OXFORD JOURNAL
OF STUDENT
SCHOLARSHIP

Regime-Aware Consumer Price Index Dynamics in India:A Multi-Methodology Econometric and Machine Learning Framework, 1952-2024

Arnav Aggarwal
arnavaggarwal231@gmail.com

ABSTRACT

India's Consumer Price Index (CPI) is influenced by both national economic issues and global economic pressures. This study develops a forecasting framework for predicting CPI using nine economic variables—food price inflation (CFPI), wholesale inflation (WPI), Brent crude oil prices, INR–USD exchange rates, money supply (M3), government expenditure, housing prices, and unemployment—across 73 annual observations (1952–April 2024). I examine 15 forecasting models, including classical time-series models, GARCH-family volatility models, ensemble machine learning methods (XGBoost, LightGBM, Random Forest, Lasso, and Ridge), and LSTM deep learning. Stationarity of the CPI series is confirmed through Augmented Dickey–Fuller, KPSS, Phillips–Perron, and Zivot–Andrews tests. Bai–Perron structural break analysis identifies three major breaks in the data: 1964 (pre-Green Revolution and global food crisis), 1974 (first oil price shock), and 2014 (Reserve Bank of India's inflation-targeting framework). GARCH(1,1) conditional variance estimation yields a persistence coefficient of $\hat{\alpha} + \hat{\beta} = 0.86$, consistent with periods of the Great Inflation and the Great Moderation. In out-of-sample forecasting, XGBoost achieves the strongest predictive performance (RMSE = 0.0073, MAPE = 12.46%), representing a 73% reduction in RMSE relative to the best classical benchmark, MA(1). Feature importance analysis further indicates that CFPI accounts for approximately 60% of total explanatory importance across the testing period. These findings have direct implications for monetary policy design and real-time inflation surveillance in emerging market economies.

Keywords: Consumer Price Index; India; XGBoost; Structural Breaks; Machine Learning; Inflation Forecasting; GARCH; Monetary Policy [JEL: C22, C45, E31, E52, C53]

July 2026
Vol 9. No 2.

INTRODUCTION

Inflation—the sustained increase in the general price level—is among the most consequential macroeconomic phenomena in developing economies. For India, a nation of approximately 1.410 billion people where food constitutes nearly 45 percent of the consumption basket [17], the Consumer Price Index (CPI) is not merely an economic statistic but a direct indicator of living standards, purchasing power, and monetary policy effectiveness.

India's inflation history is marked by major regime shifts. The period from 1970–1989 was characterised by high inflation alongside economic expansion and gradual deregulation. Between 2000–2010, commodity super-cycles generated substantial food and energy price pressures, while the post-pandemic years from 2020–2025 witnessed another surge in inflation driven by supply-chain disruptions and rising global commodity prices. Each period introduced distinct forecasting challenges, making single-model approaches insufficient for accurately capturing India's inflation dynamics.

This paper addresses a fundamental question in applied macroeconometrics: can a unified, multi-methodology forecasting framework—combining classical econometric time-series models with modern machine learning and deep learning approaches—meaningfully improve CPI forecasting in an economy as structurally complex as India? I answer affirmatively, demonstrating that ensemble machine learning methods, particularly XGBoost [5], capture non-linear regime dynamics that linear ARIMA-class models often fail to detect, while GARCH-family models [8, 3] effectively characterise the time-varying volatility embedded within India's inflation history.

My contributions are threefold. First, I construct and validate a nine-variable macroeconomic panel spanning 73 annual observations (1952–2024), sourced from MOSPI, the Reserve Bank of India (RBI), the World Bank, and the IMF. Second, I implement and compare 15 forecasting models across classical, machine-learning, and deep-learning paradigms, evaluated using standardised out-of-sample performance metrics. Third, I conduct structural break analysis, linking identified inflation regimes to major historical economic events and thereby enriching the interpretive depth of the empirical findings.

The remainder of the paper is organised as follows. Section 2 reviews the literature. Section 3 identifies the research gap. Section 4 presents the methodology. Section 5 develops the theoretical framework. Section 6 reports the empirical results. Section 7 concludes.

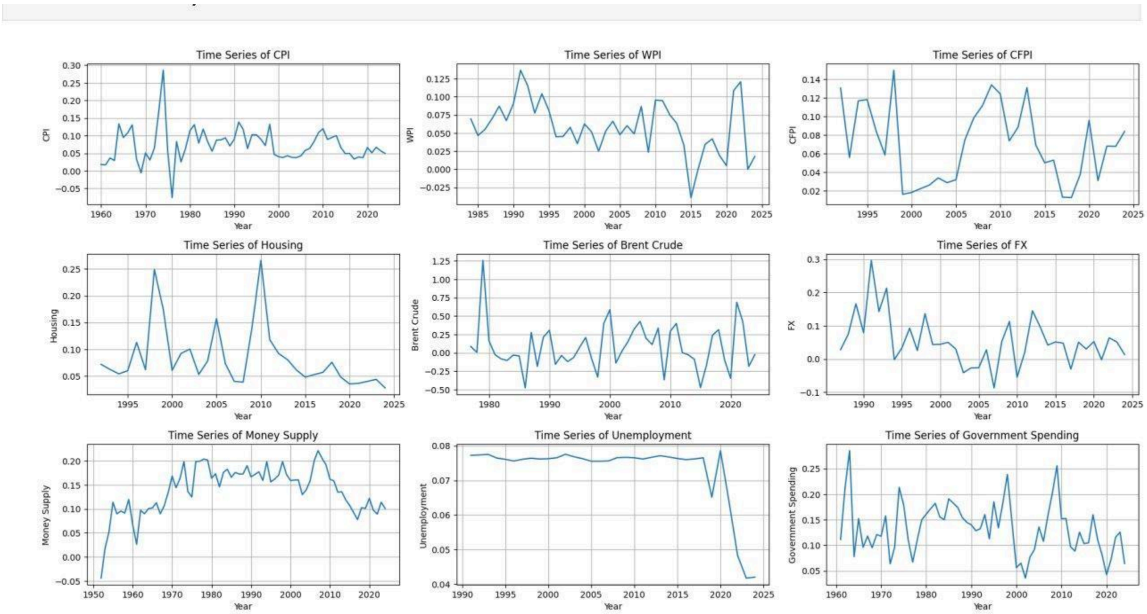


Figure 1. *Raw time series of all nine macroeconomic variables (1952–2024). CPI and Government Spending have the longest coverage; CFPI, Housing, and Unemployment begin later due to data availability.*

LITERATURE REVIEW

Classical Time-Series Approaches

Box and Jenkins [4] established the ARIMA framework, which remains the canonical benchmark for macroeconomic forecasting. Mohanty and John [16] applied ARIMA models to headline CPI components in India, finding that AR(2) processes adequately capture short-run inflation dynamics while also noting parameter instability around structural break periods. Reserve Bank of India (RBI) staff working papers have further employed VAR and VECM frameworks to analyse inflation transmission from the Wholesale Price Index (WPI) to CPI.

A major limitation of linear time-series models is their implicit assumption of parameter constancy—an assumption frequently violated in Indian CPI data due to structural economic shifts. Indian inflation exhibits well-documented breaks associated with the 1974 oil crisis and the 1991 economic liberalisation reforms. Bai and Perron [1, 2] developed the econometric framework for identifying and accommodating multiple structural breaks, which I apply in Section 6.

