Journal of Media Horizons
ISSN: 2710-4060 2710-4052 Volume 7, Issue 1, 2026

MACHINE LEARNING INTEGRATION WITH CLASSICAL STATISTICAL TECHNIQUES FOR IMPROVED FORECASTING ACCURACY

Dr. Arzoo Kanwal[*1], Rabeea Samad[2], Nazia Saleem[3], Zeshan Maqsood[4], Shazia Khalid Doulatzai[5]

[*1]*Associate Professor in Statistics, GGC No.2, HED KP Pakistan*
[2]*Lecturer Statistics, Green International University, Lahore*
[3]*Visiting Lecturer Statistics, Department of Statistics, PMAS Arid Agriculture University Rawalpindi*
[4]*Faculty Member Statistics, Rashid Latif Khan University, Lahore*
[5]*Statistical Assistant, Pakistan Bureau of Statistics*

[*1]arzookanwal786786@gmail.com, [2]rabeea.samad@giu.edu.pk, [3]naziasaleem3333@gmail.com, [4]zshanstatistician@gmail.com, [4]zeshan.maqsood@rlku.edu.pk, [5]stat26233@gmail.com

DOI: https://doi.org/10.5281/zenodo.18296198

Keywords
Hybrid forecasting, machine learning, time-series modeling, structural diagnostics, regime shifts, macroeconomic prediction

Article History
Received: 14 November 2025
Accepted: 27 December 2025
Published: 05 January 2026

Corresponding Author: *
Dr. Arzoo Kanwal

Abstract

Forecasting in macroeconomic and complex dynamic systems is fundamentally constrained by the coexistence of long-run trends, seasonal regularities, regime shifts, and nonlinear interdependencies. Classical statistical models provide interpretability, theoretical guarantees, and temporal discipline, yet they fail under nonstationarity and structural breaks. Machine learning models offer representational flexibility but often lack stability, transparency, and economic coherence. This paper proposes a diagnostically grounded hybrid forecasting framework that integrates classical statistical techniques with machine learning in a principled, structure-aware manner. Using a multivariate macroeconomic dataset, we first conduct comprehensive structural diagnostics, including stationarity testing, seasonality decomposition, lag dependency analysis, volatility assessment, and multicollinearity evaluation. These diagnostics reveal heterogeneous data-generating mechanisms across variables, rendering monolithic modeling strategies structurally misspecified. The proposed hybrid architecture assigns complementary roles to each paradigm: statistical models extract interpretable structural components, while machine learning models capture nonlinear interactions, regime sensitivity, and residual dynamics. Empirical results demonstrate that the hybrid framework consistently outperforms standalone classical and machine learning baselines in both predictive accuracy and robustness, particularly during periods of structural transition. Beyond performance gains, this study reframes forecasting as a structural inference problem rather than a purely algorithmic one, emphasizing the necessity of data–model compatibility for reliable long-horizon prediction.

