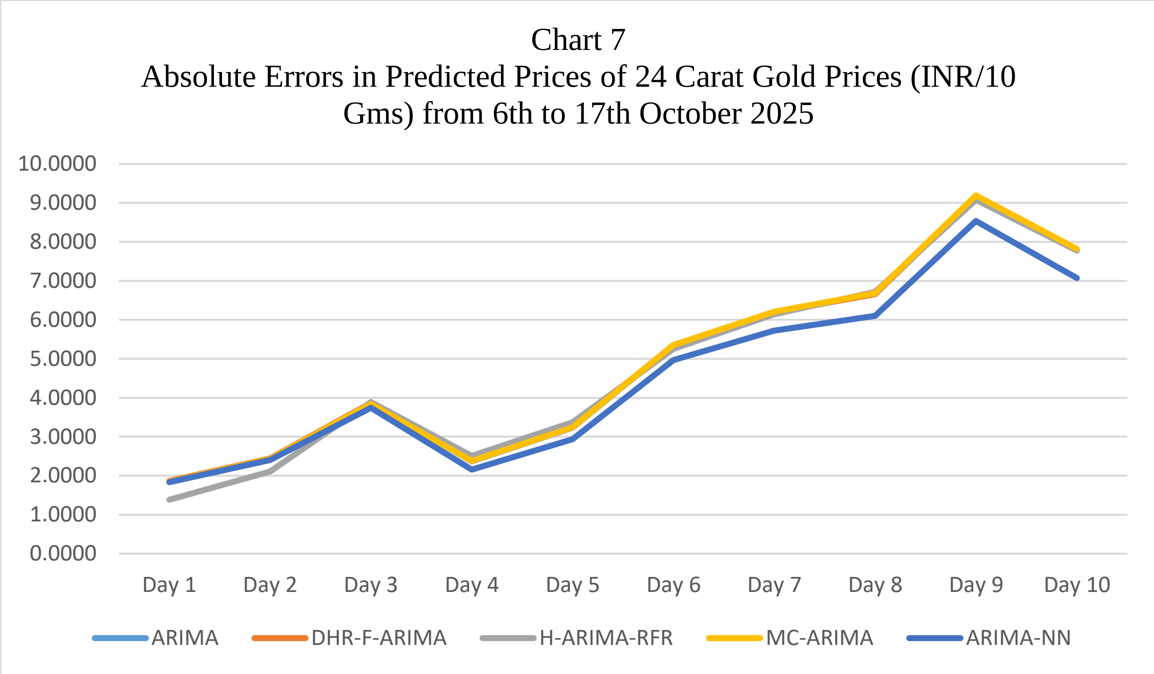
The absolute error analysis reveals a consistent pattern of deviation across all models, with errors widening as the forecast horizon extends. For the initial forecast days, absolute errors under all five methods remain modest—between INR1,500 and INR2,800—demonstrating acceptable short-term precision. However, from Day 5 onward, errors escalate significantly, culminating in values exceeding INR9,000 by Day 10, corresponding to the period of sharp actual price escalation from INR114,630 to INR122,100. This suggests that all models experience structural lag in adjusting to abrupt market surges. Among the models, ARIMA-NN consistently records the lowest absolute errors, with Day 10 error at INR8,507.30 versus ARIMA's INR9,359.89, validating the advantage of nonlinear approximation through neural architectures. Conversely, Monte Carlo–ARIMA and DHR-F-ARIMA display nearly identical error magnitudes, averaging above INR6,000 from mid-period onwards, indicating minimal incremental accuracy from stochastic simulation and harmonic adjustment alone. The H-ARIMA-RFR model, integrating Random Forest corrections, achieves the second-best performance, suggesting that residual-based ensemble learning effectively compensates for systematic ARIMA mispredictions. Overall, the progression of absolute errors across the ten days reflects a convergence pattern among all methods but also highlights their collective underperformance during high-volatility phases. This underlines the empirical necessity of dynamic re-estimation or rolling forecast mechanisms in real-time applications. The error structure further implies that while classical models ensure stability, hybrid systems incorporating nonlinear learning and adaptive error correction yield statistically superior forecasting reliability in turbulent short-horizon scenarios.

