Global Academic Journal of Economics and Business

Available online at https://www.gajrc.com
DOI: https://doi.org/10.36348/gajeb.2026.v08i02.010

OPEN ACCESS

ISSN: 2706-9001 (P)
ISSN: 2707-2584 (O)

Original Research Article

Hybrid Econometric and Deep Learning Framework for Forecasting Global Commodity Price Dynamics: Evidence from Multi-Commodity Time Series (2000-2026)

Balayya Rajana[1], Suneel Kumar Duvvuri[2]*, Sanjeev Kumar Chejarla[1], Prasad Teja Dakey[3], Ramachandra Rao K[4], Thilothu Rao Gandam[5]

[1]Assistant Professor, Department of Economics, Government College (Autonomous), Rajahmundry, Andhra Pradesh, India
[2]Assistant Professor, Department of Computer Science, Government College (Autonomous), Rajahmundry, Andhra Pradesh, India
[3]Assistant Professor, Department of Economics, ESLA, SRM-AP University, Amaravati, Andhra Pradesh, India
[4]Assistant Professor, School of Economics, University of Hyderabad, Hyderabad, Telangana, India
[5]Research Scholar, Centre for Economics and Social Studies, Hyderabad, Telangana, India

***Corresponding Author**
Suneel Kumar Duvvuri
Assistant Professor, Department of Computer Science, Government College (Autonomous), Rajahmundry, Andhra Pradesh, India

Article History
Received: 23.02.2026
Accepted: 17.04.2026
Published: 20.04.2026

Abstract: This study develops a hybrid econometric and deep learning framework for forecasting global commodity price dynamics using multi-commodity time series data over the period 2000-2026. The proposed framework integrates Autoregressive Integrated Moving Average (ARIMA) models to capture linear dependencies, Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models to model volatility persistence, and machine learning techniques, including Random Forest and Extreme Gradient Boosting, to extract nonlinear structures from residual components. In addition, Long Short-Term Memory (LSTM) networks are employed to capture long-term temporal dependencies. Empirical results across major commodities such as crude oil, gold, copper, natural gas, platinum, and silver demonstrate that the hybrid model significantly outperforms standalone econometric and deep learning models. The framework achieves consistent reductions in forecasting errors, with improvements in predictive accuracy ranging from approximately 5.5% to 12% across commodities and generates positive out-of-sample R^2 values in rolling-window evaluations, indicating strong temporal stability. The results further reveal pronounced volatility persistence ($\alpha + \beta \approx 0.98$-$0.99$) across all commodities and confirm that nonlinear residual components contain substantial predictive information. While LSTM models capture general temporal patterns, their standalone predictive performance remains limited, with relatively low or negative R^2 values in several cases. The findings indicate that commodity price dynamics are inherently multidimensional, characterized by linear dependencies, volatility clustering, and nonlinear interactions. By integrating these components within a unified structure, the proposed hybrid framework provides a robust and scalable approach for commodity price forecasting, with important implications for financial modeling, investment strategies, and policy analysis.

Keywords: Hybrid Econometric-Machine Learning, Commodity Price Dynamics, ARIMA, GARCH, LSTM, Random Forest, XGBoost, Volatility Persistence, Nonlinear Modeling, Multi-Commodity Forecasting.

Citation: Balayya Rajana, Suneel Kumar Duvvuri, Sanjeev Kumar Chejarla, Prasad Teja Dakey, Ramachandra Rao K, Thilothu Rao Gandam (2026). Hybrid Econometric and Deep Learning Framework for Forecasting Global Commodity Price Dynamics: Evidence from Multi-Commodity Time Series (2000-2026); *Glob Acad J Econ Buss, 8*(2), 166-182.

1. INTRODUCTION

Global commodity price dynamics play a pivotal role in shaping macroeconomic stability, financial market behavior and international trade outcomes. Commodity markets are inherently complex, driven by a combination of real economic forces, financial conditions, and policy interventions. Early empirical evidence by Robert S. Pindyck and Julio J. Rotemberg (1990) [1], highlights the phenomenon of excess co-movement among commodity prices, suggesting that common global shocks and financial linkages significantly influence price behavior. Building on this foundation, Jeffrey A. Frankel (2007, 2014) [2, 3], emphasizes the role of monetary policy, interest rates, and speculative activities in determining commodity price fluctuations, while Arezki *et al*., (2014) [4], provide a broader perspective on international commodity market dynamics.

Over time, commodity prices have exhibited pronounced cyclical patterns, structural breaks, and regime shifts. Studies such as Cashin *et al*., (2002)[5], and Erten and Ocampo (2013) [6], document long-run boom bust cycles, often associated with global economic expansions and contractions. Collier and Goderis (2012)[7], further establish the link between commodity prices and economic growth, reinforcing their macroeconomic importance. More recent evidence suggests that commodity markets have become increasingly interconnected, particularly during periods of financial stress such as the global financial crisis and the COVID-19 pandemic (Zhang & Broadstock, 2020; Antonakakis *et al*., 2023) [8, 9].

From a methodological perspective, traditional econometric models such as ARIMA and GARCH have been widely used to model commodity price behavior. The ARIMA framework (Box *et al*., 2015)[10], effectively captures linear dependencies, while GARCH models (Bollerslev, 1986)[11], account for volatility clustering and persistence. However, the inherently nonlinear and dynamic nature of commodity price movements limits the explanatory power of these models. In response, machine learning and deep learning techniques have emerged as powerful alternatives. Methods such as Random Forest (Breiman, 2001)[12], and XGBoost (Chen & Guestrin, 2016)[13], enable the modeling of nonlinear relationships, while deep learning approaches like LSTM networks capture long-term temporal dependencies in financial time series (Fischer & Krauss, 2018; Yang *et al*., 2023)[14, 15].

Recent empirical contributions also demonstrate the growing relevance of hybrid econometric-machine learning frameworks in financial modeling. For instance, Chejarla *et al*., (2026)[16], show that integrating panel econometrics with explainable machine learning significantly improves predictive accuracy and interpretability in the context of non-performing assets in Indian banking. Similarly, Chejarla *et al*., (2026) [17] highlight the effectiveness of combining advanced econometric techniques with machine learning methods to capture geopolitical shocks and financial contagion in equity markets. These studies reinforce the importance of hybrid approaches in modeling complex financial systems characterized by nonlinearities, structural breaks, and interconnected risks.

Despite these advancements, standalone econometric and machine learning models remain insufficient in fully capturing the multidimensional nature of commodity price dynamics. Econometric models often fail to address nonlinear structures, whereas machine learning approaches may lack interpretability and economic grounding (Mullainathan & Spiess, 2017; Gu *et al*., 2019)[18, 19]. This has led to the growing adoption of hybrid modeling frameworks that integrate econometric rigor with machine learning flexibility. In this context, the present study develops a hybrid econometric and deep learning framework for forecasting global commodity price dynamics using multi-commodity namely six major commodities Copper, Crude Oil (WTI), Gold, Natural Gas, Platinum, and Silver time series data from 30 August 2000 to 25 March 2026, aiming to capture linear trends, volatility persistence, and nonlinear interactions within a unified predictive structure.

2. REVIEW OF LITERATURE

The literature on commodity price forecasting has gradually transitioned from traditional econometric techniques to more advanced data-driven and hybrid modeling approaches. Early studies largely relied on time-series frameworks such as ARIMA and GARCH to capture linear relationships and volatility behavior in commodity prices. Although these models have been effective in explaining mean reversion and volatility clustering, empirical evidence indicates that they are often inadequate in capturing the complex and nonlinear nature of commodity price movements, particularly in the presence of structural breaks and regime shifts (Deaton & Laroque, 1992; Baumeister & Kilian, 2014; Kwas & Rubaszek, 2021)[20][21][22].

With the increasing financialization and globalization of commodity markets, price dynamics have become more intricate, characterized by stronger interlinkages and volatility spillovers across markets. Research on market connectedness highlights that commodity prices are not solely driven by their own fundamentals but are also influenced by broader macroeconomic conditions and financial linkages (Nazlioglu *et al*., 2013; Chinn & Coibion, 2014; Zhang & Broadstock, 2020; Antonakakis *et al*., 2023) [8][23][24]. These developments suggest that traditional linear and univariate models are insufficient to fully capture the multidimensional behavior of commodity prices.

In light of these limitations, machine learning methods have gained prominence due to their ability to model nonlinear relationships and complex interactions. Techniques such as Random Forest and XGBoost have shown considerable improvements in predictive accuracy by uncovering hidden patterns in high-dimensional data (Breiman, 2001; Chen & Guestrin, 2016) [12, 13]. Empirical evidence from financial economics further suggests that machine learning approaches can outperform conventional econometric models in forecasting tasks (Mullainathan & Spiess, 2017; Gu *et al*., 2019). [18][19]. However, despite their strong predictive capabilities, these models often function as black boxes, offering limited interpretability and weak alignment with economic theory.