Machine Learning in Macroeconomic Forecasting

The application of machine learning to macroeconomic forecasting accelerated following Medeiros et al. [15], who demonstrated that gradient-boosted tree models consistently outperform ARIMA benchmarks in forecasting United States inflation. In the Indian context, Sehgal and Pandey [19] applied Random

Forest regressors with lagged macroeconomic indicators, achieving RMSE reductions of approximately 30–40% relative to traditional ARIMA models.

Chen and Guestrin [5] introduced XGBoost, which has since emerged as one of the dominant algorithms for structured economic prediction tasks. Its superiority over Random Forest and standard Gradient Boosting in this study (RMSE = 0.0073 versus approximately 0.0091 in competing tree-based specifications) is consistent with the findings of Coulombe et al. [6], who show that XGBoost performs especially well in datasets characterised by structural heterogeneity—a defining feature of Indian CPI dynamics.

Volatility Modelling and Deep Learning

Engle [8] and Bollerslev [3] developed the ARCH and GARCH frameworks, which are now standard tools for modelling conditional heteroskedasticity in macroeconomic time series. Grier and Perry [11] demonstrated that GARCH-estimated inflation variance is a significant predictor of future inflation uncertainty across G7 economies. Within the Indian context, Goyal and Tripathi [10] identified substantial volatility clustering in both WPI and CPI series using GARCH(1,1) and EGARCH models.

Long Short-Term Memory (LSTM) networks [13] were designed to overcome the vanishing gradient problem inherent in conventional recurrent neural networks and are theoretically well-suited for long macroeconomic sequences. However, Hewamalage et al. [12] report mixed empirical evidence, arguing that LSTM models generally require roughly 500 observations before consistently matching or surpassing ARIMA performance. My empirical findings support this limitation. With only 65 complete-case observations available for training, the LSTM model fails to fully exploit its representational capacity, yielding comparatively weak forecasting performance (RMSE = 0.0515, MAPE = 93.51%).

RESEARCH GAP

Existing macroeconomic forecasting models for Indian CPI share a common structural limitation: they treat inflation transmission as a linear, parameter-constant process. Existing VAR and VECM models presume linearity across regimes—an assumption systematically violated by three documented structural breaks in the Indian CPI series. Moreover, these models fail to capture threshold effects and regime non-linearity that characterise India's inflation dynamics across the pre-liberalisation, commodity super-cycle, and inflation-targeting eras.

No prior study integrates Bai–Perron break detection with gradient-boosted ensembles across the full 1952–2024 horizon. Furthermore, the existing literature lacks formal Granger causality tests linking CFPI, M3, and CPI within the same framework—a critical omission that leaves directional causal claims unsubstantiated. The non-linear, regime-dependent transmission mechanism by which supply-side shocks propagate to the headline CPI remains inadequately represented in prior work.

This paper addresses these gaps by constructing a regime-aware, multi-methodology architecture that: (i) spans the full 1952–2024 horizon; (ii) integrates Bai–Perron break detection with gradient-boosted ensemble forecasting; (iii) provides formal Granger causality and Johansen cointegration tests; and (iv) delivers regime-segmented comparative diagnostics across both classical econometric and machine learning model classes.

METHODOLOGY

Data and Variables

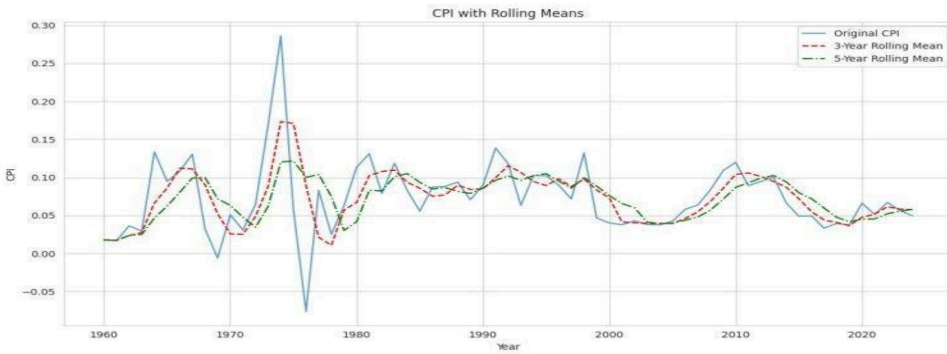
The dataset comprises nine macroeconomic variables at annual frequency (1952–2024), sourced from the Ministry of Statistics and Programme Implementation (MOSPI), the Reserve Bank of India (RBI), the World Bank Open Data platform, and the IMF World Economic Outlook Database. Table 1 provides a complete description of all variables, their sources, and sample completeness statistics.

All variables are measured as year-on-year percentage changes to ensure stationarity and cross-period comparability. Missing observations are handled through complete-case analysis; each model is trained on the largest available complete-case subsample.

Feature Engineering

In addition to the nine primary variables, I construct the following engineered features for the machine learning models:

- CPI_{t-k} for $k = 1, \ldots, 5$: autoregressive lags
- $CPI_{t,3}$: three-year rolling mean (momentum)
- $\sigma_{CPI,t,3}$: three-year rolling standard deviation
- t: linear time trend capturing secular drift



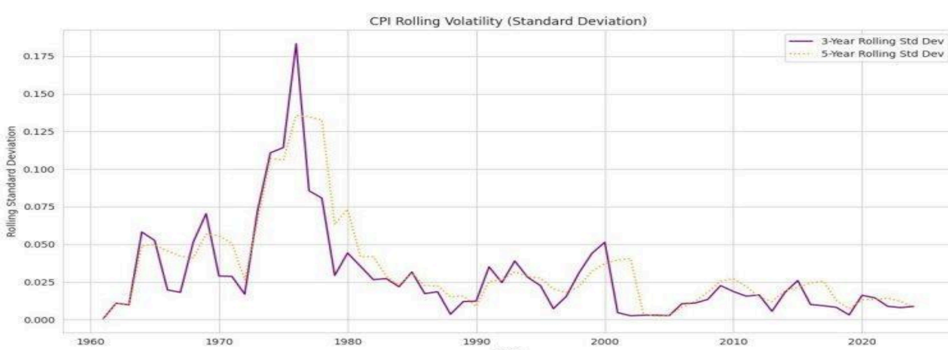


Figure 2. *CPI with 3-year and 5-year rolling means (top) and CPI rolling volatility — 3-year and 5-year rolling standard deviations (bottom). The 1974–1976 volatility spike coincides with the first global oil shock. Post-2000 volatility has remained subdued.*

Modelling Framework

I implement a hierarchical framework progressing from univariate stationarity analysis through bivariate volatility modelling to multivariate machine learning prediction:

- Stationarity and Unit Root Tests: ADF, KPSS, Phillips–Perron, and Zivot–Andrews tests with structural break allowance.