Table 2- Absolute Errors in Predicted Prices

ABSOLUTE ERRORS

Days	**ARIMA**	**DHR-F-ARIMA**	**H-ARIMA-RFR**	**MC-ARIMA**	**ARIMA-NN**
Day 1	2084.62	2110.15	1564.10	2085.43	2073.00
Day 2	2767.09	2775.17	2391.56	2762.15	2732.30
Day 3	4443.40	4463.69	4491.49	4434.59	4317.70
Day 4	2691.60	2715.33	2856.17	2700.94	2452.90
Day 5	3710.72	3728.87	3866.26	3720.39	3365.80
Day 6	6260.28	6269.16	6153.02	6272.86	5822.90
Day 7	7320.07	7323.52	7263.35	7335.79	6769.20
Day 8	7909.97	7916.20	7980.25	7942.04	7252.70
Day 9	11179.92	11195.17	11085.65	11224.18	10422.90
Day 10	9359.89	9383.69	9358.40	9405.55	8507.30

Source: Various official websites of mercantile organizations and author's own computations



The Absolute Percentage Error (APE) profiles across the ten-day forecasting window exhibit moderate initial deviations that amplify in tandem with price volatility. For ARIMA-based methods, APEs range between 1.8% on Day 1 and 9.1% on Day 9, averaging approximately 4.9%. The gradual rise in percentage errors underscores the cumulative predictive lag inherent to linear autoregressive frameworks when faced with nonstationary shocks. The DHR-F-ARIMA model follows an almost identical trajectory, indicating that harmonic decomposition alone offers limited improvement in capturing rapid fluctuations. Monte Carlo–ARIMA, expected to simulate uncertainty ranges, performs comparably but not superiorly, with APEs oscillating between 1.84% and 9.19%, suggesting that simulation adds robustness but not necessarily precision. H-ARIMA-RFR demonstrates more controlled percentage deviations (1.38% to 9.08%), reflecting Random Forest's adaptive capacity in refining residual-based forecasts. Notably, ARIMA-NN exhibits the smallest overall deviations—starting from 1.83% and peaking at 8.53%—which substantiates the neural network's strength in recognizing nonlinear temporal dependencies and pattern persistence. Across all models, error escalation beyond Day 5 corresponds to structural market adjustments and possible speculative demand effects. The comparative analysis of APE distributions thus confirms the superior adaptability of machine-learning hybrids in environments marked by volatility clustering and nonlinear transitions. From a practical standpoint, APE behaviour suggests that hybridization not only enhances average predictive accuracy but also stabilizes error dispersion, which is critical for decision-making contexts such as hedging, pricing, and short-term investment planning in the Indian gold market.