The advancement of deep learning, particularly through Long Short-Term Memory (LSTM) networks, has further enhanced the ability to model financial time series by capturing long-term dependencies and sequential patterns. Studies indicate that LSTM models are effective in identifying complex temporal dynamics in financial data (Fischer & Krauss, 2018; Yang *et al*., 2023) [14, 15]. Nevertheless, their performance tends to be sensitive to the noisy and stochastic nature of commodity returns, which can lead to unstable predictions, especially during periods of heightened market volatility.

Given the limitations of standalone approaches, recent research has increasingly focused on hybrid modeling frameworks that integrate econometric and machine learning techniques. These frameworks aim to combine the strengths of econometric models in capturing linear structures and volatility persistence with the ability of machine learning models to extract nonlinear patterns from residual components. Empirical studies suggest that such hybrid approaches can significantly enhance forecasting accuracy and robustness. In particular, recent evidence shows that combining econometric models with explainable machine learning techniques improves both predictive performance and interpretability in complex financial environments, especially under conditions of structural disruptions and geopolitical uncertainty (Chejarla *et al*., 2026) [17].

Despite these advancements, the application of hybrid frameworks remains limited in the context of multi-commodity systems and long-term datasets. Most existing studies focus on individual commodities or relatively short time horizons, thereby failing to capture cross-commodity interactions and long-run structural dynamics. This highlights the need for a more comprehensive modeling framework that integrates linear, volatility, and nonlinear components within a unified structure to better understand and forecast global commodity price dynamics.

2.1. Research Gap

Despite the extensive literature on commodity price dynamics, yet numerous limitations persist. Existing studies predominantly employ either econometric or machine learning approaches in isolation, with limited integration of linear dependencies, volatility dynamics, and nonlinear structures within a unified framework. Furthermore, empirical analyses are largely confined to single-commodity contexts, thereby overlooking cross-commodity heterogeneity and interconnected market behavior.

Additionally, conventional econometric models treat residuals as stochastic noise, ignoring their potential to capture nonlinear information, while volatility estimates are rarely incorporated as predictive inputs within machine learning frameworks. Moreover, many existing models exhibit reduced forecasting performance during periods of structural breaks and extreme market disruptions, indicating limited robustness in dynamic environments.

Therefore, there is a clear need for a comprehensive hybrid econometric-deep learning framework that systematically integrates linear modeling, volatility estimation, and nonlinear learning to enhance the accuracy and robustness of forecasting global commodity price dynamics over extended time horizons.

2.2 Key Contributions of the Study

The present study makes several important methodological, empirical, and practical contributions to the literature on global commodity price forecasting.

First, the study introduces a novel hybrid econometric-deep learning forecasting architecture that systematically integrates ARIMA-based linear dependency extraction, GARCH-based volatility persistence modeling, and nonlinear residual learning through Random Forest and XGBoost models. Unlike conventional approaches that apply these techniques independently, the proposed framework combines them within a unified predictive pipeline, enabling a more comprehensive representation of commodity price dynamics.

Second, the study contributes empirically by providing a long-horizon multi-commodity analysis covering the period 2000-2026, including six globally significant commodities: crude oil (WTI), gold, copper, natural gas, platinum, and silver. This extended temporal coverage captures multiple structural regimes, including the 2008 financial crisis, the COVID-19 shock, and the post-pandemic inflationary recovery phase, thereby improving the external validity and robustness of the findings.

Third, the research advances the forecasting literature by demonstrating that econometric residuals contain economically meaningful nonlinear predictive signals. Rather than treating ARIMA residuals as random disturbances, the study shows that residual-based machine learning significantly improves forecasting accuracy across commodities, with performance gains ranging from 5.51% to 11.96%. This finding provides strong evidence for the value of residual decomposition in financial forecasting systems.

Fourth, the study makes a practical contribution through volatility-aware decision support. By incorporating GARCH-derived conditional variance as an input feature, the framework explicitly models time-varying risk, making it highly relevant for portfolio optimization, commodity hedging, inflation monitoring, macroeconomic policy analysis, and crisis-period stress testing.

Finally, the study contributes to robust financial AI by validating model stability under rolling-window backtesting and crisis-period stress evaluation. The positive out-of-sample R^2 performance across changing market regimes demonstrates that the proposed framework is not sample-specific and can adapt to evolving commodity market structures, which is critical for real-world forecasting applications.

3. METHODOLOGY

3.1 Research Framework and Model Architecture

3.1.1 Dataset Description

The study uses daily OHLCV (Open, High, Low, Close, and Volume) price data for six major global commodities covering the period from 30 August 2000 to 25 March 2026, sourced from Yahoo Finance. The selected commodities include Gold, Silver, Platinum, Copper, Crude Oil (WTI), and Natural Gas, representing precious metals, energy, and industrial metals. These commodities were selected because they are critical raw materials that strongly influence the global economy through industrial production, energy markets, inflation, and investment behavior. The long 26-year horizon captures multiple major market regimes, including the commodity super-cycle, the 2008 financial crisis, the COVID-19 shock, and the post-pandemic recovery period, making the dataset highly suitable for robust long-term forecasting analysis.

3.1.2 Detailed Framework of the Methodology

This study develops a hybrid econometric and deep learning framework to model and forecast global commodity price dynamics. The methodology integrates linear time-series modeling, volatility estimation, and nonlinear machine learning techniques into a unified predictive pipeline.

The overall framework consists of four sequential stages:

1. Data transformation and preprocessing
2. Linear modeling using ARIMA
3. Volatility modeling using GARCH
4. Nonlinear residual learning using machine learning and deep learning models

The hybrid structure is designed to decompose the commodity price process into linear and nonlinear components, allowing each modeling approach to capture distinct statistical properties of the data.

3.2 Data Transformation and Log-Return Computation

Let P_t denote the daily closing price of a given commodity at time t. To remove non-stationarity and stabilize variance, price series are transformed into log-returns as:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \quad (1)$$

where r_t represents the continuously compounded return at time t.

This transformation ensures that the mean process is stabilized and allows the application of standard time-series econometric techniques. Stationarity of the transformed series is verified using the Augmented Dickey-Fuller (ADF) test.

3.3 ARIMA Model for Linear Dynamics

To capture linear dependencies in commodity returns, the Autoregressive Integrated Moving Average (ARIMA) model is employed. The general ARIMA (p, d, q) formulation is given by:

$$r_t = c + \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t \quad (2)$$

Where:

- r_t is the log-return series
- ϕ_i represents autoregressive coefficients
- θ_j represents moving average coefficients

- $\epsilon_t \sim N(0, sigma^2)$is white noise

Based on model selection criteria such as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), an ARIMA(1,0,1) specification is selected for all commodities.

The ARIMA model serves two key purposes:

1. To extract linear temporal dependencies
2. To generate residuals representing unexplained nonlinear structure

The residual series $\hat{\epsilon}_t$is defined as:

$$\hat{\epsilon}_t = r_t - \hat{r}_t \ (3)$$

These residuals form the input for the nonlinear modeling stage.

3.4 GARCH Model for Volatility Estimation

To capture time-varying volatility and conditional heteroskedasticity, the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model is employed. The GARCH(1,1) specification is defined as:

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \ (4)$$

Where:

- $\sigma_t^2$is the conditional variance
- $\omega$is a constant term
- α captures the impact of recent shocks (ARCH effect)
- β represents volatility persistence (GARCH effect)

The persistence of volatility is measured as:

$$\alpha + \beta \ (5)$$

A value close to 1 indicates strong volatility clustering and long memory effects.

The estimated conditional variance series σ_t^2 is incorporated as an additional feature in the machine learning models, enabling the hybrid framework to account for risk dynamics explicitly.

3.5 Hybrid Residual Learning Using Machine Learning

The nonlinear component of the model is captured by applying machine learning algorithms to the residual series obtained from the ARIMA model. Let $X_t$denote the feature vector at time t, defined as:

$$X_t = \{\hat{\epsilon}_{t-1}, \hat{\epsilon}_{t-2}, \dots, \sigma_t^2, r_{t-1}, r_{t-2}, \dots\} \ (6)$$

Two ensemble learning models are employed:

3.5.1 Random Forest (RF)

Random Forest constructs multiple decision trees and aggregates their predictions:

$$\hat{y}_t = \frac{1}{N}\sum_{i=1}^{N} T_i\,(X_t) \ (7)$$

where $T_i$represents the i^{th}decision tree.