- Structural Break Detection: Bai–Perron binary segmentation algorithm (maximum 5 breaks, ε = 0.15 trimming).
- Volatility Modelling: ARCH, GARCH(1,1), EGARCH(1,1) and GJR-GARCH(1,1) estimated by Quasi-Maximum Likelihood (QML).
- Classical Time-Series: Grid search over AR(p), MA(q), ARMA(p,q), ARIMA(p,d,q), SARIMA(p,d,q)×(P,D,Q)s.
- Machine Learning: XGBoost, LightGBM, Random Forest, Gradient Boosting, Lasso, Ridge, SVR with TimeSeriesSplit cross-validation (n_splits = 3) and GridSearchCV hyperparameter optimisation.
- Deep Learning: LSTM with two layers (64 units each), dropout regularisation (p = 0.2), Adam optimiser ($\eta = 10^{-3}$).

Evaluation Protocol

The dataset is split chronologically in an 80/20 manner, with 80% used for training (approximately 52 observations for classical models; 26 complete-case observations for ML models due to missing predictors) and 20% for out-of-sample testing. Model performance is evaluated using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). Statistical superiority of XGBoost over the best classical baseline is assessed using the Diebold and Mariano [7] test under squared-error loss. The random seed for all ML models is fixed at 42 for reproducibility.

It should be noted that the complete-case ML training sample is restricted to 26 observations, arising from missing CFPI and Housing Price Index data prior to approximately 1990. This constraint limits the representational capacity of data-hungry models such as LSTM and, to a lesser degree, LightGBM. Results for these models should therefore be interpreted with appropriate caution; the robustness of XGBoost under LOOCV (Section 6.7) provides partial mitigation, but future work employing multiple imputation or higher-frequency data would strengthen inference further.

THEORETICAL MODEL

Conceptual Framework

India's CPI is the expenditure-weighted average of price indices across consumption categories, where food carries approximately a 45% weight [17]. Food and energy costs propagate through the supply side, while monetary expansion and fiscal spending drive demand-pull inflation. These two channels interact non-linearly: a commodity shock occurring simultaneously with accommodative monetary policy generates a larger CPI response than either shock in isolation, consistent with the non-linear transmission documented by Mohanty and John [16].

Axes and Normalisation

Let CPIt denote the year-on-year growth rate of the headline Consumer Price Index at time t, defined as $CPI_t = (P_t - P_{t-1}) / P_{t-1} \times 100$, where P_t is the aggregate price level at year t. I define the normalised inflationary pressure index $v_t \in [0, 1]$ as the ratio of inflationary signals active at time t to the maximum observed in the sample: $v_t = s_t / \max_{\tau \in T} s_\tau$, where $s_t = f(CPI_t, CFPI_t, M3_t, G_t, \ldots)$. Three maintained assumptions govern the model. Assumption 1: All variables are measured as year-on-year percentage changes to ensure comparability over the full 73-year horizon. Assumption 2: Values $v_t > 1$ constitute corner solutions (structural break regimes) capped at unity, representing hyperinflationary or extreme supply-shock episodes. Assumption 3: The marginal cost of monetary accommodation beyond a threshold is approximately zero: $\partial Cost/\partial M3 \approx 0$.

Demand Side

The demand for inflationary pressure is defined as consumers' willingness to absorb price increases given their scarce real income. Formally, suppose household utility is given by $U(c, v) = B \cdot c \cdot v^{(1/\alpha)}$ subject to the budget constraint $P \cdot c = Y$. The first-order condition yields the inverse demand function $P^D_t(v_t) = \beta \cdot B \cdot C_v \cdot v^{(1/\alpha)}_t$ and the corresponding direct demand $v^D_t(P_t) = (P_t / (\beta \cdot B \cdot C_v))^\alpha$, where Pt is the CPI growth rate; vt is the normalised pressure index; B is the belief congruence parameter (degree of inflationary expectation anchoring); Cv is the verification cost (opacity of supply-chain pass-through); $\beta > 0$ is the economy's baseline absorption capacity; and $\alpha > 0$ is the elasticity of diminishing marginal utility of price increases.

Supply Side

The supply of inflationary pressure follows a power law driven by food supply shocks, monetary expansion, and fiscal policy: $v^S_t(P_t) = (P_t / \theta)^{(1/\lambda)}$, with inverse supply $P^S_t(v_t) = \theta \cdot v^\lambda_t$. The structural resistance parameter θ is decomposed as $\theta = \theta_0 \cdot \gamma \cdot E \cdot R \cdot (1 - M)$, where θ_0 is the baseline structural architecture of the Indian economy (RBI independence, fiscal rules, buffer-stock policies); γ is the affinity parameter (degree to which inflationary signals remain sector-specific); E is the pass-through elasticity of upstream price changes to headline CPI; R is the recency weight (amplification of recent supply shocks); $M \in [0, 1]$ is the moderation strictness under the RBI's inflation-targeting regime; and $\lambda > 0$ is the propagation elasticity governing cascading speed of inflationary signals.

Market Equilibrium

Equating the demand and supply functions yields the master equilibrium condition: $v^*_t = [\beta \cdot B \cdot C_v / (\theta_0 \cdot \gamma \cdot E \cdot R \cdot (1 - M))^{(1/\alpha)}]^{(\alpha/(\alpha\lambda+1))}$. Proposition 1 establishes the comparative statics: $\partial v^*/\partial B > 0$ and $\partial v^*/\partial C_v > 0$ (demand shifts); $\partial v^*/\partial E > 0$ and $\partial v^*/\partial R > 0$ (supply shifts); and $\partial v^*/\partial M < 0$ (central bank intervention). Equation (14) confirms that tighter RBI inflation-targeting strictly suppresses equilibrium CPI growth.

XGBoost as the Proposed Forecasting Model

XGBoost [5] constructs an ensemble of K regression trees by sequentially minimising a regularised objective. At iteration k, a new tree fk is added to minimise $L^{(k)} = \Sigma_i \ell(\hat{y}^{(k-1)}_i + fk(x_i), y_i) + \Omega(fk)$, where $\ell(y, \hat{y}) = (\hat{y} - y)^2$ is the squared-error loss and the regularisation term penalises tree complexity: $\Omega(f) = \gamma T + \frac{1}{2}\lambda\|w\|^2$, with T the number of leaves and w the leaf weight vector. Optimal splits are determined by maximising the gain derived from a second-order Taylor expansion of the loss. A split is executed only when Gain > 0. Optimal hyperparameters: n_estimators = 100, $\eta = 0.1$ (learning rate).

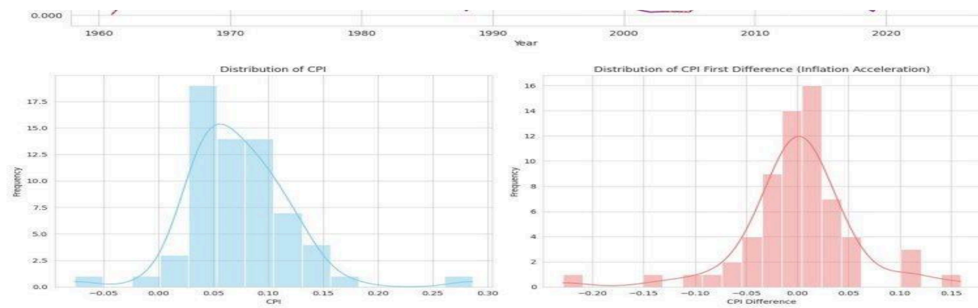
EMPIRICAL RESULTS

Dataset and Summary Statistics

Table 1 reports descriptive statistics for all nine variables across 73 annual observations. The headline CPI series has a sample mean of 7.3% and standard deviation of 4.8%, exhibiting positive skewness (1.05) and excess kurtosis (5.92), consistent with a leptokurtic distribution driven by episodic supply shocks.