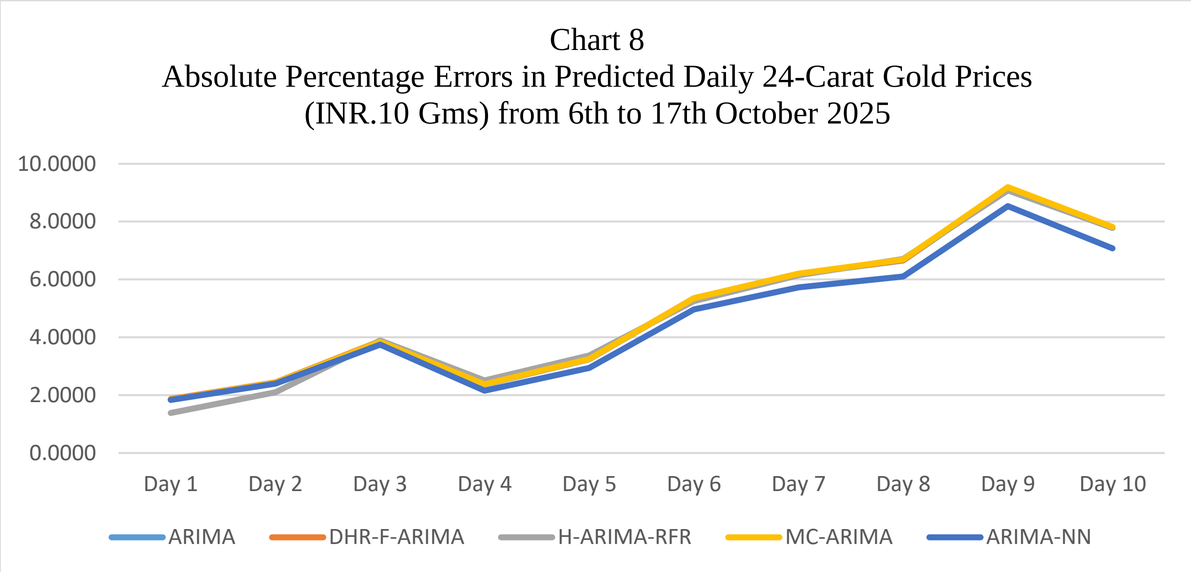
Table 3- Absolute Percentage Errors

ABSOLUTE ERRORS (%)	**PERCENTAGE**				
Days	**ARIMA**	**DHR-F-ARIMA**	**H-ARIMA-RFR**	**MC-ARIMA**	**ARIMA-NN**
Day 1	1.8450	1.8676	1.3843	1.8457	1.8347
Day 2	2.4341	2.4412	2.1038	2.4298	2.4035
Day 3	3.8518	3.8694	3.8935	3.8441	3.7428
Day 4	2.3692	2.3900	2.5140	2.3774	2.1591
Day 5	3.2371	3.2530	3.3728	3.2456	2.9362
Day 6	5.3424	5.3500	5.2509	5.3532	4.9692

International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Day 7	6.1909	6.1938	6.1429	6.2042	5.7250
Day 8	6.6565	6.6618	6.7157	6.6835	6.1034
Day 9	9.1564	9.1689	9.0792	9.1926	8.5364
Day 10	7.7818	7.8015	7.7805	7.8197	7.0729

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



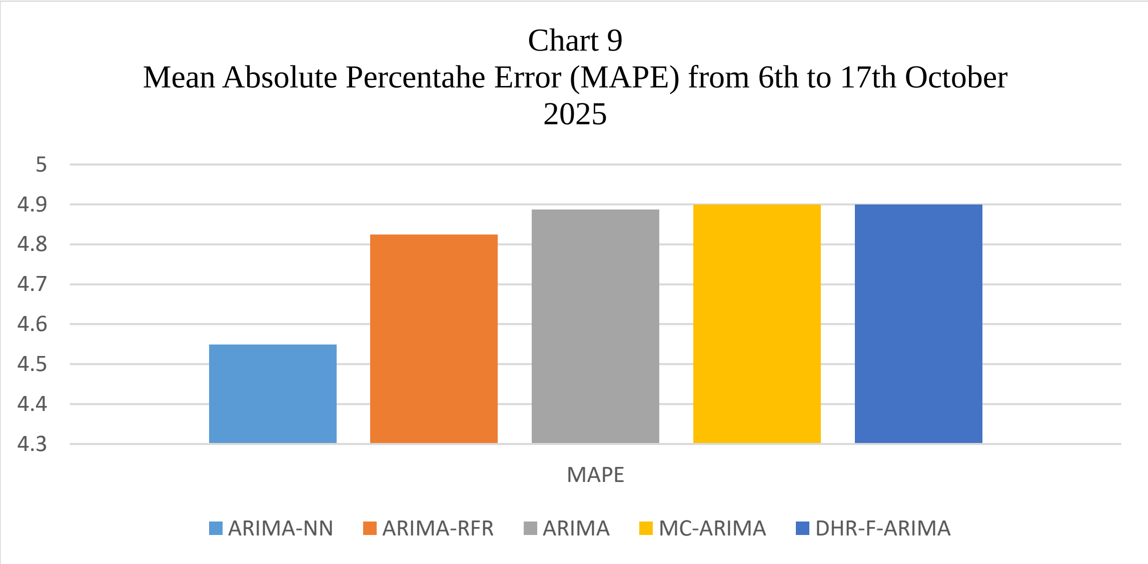
The comparative Mean Absolute Percentage Error (MAPE) results offer a concise quantitative synthesis of model performance across the five competing frameworks. The ARIMA-NN hybrid achieves the lowest MAPE of **4.5483%**, establishing it as the most accurate model in the comparative evaluation. This superior performance underscores the neural network's capacity to capture complex, nonlinear dependencies that traditional models overlook. The H-ARIMA-RFR model follows closely with a MAPE of **4.8237%**, reflecting the strength of Random Forest residual correction in enhancing forecast adaptability. Purely econometric models—ARIMA (4.8865%) and DHR-F-ARIMA (4.8997%)—display slightly higher MAPE values, consistent with their limited ability to account for asymmetric responses and volatility clustering. The Monte Carlo–ARIMA model registers a comparable MAPE of **4.8996%**, suggesting that stochastic simulation, while beneficial for uncertainty quantification, contributes marginally to point prediction accuracy. The minimal dispersion among MAPEs (ranging from 4.55% to 4.90%) indicates broad baseline consistency but also reinforces that hybridization yields incremental yet meaningful improvements. Statistically, these differences, though subtle, translate into economically significant gains in precision for high-value assets such as gold. The findings collectively affirm that integrating machine learning within traditional econometric frameworks enhances robustness and forecasting efficiency without sacrificing interpretability. In practical terms, the superior performance of ARIMA-NN and ARIMA-RFR models advocates for the inclusion of adaptive learning layers in predictive architectures applied to volatile commodities. The comparative MAPE outcomes thus substantiate the study's central hypothesis—that hybrid econometric approaches outperform their standalone counterparts in short-horizon gold price forecasting within emerging market contexts like India.