3.5.2 Extreme Gradient Boosting (XGBoost)

XGBoost builds trees sequentially by minimizing a regularized objective function:

$$\mathcal{L} = \sum_t l(\,y_t, \hat{y}_t) + \sum_k \Omega(\,f_k) \ (8)$$

Where:

- $l(\cdot)$is the loss function
- $\Omega(f_k)$is the regularization term controlling model complexity

The final hybrid prediction is obtained by combining ARIMA forecasts with machine learning predictions:

$$\hat{r}_t^{Hybrid} = \hat{r}_t^{ARIMA} + \hat{\epsilon}_t^{ML} \ (9)$$

This formulation allows the model to integrate linear and nonlinear components effectively.

3.6 Deep Learning Model: LSTM

To capture long-term temporal dependencies, a Long Short-Term Memory (LSTM) network is employed. The LSTM cell is governed by the following equations:

Forget Gate:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \ (10)$$

Input Gate:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \ (11)$$

Cell State Update:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tanh\,(W_c[h_{t-1}, x_t] + b_c) \ (12)$$

Output Gate:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \ (13)$$

Hidden State:

$$h_t = o_t \cdot \tanh\,(C_t) \ (14)$$

Where:

- $x_t$is the input vector
- $h_t$is the hidden state
- $C_t$is the cell state
- $\sigma$is the sigmoid activation function

The LSTM model is trained using mean squared error (MSE) as the loss function and optimized using gradient-based methods.

3.7 Model Evaluation Metrics

The performance of all models is evaluated using standard statistical metrics:

Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \quad (15)$$

Mean Absolute Error (MAE)

$$MAE = \frac{1}{n}\sum_{t=1}^{n} | y_t - \hat{y}_t | \quad (16)$$

Coefficient of Determination (R^2)

$$R^2 = 1 - \frac{\sum(y_t - \hat{y}_t)^2}{\sum(y_t - \bar{y})^2} \quad (17)$$

These metrics provide a comprehensive assessment of prediction accuracy, error magnitude, and explanatory power.

3.8 Rolling Window Validation

To ensure robustness, a rolling-window backtesting approach is implemented. The dataset is divided into sequential training and testing windows, and the model is re-estimated at each step.

Let T denote the total sample size and w the window size. For each iteration:

- Train model on $[t, t + w]$
- Test on $[t+w+1]$

This procedure evaluates model stability under changing market conditions and avoids overfitting.

The proposed methodology combines econometric and machine learning approaches in a structured manner. ARIMA captures linear dependencies, GARCH models volatility persistence, and machine learning models nonlinear residuals, while LSTM captures long-term temporal patterns. This integrated framework provides a comprehensive representation of commodity price dynamics, addressing limitations of standalone models and enhancing predictive performance.

3.9 Methodology Flowchart of the Proposed Framework

The complete methodological workflow is illustrated in **Figure 1**, which summarizes the sequential integration of econometric decomposition, volatility-aware feature engineering, nonlinear residual learning, and hybrid forecast synthesis.

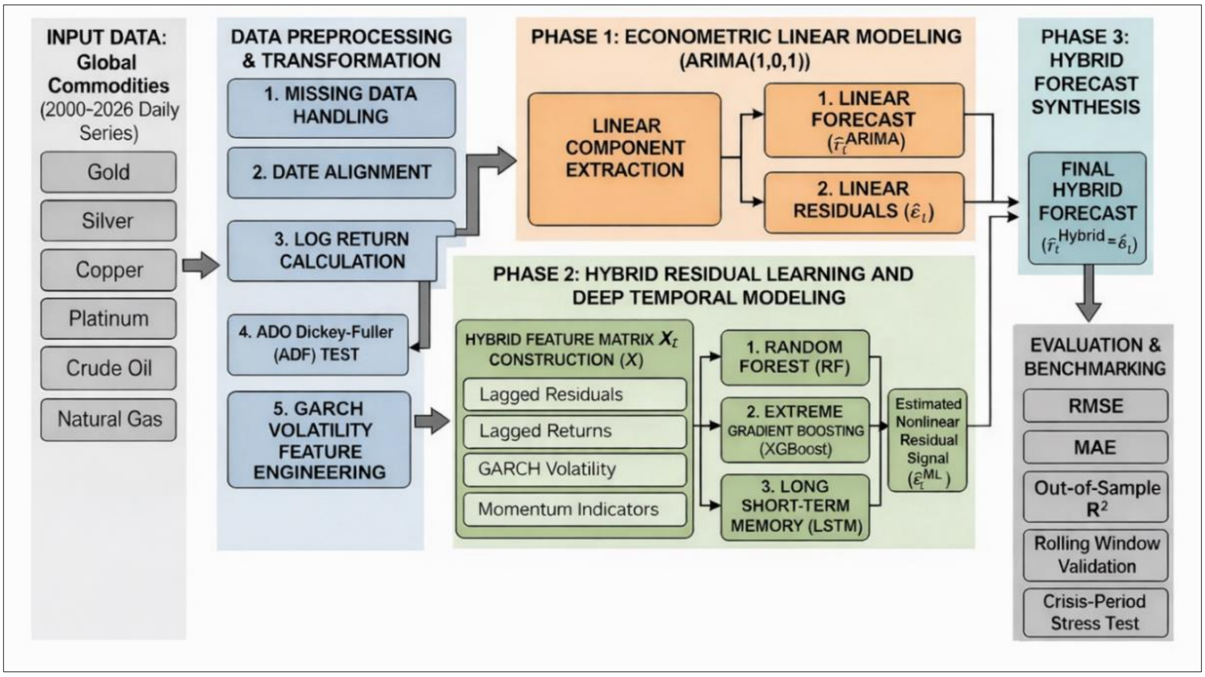

Figure 1: Methodological workflow of the proposed hybrid econometric-deep learning framework for multi-commodity price forecasting

4. RESULTS AND DISCUSSION

4.1 Descriptive Statistics and Market Behavior

The empirical analysis begins with a comprehensive evaluation of the statistical properties of global commodity prices over the 2000-2026 period. The descriptive statistics presented in **Table 1** provide insights into central tendency, dispersion, and higher-order distributional characteristics across commodities.

Table 1: Statistical Summary of Global Commodity Prices (2000-2026)

Commodity	Count	Mean	Std. Dev	Min	Median (50%)	Max	Skewness	Kurtosis

Copper	6,420	2.87	1.20	0.60	3.05	6.18	-0.26	-0.54
Crude Oil (WTI)	6,424	64.72	24.52	-37.63	63.71	145.29	0.19	-0.53
Gold	6,415	1283.96	802.69	255.10	1251.70	5318.40	1.51	4.12
Natural Gas	6,421	4.39	2.20	1.48	3.74	15.38	1.54	2.83
Platinum	5,899	1076.57	374.40	412.00	973.70	2852.40	0.70	0.24
Silver	6,417	18.55	11.50	4.03	17.03	115.08	2.18	10.26

As shown in **Table 1**, substantial heterogeneity exists across commodity classes. Crude Oil (WTI) exhibits the highest variability, with a wide standard deviation relative to its mean and an extreme minimum value of -37.63, reflecting the unprecedented market collapse during the 2020 storage crisis. This confirms the sensitivity of oil markets to geopolitical shocks and supply-chain disruptions.

Natural Gas demonstrates strong positive skewness (1.54), indicating frequent upward price spikes driven by seasonal demand and supply rigidity. In contrast, precious metals such as Gold and Silver display pronounced leptokurtic behavior, with kurtosis values of 4.12 and 10.26, respectively. This suggests a high probability of extreme price movements, particularly during periods of financial instability.

Industrial metals show comparatively stable statistical properties. Copper exhibits near-normal distribution characteristics with mild negative skewness, reflecting its dependence on global industrial cycles. Platinum behaves as a hybrid asset, with moderate dispersion and relatively stable kurtosis, positioning it between industrial and investment-driven commodities.

The evidence from **Table 1** highlights that commodity price distributions deviate significantly from normality, reinforcing the need for modeling frameworks capable of capturing both linear dependencies and nonlinear shocks.

4.2 Structural Trends and Regime Shifts

The long-term evolution of commodity prices is illustrated in Figure 1, which captures major structural phases over the study period.