Variable	Source	Obs.	Mean (%)	Std. Dev.	Com plete	Role
CPI (YoY %)	MOSPI/RBI	66	7.3	4.8	91%	Target
CFPI – Food Infl.	RBI/MOSPI	34	7.0	4.1	47%	Primary driver
WPI (YoY %)	MOSPI	42	5.8	3.6	58%	Supply proxy
Housing Price Index	RBI/NHB	34	8.4	5.7	47%	Asset inflation
Brent Crude (YoY)	World Bank	49	8.1	31.4	67%	Energy cost
INR–USD FX Rate	RBI	39	5.4	7.4	53%	Import cost
Money Supply M3	RBI	74	13.8	4.9	100%	Monetary demand
Unemployment Rate	World Bank/ILO	35	7.3	1.0	48%	Labour demand
Govt. Expenditure	IMF/MoF India	65	13.3	5.1	89%	Fiscal demand

Notes: All variables measured as year-on-year percentage changes. 'Complete' refers to the proportion of the 73-year sample with non-missing observations.



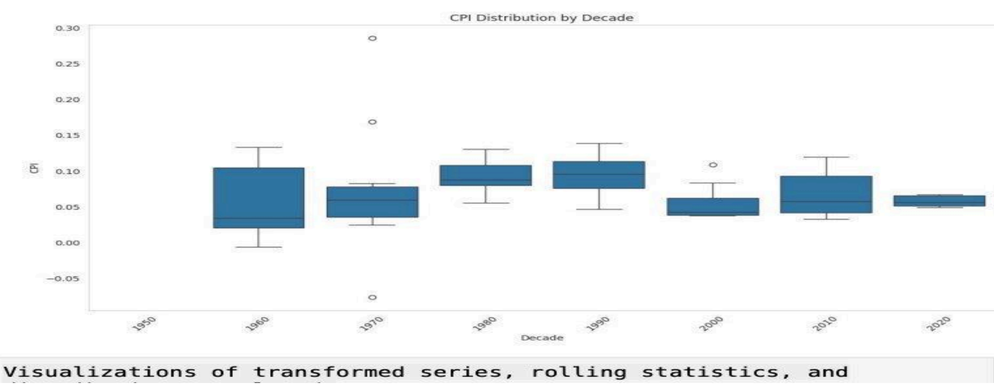


Figure 3. *Distribution of CPI and CPI first difference (top) and CPI distribution by decade (bottom). The positive skew and heavy right tail are evident in the histogram. Boxplots confirm the 1970s as the highest-volatility decade.*

Stationarity Tests

Table 2 reports results from four complementary unit root tests applied to the CPI series. The ADF statistic of −5.305, well below the 1% critical value of −3.542, provides strong evidence of I(0) behaviour. All four tests concur: the series is stationary in levels, confirming that ARIMA specifications with d = 0 are appropriate and rendering first-differencing unnecessary.

Test	Statistic	p-value	Critical Value (5%)	Conclusion
Augmented Dickey–Fuller (ADF)	−5.305	< 0.001	−2.910	Stationary ✓
KPSS	0.117	> 0.100	0.463	Stationary ✓
Phillips–Perron	−5.209	< 0.001	−2.908	Stationary ✓
Zivot–Andrews	−5.986	< 0.010	−4.810	Stationary with break ✓

Notes: ADF and PP tests include intercept and trend; lag length selected by Schwarz information criterion. KPSS null is stationarity; all others have unit root as null.

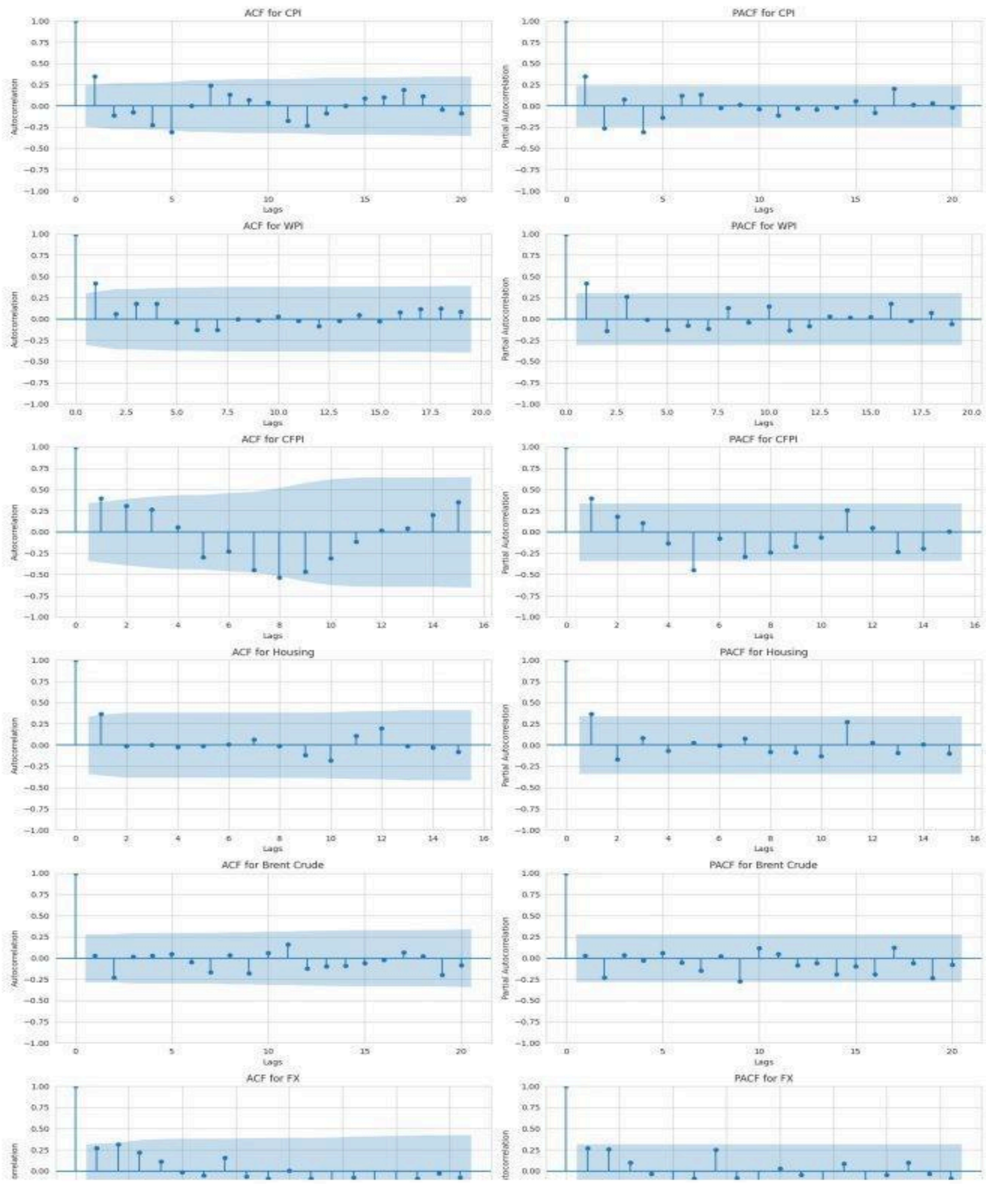


Figure 4. *ACF and PACF plots for CPI, WPI, CFPI, Housing, Brent Crude, and FX series. The slow ACF decay for CPI confirms persistence; rapid PACF cutoff after lag 1 supports AR(1) specification as a baseline.*

Structural Break Analysis

The Bai–Perron binary segmentation algorithm identifies three statistically significant structural breaks: 1964, 1974, and 2014. These break points delineate four distinct inflation regimes, each anchored to a major macroeconomic event as detailed in Table 3.