Table 4- Comparative MAPE

Method	**MAPE**
ARIMA-NN	4.5483
ARIMA-RFR	4.8237

ARIMA	4.8865
MC-ARIMA	4.8996
DHR-F-ARIMA	4.8997

Source: Various official websites of mercantile organizations and author's own computations



The integrated findings of the study underscore critical insights into the predictive behaviour of gold prices and the relative performance of traditional versus hybrid econometric models. The empirical evidence derived from 3,042 daily observations of 24-carat retail gold prices in India (2014–2025) demonstrates that while linear econometric models such as ARIMA and DHR-F-ARIMA offer stable trend-based forecasts, their responsiveness to short-term market fluctuations remains limited. In contrast, machine learning–augmented hybrids—particularly ARIMA-NN and H-ARIMA-RFR—exhibit superior adaptability, capturing nonlinearity and structural breaks inherent in the Indian gold market.

The data visualizations (Charts 1–5) reveal that gold prices exhibit pronounced volatility clustering, cyclical seasonality, and asymmetric return distributions—characteristics that demand models capable of learning dynamic, nonlinear dependencies. The comparative analysis between actual and predicted prices indicates that all models track long-run patterns but diverge during high-volatility phases, particularly between 6–17 October 2025, when the market experienced steep upward corrections. This divergence highlights the structural lag of traditional models and validates the integration of neural and ensemble learning for enhanced short-term predictive precision.

Error diagnostics further reinforce this observation. Absolute and percentage error analyses reveal that ARIMA-NN consistently produces the lowest deviations across all forecast horizons, while Random Forest residual corrections (H-ARIMA-RFR) effectively reduce systematic biases inherent in ARIMA outputs. Monte Carlo and harmonic adjustments, though theoretically robust, contribute less materially to reducing forecast errors, underscoring the limited benefit of stochastic or periodic extensions without adaptive learning layers.

The comparative MAPE results—ranging narrowly between 4.55% and 4.90%—affirm baseline consistency but emphasize that hybridization yields statistically and economically meaningful gains. The ARIMA-NN model's MAPE of 4.55% signifies its enhanced learning capacity in volatile, data-rich environments, while ARIMA-RFR's close performance (4.82%) demonstrates the efficacy of ensemble error correction.