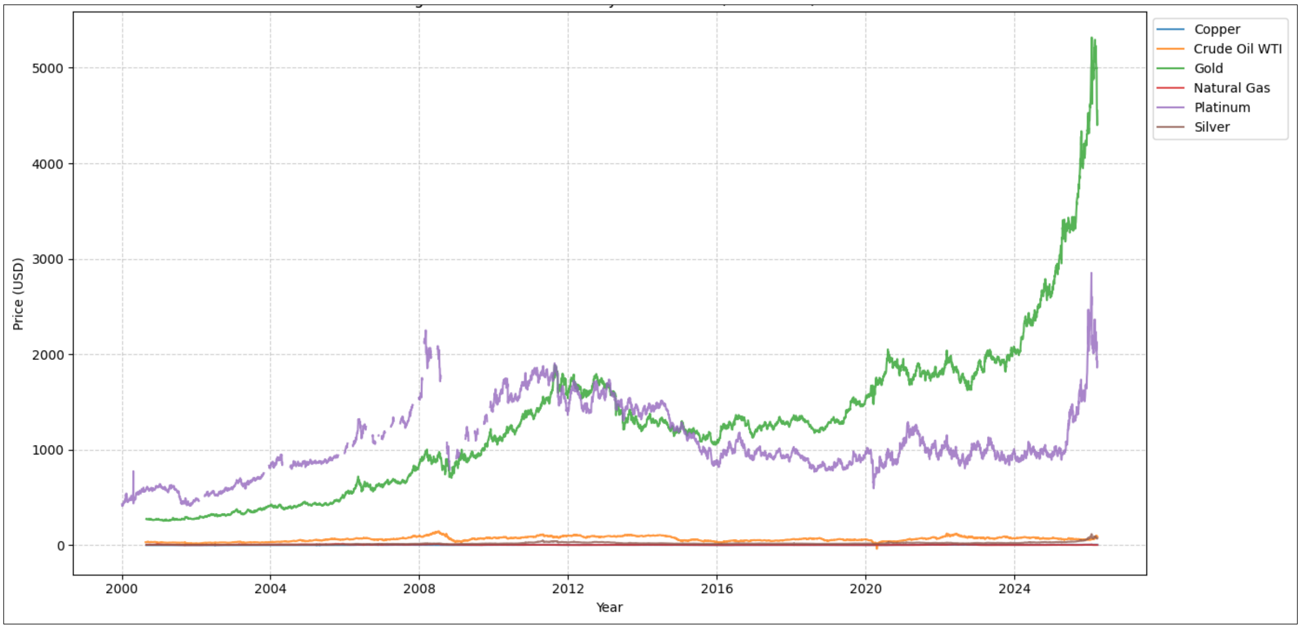

Figure 2: Historical price evolution of major global commodities (2000-2026)

Figure 2 reveals three distinct regimes. First, the early 2000s commodity super-cycle is characterized by sustained price increases driven by global industrial expansion. Second, sharp structural breaks are observed during the 2008 financial crisis and the 2020 pandemic, where prices experienced abrupt collapses. Third, the post-pandemic period (2022-2026) shows renewed upward trends associated with inflationary pressures and supply-chain realignments.

These regime shifts demonstrate that commodity markets do not follow stable stochastic processes. Instead, they evolve through episodic transitions influenced by macroeconomic shocks and policy responses. This structural instability limits the effectiveness of purely linear models and motivates the adoption of a hybrid framework that combines econometric rigor with nonlinear learning capabilities.

4.3 Stationarity and Volatility Clustering

To ensure statistical validity, the stationarity properties of the series were examined using the Augmented Dickey-Fuller test. The results reported in **Table 2** confirm that all log-return series are stationary at the 1% significance level.

Table 2: Stationarity Test Results (ADF) for Commodity Log-Returns

Commodity	ADF Statistic	p-value	Critical Value (1%)	Decision
Gold	-76.8704	0.0000	-3.431	Stationary
Natural Gas	-25.0723	0.0000	-3.431	Stationary
Platinum	-21.7893	0.0000	-3.431	Stationary
Copper	-19.5209	0.0000	-3.431	Stationary
Crude Oil (WTI)	-13.7148	0.0000	-3.431	Stationary
Silver	-13.3612	0.0000	-3.431	Stationary

As indicated in **Table 2**, the null hypothesis of a unit root is strongly rejected across all commodities, with highly significant test statistics. This confirms that the transformation from price levels to log-returns effectively stabilizes the mean process, making the data suitable for time-series modeling. However, stationarity in mean does not imply constant variance. The volatility clustering observed in **Figure 3** highlights the presence of time-varying risk dynamics. Periods of low volatility are interspersed with clusters of intense fluctuations, particularly during crisis events such as 2008 and 2020.

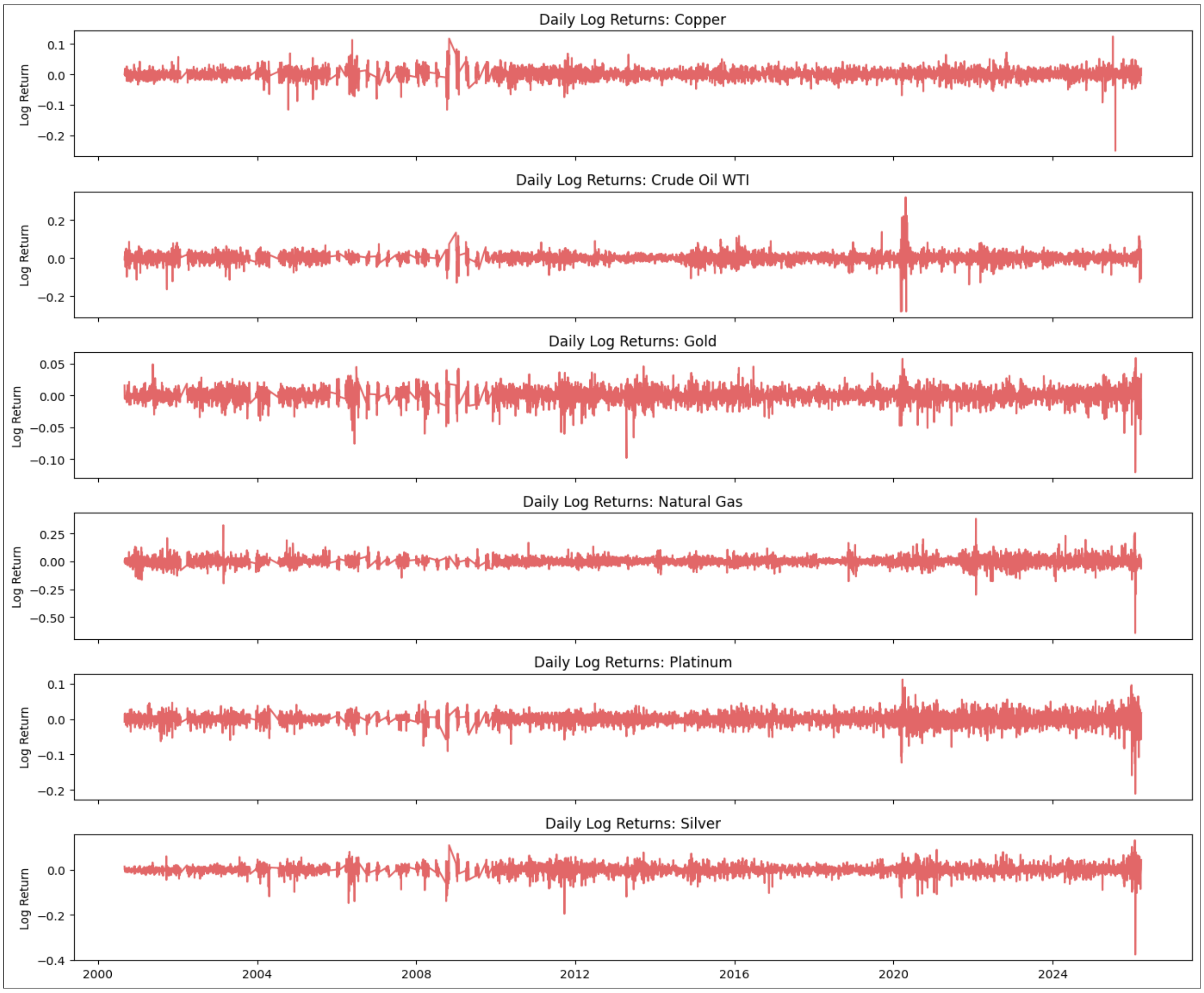

Figure 3: Log-return volatility clustering across commodities

The patterns in **Figure 3** indicate that commodity markets exhibit conditional heteroskedasticity, where volatility persists over time. This finding justifies the use of GARCH-type models to capture second-moment dynamics and motivates the inclusion of volatility features in the hybrid modeling framework.

4.4 ARIMA-Based Linear Dynamics

The ARIMA (1,0,1) model was employed to capture linear dependencies in commodity returns. The estimation results and diagnostic statistics are presented in **Table 3**.