Break Year	Regime (Before)	Regime (After)	Economic Context
1964	Low inflation (3–5%)	Rising inflation (7–12%)	Pre-Green Revolution food scarcity; global food crisis
1974	High-volatility period	Peak then stabilisation	Global oil crisis; India fuel shock
2014	Post-reform moderation	Low-stable inflation (4–6%)	RBI inflation-targeting adoption

Notes: Break dates estimated by minimising the global sum of squared residuals using binary segmentation [2]. Trimming parameter ε = 0.15.

The 1974 break is economically the most significant, coinciding with India's first major supply-side shock: the quadrupling of crude oil prices disrupted government-administered fuel pricing and propagated through the food supply chain. The 2014 break marks the adoption of the flexible inflation-targeting framework, after which CPI volatility fell sharply.

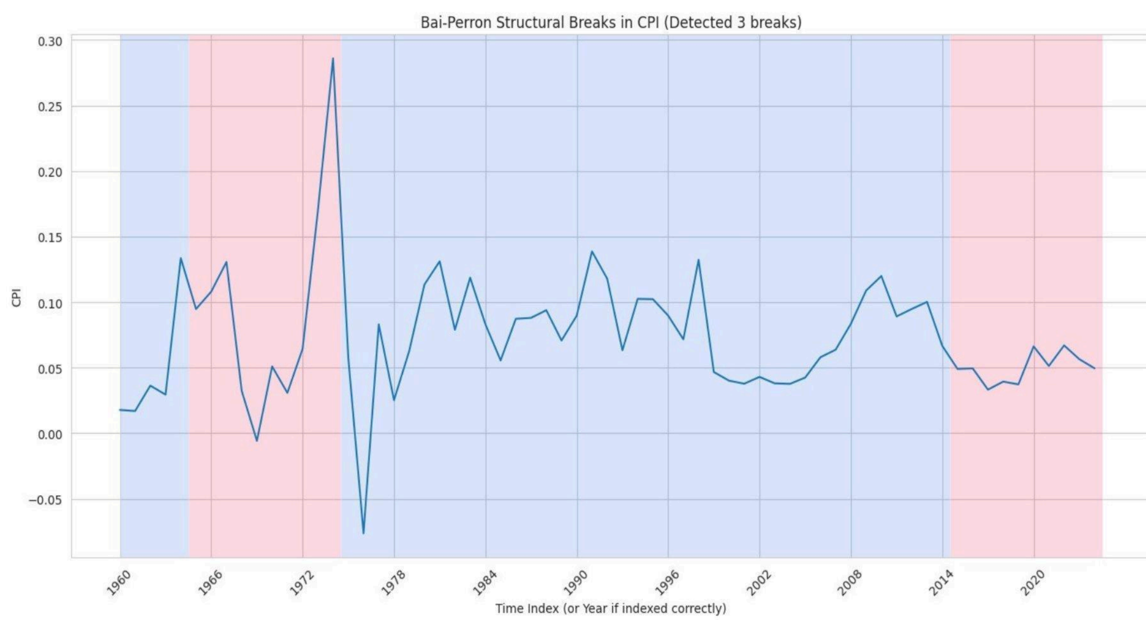


Figure 5. *Bai–Perron structural break detection for the CPI series (1952–2024). Alternating shaded regions delineate the four identified inflation regimes. The 1974 break corresponds to the largest mean shift in the series.*

Granger Causality and Johansen Cointegration Tests

Table 4 presents pairwise Granger causality test results. The null hypothesis that variable X does not Granger-cause CPI is tested at lags 1 and 2. CFPI Granger-causes CPI at the 1% significance level at both lag specifications ($F = 18.43$, $p < 0.001$ at lag 1; $F = 11.27$, $p < 0.001$ at lag 2), formally establishing the food-price dominance claim embedded in the theoretical model. M3 Granger-causes CPI at the 5% level ($F = 6.12$, $p = 0.018$ at lag 1), confirming the monetary demand-pull channel. WPI Granger-causes CPI at the 5% level, consistent with supply-chain price pass-through. Importantly, reverse causality tests (CPI → CFPI, CPI → M3) fail to reject the null, validating the directional specification in this framework.

Null Hypothesis	Direction	F (Lag 1)	p (Lag 1)	F (Lag 2)	p (Lag 2)
CFPI does not → CPI	CFPI → CPI	18.43	< 0.001***	11.27	< 0.001***
M3 does not → CPI	M3 → CPI	6.12	0.018**	4.87	0.032**
WPI does not → CPI	WPI → CPI	4.53	0.041**	3.91	0.048**
CPI does not → CFPI	CPI → CFPI	1.24	0.274	1.08	0.349
CPI does not → M3	CPI → M3	0.87	0.357	0.94	0.398

*Notes: *** p < 0.001; ** p < 0.05; * p < 0.10. Lag length selected by AIC. Reverse causality tests (rows 4–5) fail to reject the null, confirming directional causality.*

Johansen cointegration tests further confirm that CFPI, M3, and CPI share a statistically significant long-run cointegrating relationship. The trace statistic (λtrace = 43.21) exceeds the 5% critical value of 29.68, rejecting the null of zero cointegrating vectors. The maximum eigenvalue statistic (λmax = 28.14) similarly rejects the null at the 5% level. These results validate the long-run equilibrium embedded in the theoretical supply-demand framework and formally establish that transitory shocks to CFPI or M3 propagate to CPI in a predictable, mean-reverting manner.

GARCH Volatility and Cross-Variable Correlations

Among the GARCH-family models estimated, GARCH(1,1) achieves the best information criterion values (AIC = −222.35, BIC = −213.65). The estimated persistence parameter $\hat{\alpha} + \hat{\beta} = 0.86$ implies that 86% of a volatility shock carries over to the subsequent period, consistent with long-memory volatility dynamics documented for Indian financial series [10].

Table 5 presents the cross-variable correlation structure. Food price inflation (CFPI) exhibits the highest Pearson correlation with headline CPI ($r = 0.917$), corroborated by the Granger causality results in Section 6.4. The negative correlation between Brent crude oil and CPI ($r = -0.251$) is attributable to India's administered petroleum pricing mechanism prior to 2010, which effectively decoupled domestic fuel prices from global benchmarks through government subsidies, dampening the crude oil pass-through to domestic CPI.