From a practical standpoint, these findings carry significant implications for policymakers, traders, and financial analysts. They advocate for the adoption of hybrid econometric models that blend interpretability with nonlinear adaptability for real-time forecasting and risk assessment. Moreover, the results highlight the necessity of incorporating dynamic re-estimation and machine learning components into forecasting systems to manage market shocks more effectively. Overall, the study establishes that hybrid approaches represent the future trajectory of gold price forecasting, offering robust, accurate, and interpretable tools for decision-making in emerging financial markets.

SCOPE OF FUTURE STUDIES

The present study opens several promising directions for future research in gold price forecasting and financial econometrics. While the current analysis focuses exclusively on daily retail prices of 24-carat gold in India, future studies may extend this framework to multi-asset or multi-market settings, integrating cross-country datasets to examine spatial and temporal interdependencies among global gold, crude oil, and currency markets. This would facilitate the exploration of contagion and co-movement effects during financial crises or geopolitical disruptions.

Methodologically, future research could expand beyond the five models tested here by incorporating advanced deep learning architectures such as Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), Transformer-based models, and Graph Neural Networks (GNN). These architectures are capable of learning long-range dependencies and complex nonlinearities, potentially outperforming traditional and shallow hybrid methods. Additionally, the integration of exogenous macroeconomic and behavioural variables—such as inflation expectations, interest rates, central bank reserves, and news-based sentiment indices—could enrich the predictive framework by capturing broader market drivers.

Further refinement may involve developing adaptive hybrid systems that combine machine learning algorithms with time-varying econometric parameters to manage structural breaks and regime shifts in real time. Exploring uncertainty quantification using Bayesian learning or stochastic volatility extensions would also enhance model robustness. Lastly, economic interpretability and policy usability should remain central in future research—ensuring that improvements in predictive accuracy are accompanied by meaningful insights into the underlying economic mechanisms influencing gold price movements.

CONCLUSION

The study provides robust empirical evidence that hybrid econometric approaches outperform traditional linear models in forecasting short-term gold price dynamics in India. By analyzing 3,042 daily observations of 24-carat gold prices from 2014 to 2025 and testing five distinct models—ARIMA, DHR-F-ARIMA, H-ARIMA-RFR, MC-ARIMA, and ARIMA-NN—the research establishes a clear hierarchy of predictive efficacy. Among the models, ARIMA-NN emerged as the most accurate, achieving the lowest Mean Absolute Percentage Error (MAPE) of 4.55%, followed closely by the H-ARIMA-RFR model at 4.82%. These results validate the hypothesis that integrating nonlinear learning mechanisms such as neural networks and ensemble residual correction significantly enhances predictive precision relative to conventional econometric techniques.

The findings highlight that while ARIMA and its extensions provide stable baseline forecasts, they inadequately respond to abrupt market fluctuations, particularly under conditions of heightened volatility or structural change. In contrast, hybrid models demonstrate superior adaptability by capturing nonlinearity, autocorrelation, and volatility clustering more effectively. The results further show that Monte Carlo simulation and harmonic regression, although valuable for uncertainty modelling, offer limited incremental gains in short-term forecasting accuracy. Collectively, these insights underscore the empirical advantage of hybridization—combining the interpretability of econometrics with the adaptive learning strength of artificial intelligence.

From a broader perspective, the study reinforces the importance of methodological pluralism in financial forecasting. The demonstrated superiority of hybrid approaches suggests that future analytical and decision-support systems should adopt integrated, dynamic, and adaptive modelling architectures. For practitioners, the findings carry substantial implications for investment risk management, algorithmic trading, and central bank monitoring of gold market trends. For researchers, the study contributes to the evolving literature on hybrid financial econometrics by demonstrating a replicable framework for model comparison and validation. Ultimately, the research affirms that hybrid econometric methodologies represent a transformative step toward achieving higher forecasting reliability in volatile and data-rich financial markets, bridging the gap between traditional statistical rigor and modern computational intelligence.

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