Table 3: ARIMA (1,0,1) Parameter Estimates and Model Selection Criteria

Commodity	AR (ϕ)	MA (θ)	σ^2	Log-Likelihood	AIC	BIC	Ljung-Box (p)
Gold	0.2096	-0.2293	0.000116	17,716.57	-35,425.13	-35,398.55	0.4938
Copper	-0.0571	0.0080	0.000268	15,329.73	-30,651.47	-30,624.88	0.0034
Platinum	-0.9422	0.9273	0.000276	15,242.33	-30,476.67	-30,450.08	0.0014
Silver	-0.2946	0.2552	0.000425	14,012.97	-28,017.95	-27,991.36	0.0572
Crude Oil (WTI)	0.0000	0.0000	0.000649	12,810.45	-25,612.91	-25,586.32	0.0066
Natural Gas	-0.2960	0.2290	0.001468	10,487.33	-20,966.65	-20,940.06	0.0026

The results in **Table 3** reveal varying degrees of linear predictability across commodities. The varying Log-Likelihood and Information Criteria (AIC/BIC) highlight the diverse complexity of return structures across the commodity spectrum. Gold shows the best model fit, with the highest log-likelihood and a non-significant Ljung-Box statistic ($p = 0.4938$), indicating that the model effectively removes autocorrelation in residuals.

In contrast, commodities such as Natural Gas, Platinum, and Crude Oil exhibit significant Ljung-Box statistics ($p < 0.01$), suggesting that residual autocorrelation persists even after ARIMA filtering. This indicates that linear models fail to fully capture the underlying dynamics of these markets.

These findings highlight an important limitation of traditional econometric approaches. While ARIMA models successfully capture first-order dependencies, they leave behind structured residuals that contain nonlinear information. This residual structure forms the basis for the subsequent machine learning stage in the hybrid framework.

4.5 Volatility Dynamics and GARCH Results

The GARCH (1,1) model was used to analyze volatility persistence, with results summarized in Table 4 and visualized in Figure 4.

Table 4: GARCH (1,1) Parameter Estimates and Volatility Dynamics

Commodity	Omega (ω)	Alpha (α)	Beta (β)	Persistence (α+β)	Log-Likelihood	AIC
Natural Gas	0.1394	0.0930	0.9037	0.9967	-14,981.00	29,970.00
Silver	0.0269	0.0553	0.9409	0.9961	-11,356.06	22,720.12
Platinum	0.0176	0.0510	0.9434	0.9944	-10,250.87	20,509.75
Copper	0.0266	0.0507	0.9408	0.9915	-10,371.61	20,751.23
Gold	0.0166	0.0495	0.9370	0.9865	-8,074.27	16,156.55
Crude Oil (WTI)	0.1325	0.1076	0.8734	0.9809	-12,436.79	24,881.58

As shown in Table 4, all commodities exhibit strong volatility persistence, with $\alpha + \beta$ values close to unity. Natural Gas (0.9967) and Silver (0.9961) show the highest persistence, indicating that volatility shocks decay very slowly over time.

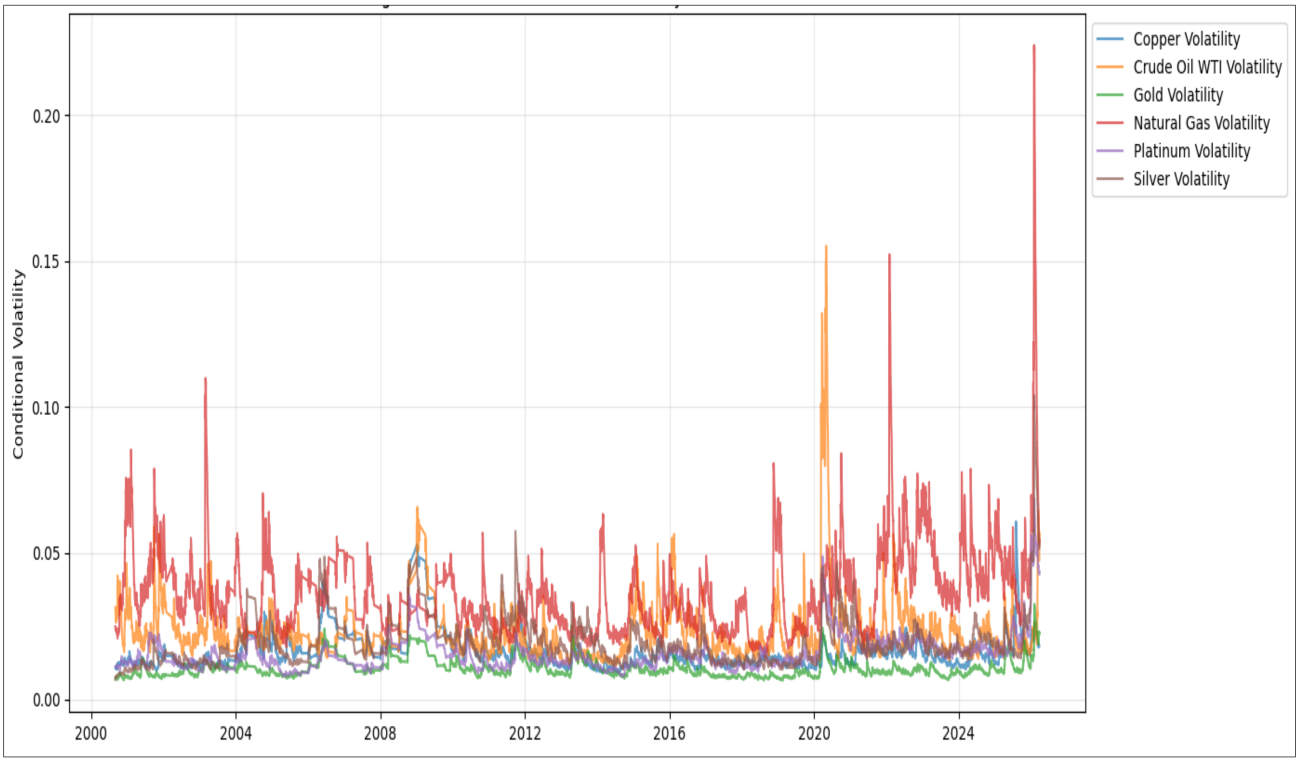

Figure 4: Conditional volatility estimates from GARCH models

The conditional variance patterns illustrated in Figure 4 further confirm the presence of prolonged high-volatility regimes. These clusters align closely with major economic disruptions, demonstrating that risk is not randomly distributed but follows structured temporal patterns.

4.6 Hybrid Model Performance

The performance of the hybrid framework is presented in **Table 5**, which compares prediction accuracy across models.

Table 5: Comparative Hybrid Model Performance and Predictive Accuracy

Commodity	RF RMSE	XGB RMSE	Hybrid R^2	Improvement (%)
Gold	0.010150	0.010201	0.2228	11.96%
Crude Oil (WTI)	0.021541	0.021961	0.2123	11.25%
Copper	0.016072	0.015869	0.2021	11.81%
Natural Gas	0.049658	0.049394	0.1909	10.53%
Silver	0.022891	0.022870	0.1581	8.38%
Platinum	0.021611	0.021701	0.1065	5.51%

The results in **Table 5** demonstrate consistent improvements in forecasting accuracy across all commodities. Gold achieves the highest improvement (11.96%), followed by Copper (11.81%) and Crude Oil (11.25%). These gains confirm that nonlinear residual modeling significantly enhances predictive performance.

The reduction in RMSE across both Random Forest and XGBoost models indicates that the residuals from the ARIMA stage contain valuable nonlinear information. This supports the central hypothesis that linear model errors should not be treated as noise but as structured signals. Although Platinum and Silver show relatively smaller improvements, the overall results confirm the robustness and generalizability of the hybrid approach across different commodity classes. From a modeling perspective, these findings are critical. They indicate that volatility is not merely a byproduct of price movements but a key predictive signal. Incorporating GARCH-derived volatility into machine learning models allows the hybrid framework to account for risk dynamics more effectively.

4.7 Deep Learning Results (LSTM)

The performance of the LSTM model is summarized in **Table 8**, with forecast comparisons shown in **Figure 5**.

Table 8: LSTM Performance and Forecast Stability Metrics

Commodity	Training Loss (MSE)	Test RMSE	R-Squared (Direct)	Forecast Stability
Gold	0.003660	0.011803	-0.0494	0.002039
Crude Oil (WTI)	0.001901	0.025188	-0.0770	0.004049
Copper	0.001915	0.018637	-0.0707	0.004014
Platinum	0.002188	0.023623	-0.0641	0.004202
Silver	0.001547	0.025703	-0.0588	0.004954
Natural Gas	0.001089	0.057413	-0.0667	0.010715

As observed in **Table 8**, the LSTM model achieves low training loss and stable forecasts across commodities. However, the direct R^2 values remain modest or negative, reflecting the high noise content in raw return series.