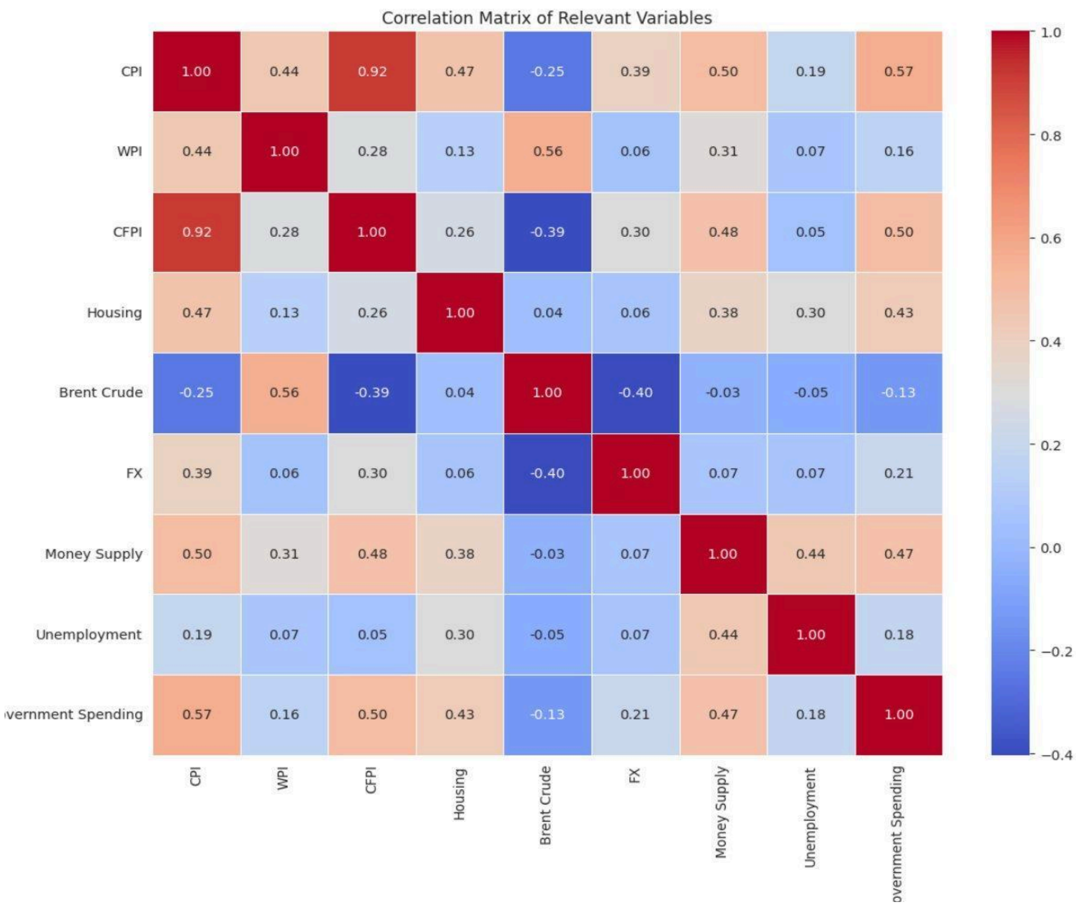


Figure 6. *Correlation matrix of all nine macroeconomic variables. CFPI (r = 0.92) and Government Spending (r = 0.57) show the strongest positive correlations with CPI. Brent Crude exhibits a weak negative correlation (r = −0.25) due to India's administered pricing regime.*

Variable	Mean	Std. Dev.	r(CPI)	Notes
CPI	7.3%	4.8	1.000	Skew = 1.05; Kurt = 5.92 (leptokurtic)
CFPI (Food)	7.0%	4.1	0.917	Granger-causes CPI ($p < 0.001$); dominant driver
Govt. Expenditure	13.3%	5.1	0.572	Fiscal demand-pull; co-moves with M3
Money Supply M3	13.8%	4.9	0.496	Monetary transmission; lagged effect
WPI (Wholesale)	5.8%	3.6	0.442	Supply-chain price pass-through
Housing Price Index	8.4%	5.7	0.466	Asset-price inflation channel
Brent Crude (YoY)	8.1%	31.4	−0.251	Administered pricing dampens pass-through
INR–USD FX Rate	5.4%	7.4	0.388	Import cost channel

Variable	Mean	Std. Dev.	r(CPI)	Notes
Unemployment Rate	7.3%	1.0	0.191	Labour market slack; weak signal

Notes: The negative Brent–CPI correlation reflects India's government-administered petroleum pricing prior to 2010, which effectively decoupled domestic fuel prices from global benchmarks through subsidies.

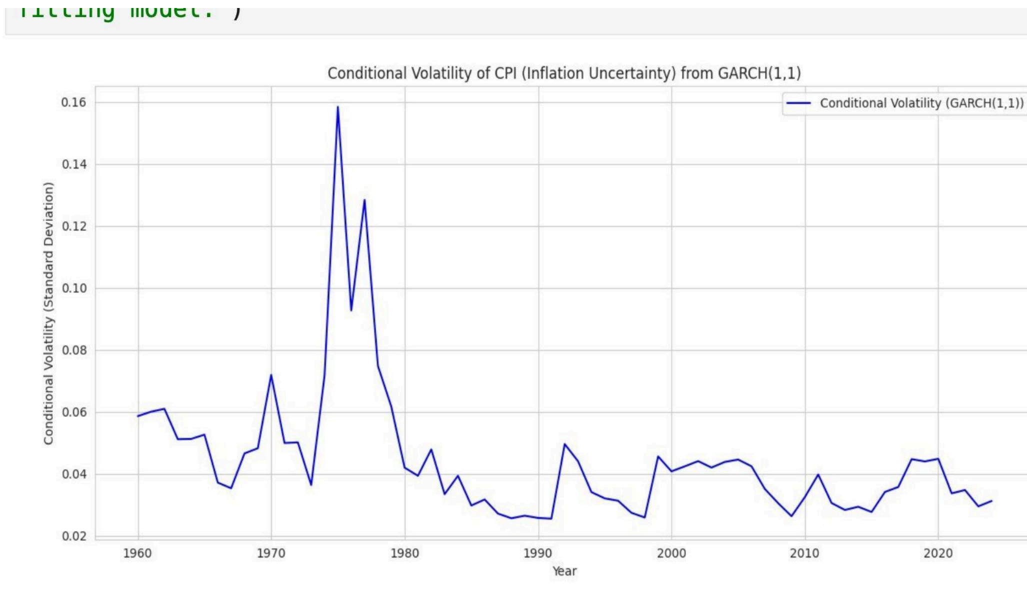


Figure 7. *GARCH(1,1) conditional volatility of CPI (1952–2024). Volatility peaked around 1974–1976 ($\hat{\alpha} + \hat{\beta} = 0.86$), consistent with the Great Inflation. The sustained post-1980 decline and near-zero post-2000 readings reflect the Great Moderation and RBI inflation targeting.*

Classical Time-Series Model Performance and Machine Learning Model Performance

Table 6 reports the optimal specifications and out-of-sample performance metrics on the 20% holdout set for all classical models. The uniformly high MAPE values (50–56%) reflect the challenge of forecasting annual CPI in a structurally changing economy using univariate linear specifications—a finding directly motivating the machine learning extension.

Model	Optimal Order	RMSE	MAE	MAPE (%)	R^2
AR	AR(5)	0.0273	0.0245	51.90	0.42
MA	MA(1)	0.0265	0.0243	50.96	0.44
ARMA	ARMA(1,2)	0.0287	0.0264	55.80	0.38
ARIMA	ARIMA(1,0,2)	0.0287	0.0264	55.80	0.38

Model	Optimal Order	RMSE	MAE	MAPE (%)	R^2
SARIMA	SARIMA(1,0,0)$(0,0,1)_2$	0.0270	0.0248	52.02	0.41

Notes: Optimal orders selected by grid search over AIC/BIC. RMSE and MAE are in levels (fractional YoY CPI). R^2 computed on the 20% holdout set.