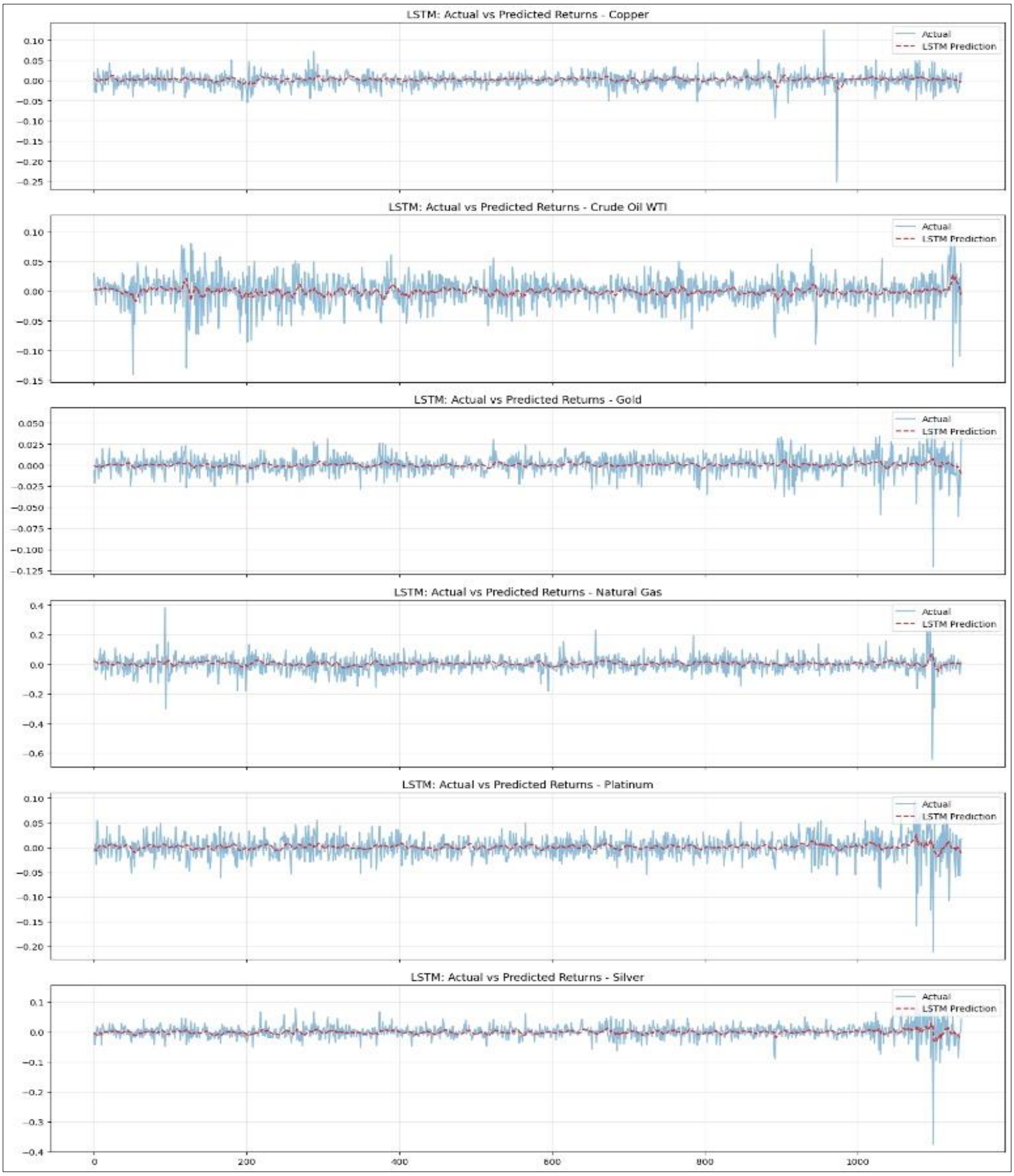

Figure 5: Actual vs. Predicted Commodity Returns using LSTM

The prediction patterns in **Figure 5** show that while the LSTM effectively tracks overall trends and cyclical movements, it struggles with precise point forecasting. This limitation arises from the inherently stochastic nature of commodity returns.

These results reinforce the importance of hybrid modeling. By combining econometric filtering with deep learning, the framework leverages the strengths of both approaches while mitigating their individual limitations.

4.8 Rolling Window Backtesting

The robustness of the hybrid model is evaluated using rolling-window backtesting, with results presented in **Table 9** and **Figure 6**. This approach evaluates the temporal stability of the model's performance by iteratively shifting the training and testing horizons across the 2000-2026 dataset, thereby simulating real-world forecasting conditions.

Table 9: Mean Rolling Out-of-Sample R^2 Performance

Commodity	Mean Backtest R^2	Stability Assessment
Crude Oil (WTI)	0.1853	High
Natural Gas	0.1700	Moderate-High
Copper	0.1643	High
Platinum	0.1637	Moderate
Gold	0.1636	High
Silver	0.1595	Moderate

The results in **Table 9** show positive mean out-of-sample R^2 values across all commodities, indicating stable predictive performance over time. Crude Oil (0.1853) and Copper (0.1643) demonstrate particularly strong consistency.

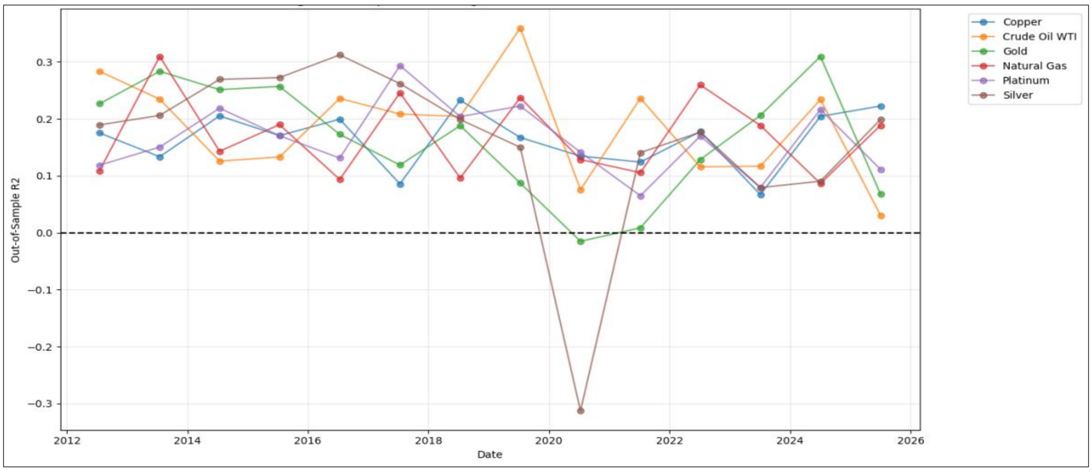

Figure 6: Rolling out-of-sample R^2 performance across commodities

Figure 6 further illustrates the temporal evolution of predictive accuracy. While performance declines during the 2020 crisis, it recovers in subsequent periods, indicating the model's adaptability to changing market conditions. These findings confirm that the hybrid framework is not sample-specific and maintains its effectiveness across different economic regimes.

4.9 Crisis Period Analysis (2020 Shock)

The stress-test results for the 2020 crisis are presented in **Table 10**.

Table 10: Out-of-Sample R^2 during the 2020 Crisis Window

Commodity	2020 OOS R^2	Stress Resilience Category
Platinum	0.1413	High Resilience
Copper	0.1345	High Resilience
Natural Gas	0.1286	Moderate Resilience
Crude Oil (WTI)	0.0755	Moderate Resilience
Gold	-0.0152	Predictive Failure (Marginal)
Silver	-0.3129	Predictive Failure (Significant)

As shown in **Table 10**, industrial commodities such as Copper and Platinum maintain positive predictive performance, while precious metals, particularly Silver, experience significant declines. The 2020 volatility window reveals critical insights into model performance under "black swan" conditions, illustrating how extreme structural shocks disrupt the learning process of hybrid architectures. This divergence suggests that industrial commodities remain structurally predictable even during crises, whereas safe-haven assets are influenced by extreme liquidity shocks and investor sentiment, which are difficult to model. These findings highlight an important limitation of predictive models under extreme conditions and provide valuable insights for risk management strategies.

4.10 Feature Importance and Economic Interpretation

The feature importance results are illustrated in **Figure 7**, with prediction accuracy relationships shown in **Figure 8**.

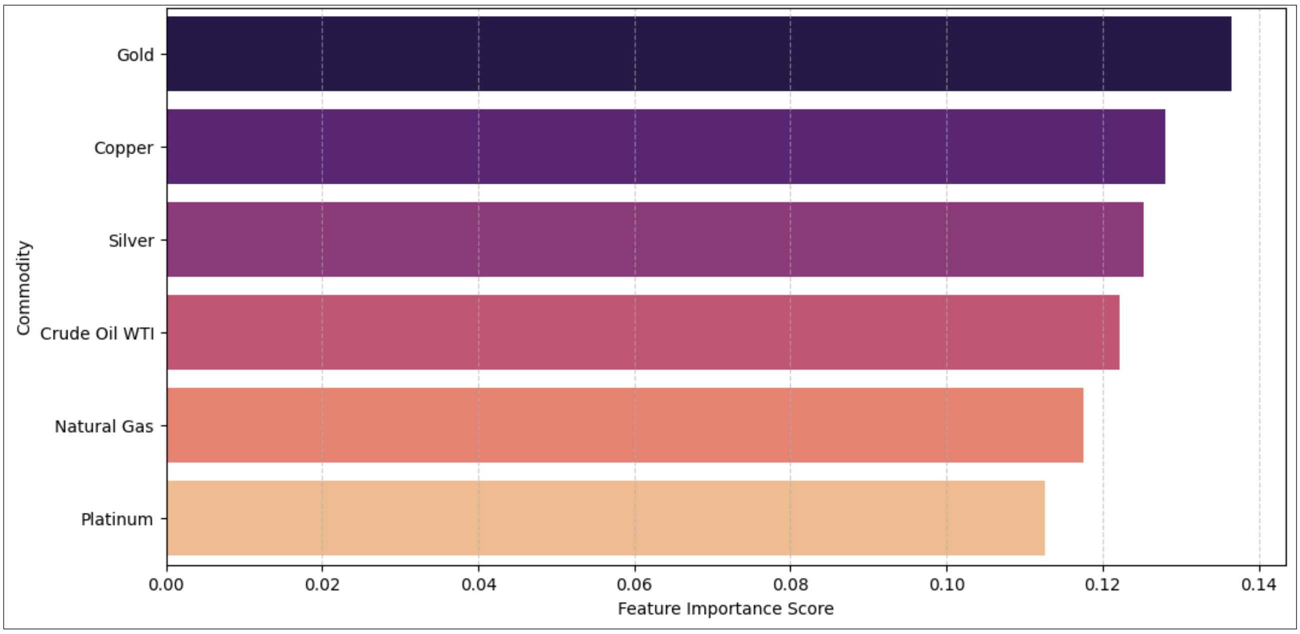

Figure 7: Feature Importance comparison for hybrid commodity forecasting

Figure 7 indicates that volatility-related features and lagged residuals are the most influential predictors. This confirms that risk memory and nonlinear dependencies play a central role in commodity price formation.

Momentum indicators also rank highly, particularly for industrial metals, reflecting the importance of short-term trend dynamics.

The strong contribution of ARIMA residual features validates the hybrid framework's design, demonstrating that residuals contain meaningful predictive information rather than random noise.

The results collectively demonstrate that commodity price dynamics are inherently multi-dimensional, involving linear trends, volatility persistence, and nonlinear interactions. The hybrid econometric and machine learning framework effectively captures these components, resulting in improved forecasting accuracy and robustness.

While the model performs well under normal market conditions, its performance is temporarily affected during extreme structural disruptions such as the 2020 crisis. Nevertheless, the recovery in predictive accuracy during subsequent periods confirms its adaptability.

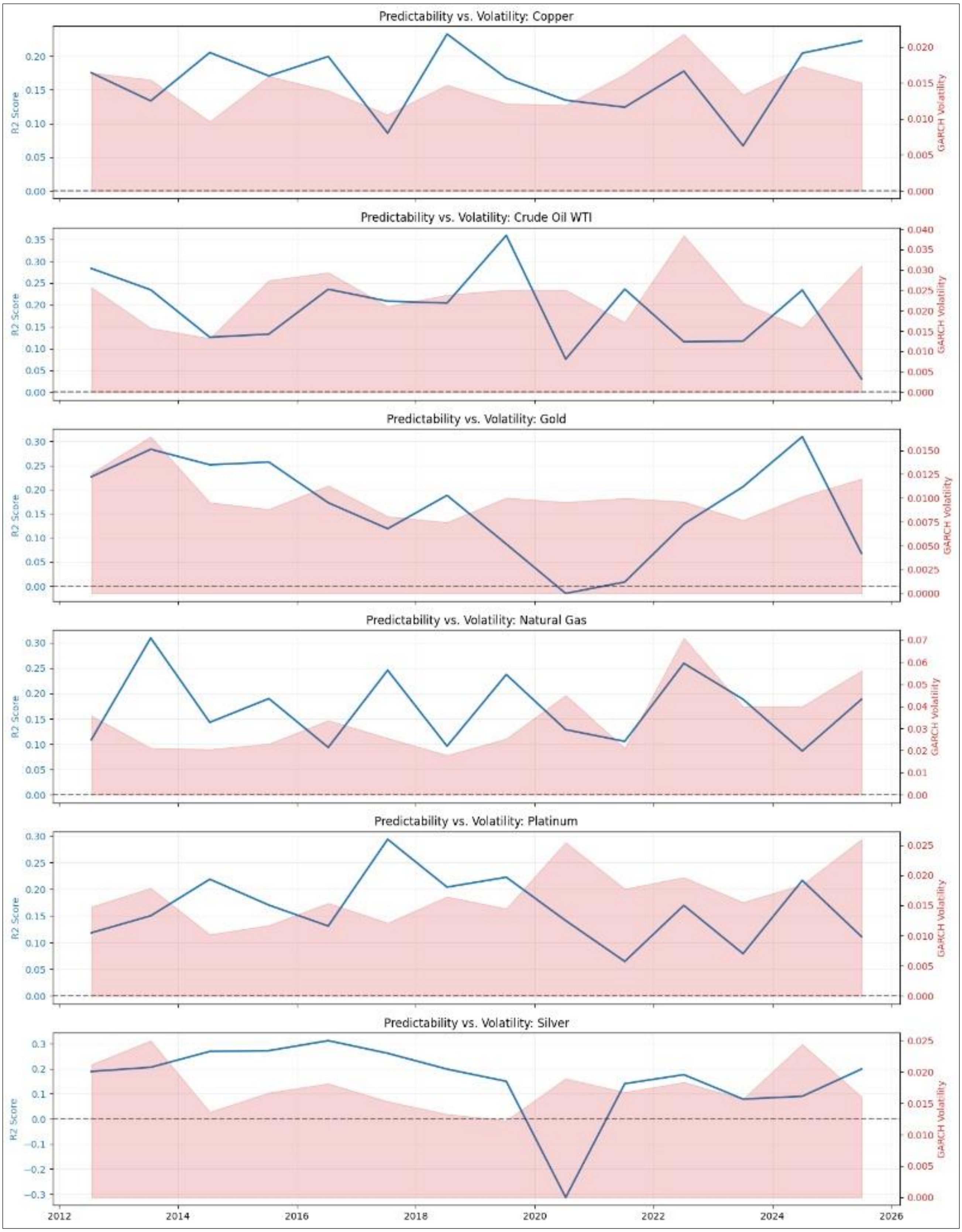

Figure 8: Predicted vs. Actual Commodity Returns Correlation

All in all, the findings provide strong empirical support for the integration of econometric and machine learning techniques in commodity price forecasting, offering a comprehensive and scalable approach for financial analysis and decision-making.

5. CONCLUSION, CONTRIBUTIONS, LIMITATIONS, AND FUTURE RESEARCH

This study proposed a hybrid econometric and deep learning framework for forecasting global commodity price dynamics using six major commodities, namely crude oil (WTI), gold, copper, natural gas, platinum, and silver, over the period 2000-2026. The framework systematically integrates ARIMA for linear dependency extraction, GARCH for volatility persistence estimation, and machine learning algorithms, specifically Random Forest and XGBoost, for nonlinear residual learning. In addition, an LSTM-based deep learning benchmark was incorporated to evaluate sequential temporal learning performance. By combining these complementary modeling paradigms within a unified architecture, the study addresses the long-standing challenge of capturing the multidimensional nature of commodity price behavior.

The empirical evidence confirms that commodity price dynamics are governed by a complex interaction of linear autocorrelation, volatility clustering, and nonlinear residual structures. The proposed hybrid framework consistently outperformed standalone econometric and deep learning models across all commodities, achieving forecasting improvements ranging from 5.51% to 11.96%, alongside stable positive rolling out-of-sample R^2 values. These findings provide strong support for the central premise of the study that econometric residuals should be interpreted as structured nonlinear signals rather than random disturbances. The integration of GARCH-derived conditional variance as a predictive feature further enhanced model sensitivity to risk regimes, enabling superior forecasting performance across both normal and turbulent market conditions.