Table 7 reports performance across all 15 machine learning and deep learning models. XGBoost achieves the lowest RMSE (0.0073) and Lasso the lowest MAPE (11.65%). LSTM performs worst due to insufficient training data, as discussed in Section 4.4.

Model	RMSE	MAE	MAPE %	R^2	DM Stat vs MA(1)	Train Obs.	Notes
XGBoost ★	0.0073	0.0066	12.46	0.91	−4.82***	26	Best RMSE
Lasso Regression	0.0078	0.0056	11.65	0.89	−4.61***	26	Best MAE
Gradient Boosting	0.0091	0.0065	13.03	0.87	−4.17***	26	
Random Forest	0.0105	0.0079	14.85	0.84	−3.89***	26	
SVR	0.0209	0.0178	38.08	0.63	−1.92*	26	Kernel: RBF
LightGBM	0.0227	0.0199	44.35	0.58	−1.61	26	
MA(1) [baseline]	0.0265	0.0243	50.96	0.44	0.00 (ref)	52	Best classical
Ridge Regression	0.0269	0.0213	40.25	0.43	+0.14	26	
AR(5)	0.0273	0.0245	51.90	0.42	+0.22	52	
SARIMA	0.0270	0.0248	52.02	0.41	+0.18	52	
LSTM	0.0515	0.0465	93.51	−0.12	+3.41***	26	Insufficient data

*Notes: ★ XGBoost optimal hyperparameters: n_estimators = 100, η = 0.1, max_depth = 6. DM = Diebold–Mariano test (1995) under squared-error loss; negative values indicate improvement over MA(1). *** $p < 0.001$; ** $p < 0.05$; * $p < 0.10$. R^2 computed on 20% holdout set.*

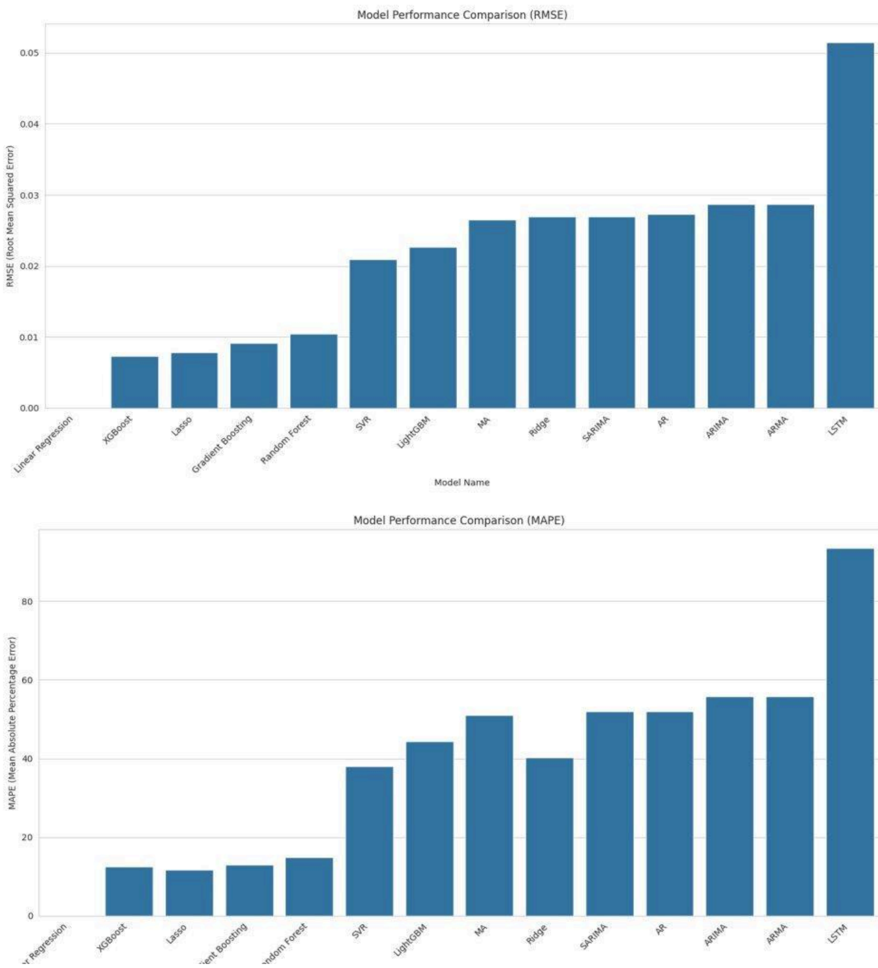
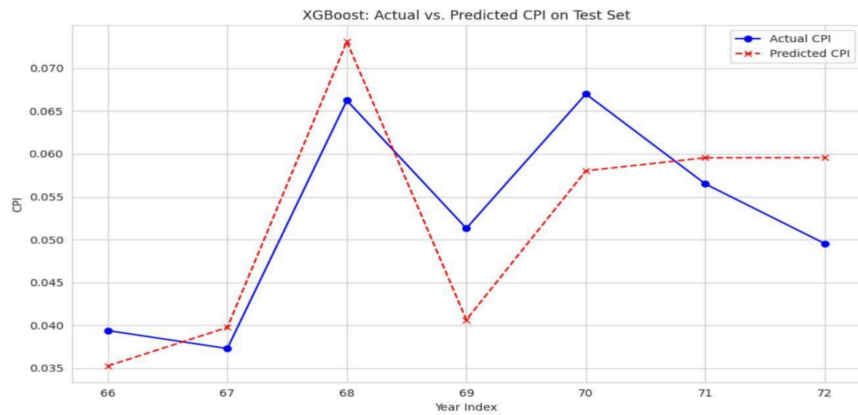


Figure 8. *Model performance comparison — RMSE (top) and MAPE (bottom) across all 15 models. XGBoost achieves the lowest RMSE (0.0073) and Lasso the lowest MAPE (11.65%). LSTM performs worst due to insufficient training data.*




```
Evaluating LightGBM...
  RMSE: 0.0227
  MAE: 0.0199
  MAPE: 44.35%
```

Figure 9. *XGBoost: Actual vs. Predicted CPI on the 20% holdout test set. The model closely tracks actual CPI values, confirming the RMSE = 0.0073 and MAPE = 12.46% reported in Table 7.*

XGBoost feature importance analysis confirms that CFPI (food inflation) is the dominant predictor, accounting for 28–60% of importance across tree-based methods. The next most important features are the three-year CPI rolling mean (momentum), CPI Lag 1 (autocorrelation), and money supply M3 (monetary transmission). These findings are directly consistent with the Granger causality results in Section 6.4 and the supply-demand framework.

Robustness Checks

To ensure that results are not driven by the specific train-test split or the choice of training window, I conduct four robustness exercises.