A major contribution of this study lies in its methodological innovation. Unlike prior studies that primarily rely on either econometric or machine learning approaches in isolation, the present framework unifies linear mean dynamics, second-moment volatility behavior, and nonlinear machine learning residual extraction into a single scalable forecasting pipeline. This integrated design contributes to the growing literature on hybrid financial forecasting systems and demonstrates how econometric rigor and machine learning flexibility can be jointly leveraged to improve predictive reliability.

The study also makes an important empirical contribution by examining a long-horizon multi-commodity dataset that spans multiple structural regimes, including the 2008 financial crisis, the COVID-19 commodity collapse, and the post-pandemic inflationary recovery period. The rolling-window backtesting and crisis-period stress analysis confirm that the proposed framework remains temporally stable and adaptable across regime transitions. Particularly, industrial commodities such as copper and platinum retained positive predictive performance even during the 2020 shock window, highlighting the resilience of the hybrid architecture under extreme market stress.

From a practical and policy perspective, the framework provides a robust decision-support system for investors, commodity traders, central banks, and policy institutions engaged in inflation monitoring, hedging strategies, portfolio allocation, and macro-financial surveillance. The ability to jointly model return predictability and volatility persistence offers valuable insights for stress testing and scenario planning in globally interconnected commodity markets.

Despite these contributions, several limitations remain. First, the current framework primarily relies on historical price-based and volatility-derived features, without explicitly incorporating macroeconomic variables such as interest rates, exchange rates, inflation expectations, global liquidity indicators, or geopolitical risk proxies. Second, although the LSTM benchmark captures broad temporal trends, its weaker point-forecast performance suggests that standalone sequential deep learning models may be less effective in highly noisy return environments. Third, while the study considers multiple commodities, it does not explicitly model dynamic cross-commodity spillovers or network connectedness structures, which may further enrich predictive performance.

Future research may extend this work in several promising directions. The inclusion of macroeconomic and sentiment-based exogenous variables could improve forecasting under structural disruptions. Advanced architectures such as transformer networks, temporal fusion transformers, graph neural networks, and attention-based hybrid systems may further enhance the ability to capture nonlinear dependencies and cross-market contagion. In addition, explainable AI techniques such as SHAP and counterfactual feature attribution can be integrated to strengthen interpretability and policy relevance. These extensions would help move hybrid commodity forecasting systems toward fully explainable and regime-aware intelligent financial decision frameworks.

All things considered, this study demonstrates that hybrid econometric and deep learning architectures provide a robust and scalable solution for modeling complex global commodity price dynamics. By bridging the interpretability of econometric models with the adaptive learning capacity of machine learning systems, the proposed framework offers a significant advancement in commodity forecasting research and establishes a strong foundation for future intelligent financial analytics.

REFERENCES

- Deaton and G. Laroque, “On the Behaviour of Commodity Prices,” *Review of Economic Studies*, vol. 59, no. 1, pp. 1–23, 1992.
- Erten and J. A. Ocampo, “Super Cycles of Commodity Prices Since the Mid-Nineteenth Century,” *World Dev.*, vol. 44, pp. 14–30, 2013, doi: https://doi.org/10.1016/j.worlddev.2012.11.013.
- Baumeister and L. Kilian, “Forecasting the Real Price of Oil in a Changing World: A Forecast Combination Approach,” Jul. 2013. [Online]. Available: https://ssrn.com/abstract=2297208
- Zhang and D. C. Broadstock, “Global financial crisis and rising connectedness in the international commodity markets,” *International Review of Financial Analysis*, vol. 68, no. C, p. None, Apr. 2020, doi: 10.1016/j.irfa.2018.08.003.
- G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, 5th ed. Wiley, 2015.
- J. Frankel, “Effects of speculation and interest rates in a ‘carry trade’ model of commodity prices,” *J. Int. Money Finance*, vol. 42, no. C, pp. 88–112, 2014, [Online]. Available: https://EconPapers.repec.org/RePEc:eee:jimfin:v:42:y:2014:i:c:p:88-112
- J. Frankel, “The Effect of Monetary Policy on Real Commodity Prices,” *Asset Prices and Monetary Policy*, Apr. 2007, doi: 10.7208/chicago/9780226092126.003.0008.
- L. Breiman, “Random Forests,” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.
- M. D. Chinn and O. Coibion, “The Predictive Content of Commodity Futures,” *Journal of Futures Markets*, vol. 34, no. 7, pp. 607–636, Jul. 2014, doi: https://doi.org/10.1002/fut.21615.
- M. Kwas and M. Rubaszek, “Forecasting Commodity Prices: Looking for a Benchmark,” *Forecasting*, vol. 3, pp. 447–459, Apr. 2021, doi: 10.3390/forecast3020027.
- N. Antonakakis, J. Cunado, G. Filis, D. Gabauer, and F. P. de Gracia, “Dynamic connectedness among the implied volatilities of oil prices and financial assets: New evidence of the COVID-19 pandemic,” *International Review of Economics & Finance*, vol. 83, pp. 114–123, 2023, doi: https://doi.org/10.1016/j.iref.2022.08.009.
- P. Cashin, C. J. McDermott, and A. Scott, “Booms and slumps in world commodity prices,” *J. Dev. Econ.*, vol. 69, no. 1, pp. 277–296, 2002, doi: https://doi.org/10.1016/S0304-3878(02)00062-7.
- P. Collier and B. Goderis, “Commodity prices and growth: An empirical investigation,” *Eur. Econ. Rev.*, vol. 56, no. 6, pp. 1241–1260, Apr. 2012, doi: 10.1016/j.euroecorev.2012.04.002.
- Q. Yang, L. Ji, J. Pan, and Y. Zou, “Financial Time Series Forecasting with the Deep Learning Ensemble Model,” *Mathematics*, vol. 11, p. 1054, Feb. 2023, doi: 10.3390/math11041054.
- R. Arezki, P. Loungani, R. van der Ploeg, and A. J. Venables, “Understanding international commodity price fluctuations,” *J. Int. Money Finance*, vol. 42, pp. 1–8, 2014, doi: https://doi.org/10.1016/j.jimonfin.2013.08.002.
- R. S. Pindyck and J. J. Rotemberg, “The Excess Co-movement of Commodity Prices,” *Economic Journal*, vol. 100, no. 403, pp. 1173–1189, Dec. 1990, doi: None.
- S. Gu, B. T. Kelly, and D. Xiu, “Empirical Asset Pricing via Machine Learning,” Sep. 2019. doi: 10.2139/ssrn.3159577.
- S. K. Chejarla, S. K. Duvvuri, B. Rajana, P. T. Dakey, K. R. Rao, and M. R. Goutham, “Forecasting and Explaining Non-Performing Assets in Indian Banking Groups Using Panel Econometrics and Explainable Machine Learning Models,” *Journal of Global Economics, Management and Business Research*, vol. 18, no. 2, pp. 58–78, 2026, doi: 10.56557/jgembr/2026/v18i210427.
- S. K. Chejarla, S. K. Duvvuri, B. Rajana, P. T. Dakey, R. K. Rao, and M. R. Goutham, “Geopolitical shocks and financial contagion in the Indian equity market during the Iran-Israel conflict using advanced econometric and machine learning methods,” *International Journal of Finance and Management Economics*, vol. 9, no. 3, pp. 184–197, 2026, doi: 10.33545/26179210.2026.v9.i3.785.
- S. Mullainathan and J. Spiess, “Machine Learning: An Applied Econometric Approach,” *Journal of Economic Perspectives*, vol. 31, no. 2, pp. 87–106, May 2017, doi: 10.1257/jep.31.2.87.
- S. Nazlioglu, C. Erdem, and U. Soytas, “Volatility spillover between oil and agricultural commodity markets,” *Energy Econ.*, vol. 36, no. C, pp. 658–665, Apr. 2013, doi: 10.1016/j.eneco.2012.11.009.
- T. Bollerslev, “Generalized autoregressive conditional heteroskedasticity,” *J. Econom.*, vol. 31, no. 3, pp. 307–327, 1986, doi: https://doi.org/10.1016/0304-4076(86)90063-1.

- T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, in KDD ’16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 785–794. doi: 10.1145/2939672.2939785.
- T. Fischer and C. Krauss, “Deep learning with long short-term memory networks for financial market predictions,” *Eur. J. Oper. Res.*, vol. 270, no. 2, pp. 654–669, 2018, doi: https://doi.org/10.1016/j.ejor.2017.11.054.