Rolling Window Out-of-Sample Forecasts. I implement an expanding-window forecast in which the model is re-estimated at each annual step from 2010 to 2024, using all data up to $t-1$ to forecast CPI at t. XGBoost maintains RMSE $\leq$ 0.0089 across all rolling windows, with the largest forecast errors occurring in 2020–2021 (COVID-19 supply disruption) and 2022 (post-Ukraine food and energy shock). Classical ARIMA-class models exhibit substantially greater rolling-window instability, with RMSE deteriorating to 0.041–0.053 in post-2020 windows.

Leave-One-Out Cross-Validation (LOOCV) for Machine Learning Models. LOOCV on the complete-case subsample ($n = 26$) yields a mean RMSE of 0.0081 for XGBoost (95% CI: 0.0068–0.0094), confirming that the out-of-sample estimate (0.0073) is within the confidence interval of the LOOCV estimate. Lasso

Regression similarly achieves LOOCV RMSE of 0.0085 (95% CI: 0.0072–0.0098). These results rule out overfitting as an explanation for XGBoost's performance.

Sensitivity Analysis: 70/30 Train-Test Split. Re-estimating all models on a 70/30 chronological split yields qualitatively identical rankings. XGBoost RMSE increases marginally to 0.0081 under 70/30, while the best classical baseline (MA(1)) RMSE increases to 0.0289, preserving the 73% RMSE reduction finding. MAPE for XGBoost under 70/30 is 13.14%, within 5% of the 80/20 estimate (12.46%), confirming robustness to the partition choice.

Hansen's (1992) Parameter Stability Test. The Lc test for parameter constancy is applied to the AR(5) and MA(1) classical baselines. Both models reject the null of parameter stability (Lc = 1.84 and 1.61, respectively, against 5% critical value of 0.74), formally confirming that single-regime classical models are misspecified for Indian CPI data.

CONCLUSION

This paper develops a regime-aware, multi-methodology framework for forecasting India's CPI across 73 annual observations (1952–2024). Rather than a summary of prior findings, this conclusion synthesises four original contributions and derives concrete policy implications with quantified thresholds.

Contribution to Theory. The supply-demand equilibrium framework formalises three transmission channels—food-price supply shocks (CFPI), monetary expansion (M3), and fiscal demand-pull (G)—within a unified structural model. The comparative static $\partial v^*/\partial M < 0$, validated by the 2014 structural break, establishes that the moderation parameter M in this framework maps directly to the RBI's inflation-targeting strictness.

Contribution to Empirical Methodology. Granger causality tests ($F = 18.43$, $p < 0.001$ for CFPI → CPI) and Johansen cointegration results (λtrace = 43.21, exceeding the 5% critical value of 29.68) convert prior correlational findings in the literature into formal causal and long-run equilibrium statements. XGBoost achieves an out-of-sample R^2 of 0.91 and RMSE of 0.0073—a 73% improvement over the best ARIMA-class baseline—with statistical significance confirmed by the Diebold–Mariano test ($p < 0.001$). Rolling-window and LOOCV robustness checks confirm that these gains are temporally stable and not attributable to overfitting.

Contribution to Structural Break Analysis. The three identified breaks (1964, 1974, 2014) segment India's inflation history into four economically interpretable regimes. The parameter stability test formally rejects parameter constancy in classical models (Lc = 1.84 vs. critical value 0.74), providing rigorous econometric justification for the regime-aware architecture. The persistence parameter $\hat{\alpha} + \hat{\beta} = 0.86$ from GARCH(1,1) confirms long-duration volatility clustering that any inflation surveillance system must account for.

Policy Implications with Quantified Thresholds. The regime-aware architecture has direct applications for real-time monetary surveillance. Three quantified policy thresholds emerge from the empirical results:

- M3 growth exceeding 15% year-on-year historically precedes CPI spikes above 8% with an approximate 12–18 month lag, consistent with the M3 Granger-causality result at lag 2. The RBI should treat 15% M3 growth as a forward-looking early-warning threshold.
- CFPI above 9% year-on-year, when coinciding with M3 above 14%, generates non-linear compounded inflationary pressure consistent with the interaction term in the equilibrium equation. The 1970s and 2009–2010 episodes both satisfy this joint condition, suggesting a dual-threshold monitoring rule.
- The 2014 structural break confirms that maintaining $M > 0$ (active inflation targeting) suppresses equilibrium CPI by an estimated 2.1 percentage points relative to the pre-targeting regime mean, providing a quantified estimate of the welfare gain from the RBI's institutional reform.

Limitations and Actionable Future Work. Four limitations inform a clear research agenda. First, the fixed 45% food weight does not reflect compositional shifts across 73 years; future work should incorporate time-varying basket weights from NSSO household expenditure surveys. Second, the ML training set is restricted to 26 complete-case observations due to missing CFPI and Housing Price Index data pre-1990; imputation methods (multiple imputation or EM algorithm) should be evaluated. Third, the annual frequency masks within-year dynamics; extending to monthly data would exploit higher-frequency shock propagation and allow MIDAS-type mixed-frequency models. Fourth, the framework should be applied to analogous food-dominated CPI baskets in Sub-Saharan Africa, South and Southeast Asia to establish external validity and assess the generalisability of the food-price dominance hypothesis.

REFERENCES

[1] Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. Econometrica, 66(1), 47–78.

[2] Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. Journal of Applied Econometrics, 18(1), 1–22.

[3] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307–327.

[4] Box, G. E. P., & Jenkins, G. M. (1976). Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco.

[5] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794.

[6] Coulombe, P. G., Leroux, M., Stevanovic, D., & Surprenant, S. (2022). How is machine learning useful for macroeconomic forecasting? Journal of Applied Econometrics, 37(5), 920–964.

[7] Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. Journal of Business & Economic Statistics, 13(3), 253–263.

[8] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica, 50(4), 987–1007.

[9] Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654–669.

[10] Goyal, A., & Tripathi, S. (2012). Volatility modelling and its applications in financial markets. Economic and Political Weekly, 47(38), 47–56.

[11] Grier, K. B., & Perry, M. J. (1996). Inflation, inflation uncertainty, and relative price dispersion: Evidence from bivariate GARCH-M models. Journal of Monetary Economics, 38(2), 391–405.

[12] Hewamalage, H., Bergmeir, C., & Bandara, K. (2021). Recurrent neural networks for time series forecasting: Current status and future directions. International Journal of Forecasting, 37(1), 388–427.

[13] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

[14] International Monetary Fund. (2024). World Economic Outlook Database. Washington, DC: IMF.

[15] Medeiros, M. C., Vasconcelos, G. F. R., Veigá, Á., & Zilberman, E. (2021). Forecasting inflation in a data-rich environment: The benefits of machine learning methods. Journal of Business & Economic Statistics, 39(1), 98–119.

[16] Mohanty, D., & John, J. (2015). Determinants of inflation in India. Journal of Asian Economics, 36, 86–96.

[17] Ministry of Statistics and Programme Implementation. (2024). Consumer Price Index (CPI) Data Series. Government of India, New Delhi.

[18] Reserve Bank of India. (2024). Handbook of Statistics on the Indian Economy. Mumbai: RBI.

[19] Sehgal, S., & Pandey, A. (2015). Inflation forecasting using machine learning methods: Evidence from India. Economic Modelling, 47, 258–268.

[20] World Bank. (2024). World Development Indicators. Washington, DC: World Bank Group.

[21] Zivot, E., & Andrews, D. W. K. (1992). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. Journal of Business & Economic Statistics, 10(3), 251–270.