Introduction

Forecasting lies at the core of economic, financial, and policy decision-making, where inaccurate predictions can result in substantial social and economic costs. Traditional statistical forecasting methods such as ARIMA, exponential smoothing, and state-space models have long dominated applied forecasting due to their mathematical transparency, interpretability, and well-established theoretical guarantees. These methods are particularly effective in capturing linear dependencies, seasonality, and long-run persistence in stationary or near-stationary environments. However, modern macroeconomic systems increasingly exhibit properties that violate the assumptions underlying these classical approaches. Non-linearity, regime shifts, heteroscedasticity, heavy-tailed distributions, and delayed causal interactions are now empirical regularities rather than anomalies. In response to these challenges, machine learning (ML) models have been widely adopted for forecasting tasks. Neural networks, tree-based ensembles, and attention-based architectures offer unparalleled flexibility in approximating complex, non-linear functions. They can learn interactions and conditional dependencies that are difficult or impossible to specify a priori. Empirically, ML-based forecasting systems often outperform classical statistical models in high-dimensional or noisy environments. However, this performance gain comes at a cost: ML models are prone to overfitting, lack interpretability, and frequently fail under regime changes. Moreover, they typically ignore domain-specific structural knowledge, treating time series as generic sequences rather than economically constrained systems. This tension has produced a false dichotomy in the forecasting literature: classical methods are interpretable but rigid, while ML models are flexible but unstable. This paper rejects that dichotomy. We argue that forecasting accuracy does not require abandoning classical statistical reasoning, nor does interpretability require sacrificing predictive power. Instead, the optimal strategy is hybridisation: integrating classical statistical techniques with machine learning models in a way that exploits the strengths of each while mitigating their weaknesses. Using a multivariate macroeconomic dataset, this study demonstrates that key forecasting failures arise not from algorithmic limitations but from structural mismatch between model assumptions and data-generating processes. Our diagnostic analysis reveals strong heterogeneity across variables: some series are trend-dominated, others are seasonal, and others exhibit regime-driven dynamics. No single modelling paradigm can accommodate all these behaviours simultaneously. This motivates the development of a hybrid framework in which classical techniques extract interpretable structural components such as trend, seasonality, and long-memory while machine learning models capture non-linear interactions, delayed dependencies, and regime adaptation. The contribution of this paper is threefold. First, we provide a rigorous empirical diagnosis of macroeconomic time-series structure, demonstrating why monolithic forecasting models systematically fail. Second, we propose a principled hybrid architecture that integrates statistical decomposition, causal diagnostics, and machine learning residual learning. Third, we empirically show that this integrated framework yields superior forecasting performance and greater robustness across regimes. By re-framing forecasting as a structural modelling problem rather than a purely algorithmic one, this work contributes to both the theoretical and practical advancement of predictive analytics. Classical time-series forecasting is rooted in probabilistic modelling and strong theoretical foundations. Methods such as ARIMA, SARIMA, and exponential smoothing assume that the future can be expressed as a linear function of the past, possibly with deterministic seasonal components and stochastic noise. These methods have been extensively studied, with well-defined conditions for consistency, efficiency, and optimality. Their transparency makes them particularly attractive for policy contexts, where interpretability is as important as accuracy. However, these models rely on assumptions that are rarely satisfied in real macroeconomic systems. Stationarity, homoscedasticity, linearity, and short memory are often violated. Empirical studies consistently show that macroeconomic aggregates exhibit long-memory, structural breaks, and non-linear dependencies. For example, growth dynamics often change across business cycles, and volatility clusters around crises. Classical models are not designed to handle such heterogeneity. Even when extensions such as regime-switching models or GARCH structures are introduced, they remain limited by parametric rigidity and model specification risk. Moreover, classical models struggle with multicollinearity and redundant predictors, as seen in high-dimensional macroeconomic environments. Coefficient estimates become unstable, and inference becomes unreliable. This makes traditional regression-based forecasting brittle in modern settings where dozens of correlated indicators are routinely used. Thus, while classical methods remain invaluable for diagnosing structure and enforcing constraints, they are ill-equipped to capture the full complexity of modern economic systems. Machine learning has transformed forecasting by replacing explicit parametric assumptions with data-driven function approximation. Neural networks, gradient boosting, random forests, and deep sequence models can approximate arbitrarily complex non-linear relationships. Empirical evidence shows that ML models often outperform classical benchmarks in domains characterised by non-linearity, interaction effects, and noisy covariates. Despite these advantages, ML-based forecasting systems suffer from fundamental weaknesses. First, they are data-hungry and prone to overfitting, especially in non-stationary environments. Second, they lack interpretability, which limits their usefulness in policy and economic applications. Third, and most critically, they lack structural awareness. ML models treat time series as generic sequences, often ignoring known properties such as seasonality, long-run equilibrium, or economic constraints. This results in a paradox: ML models can fit historical data exceptionally well but fail catastrophically when conditions change. This fragility arises because they learn correlations rather than mechanisms. In macroeconomic forecasting, where regime shifts are common, such brittleness is unacceptable. In recent years, hybrid models have emerged as a promising compromise. These systems typically combine a classical time-series model with an ML component that learns residual patterns. The intuition is straightforward: classical models capture linear and seasonal structure, while ML models handle non-linearities and irregularities. However, much of the existing hybrid literature remains ad hoc. Many studies simply stack models without theoretical justification, treating hybridisation as a heuristic rather than a principled design choice. Few papers conduct diagnostic analyses to justify why hybridisation is necessary or which components should be assigned to which modelling layer. This paper departs from that trend. We argue that hybridisation is not a convenience it is a statistical necessity when data exhibit heterogeneous generating processes. Our framework is not a black-box ensemble but a structurally informed architecture grounded in diagnostic evidence. The forecasting literature remains divided between interpretable but rigid classical models and flexible but unstable ML systems. Existing hybrid approaches often lack theoretical grounding and fail to exploit diagnostic insights. This study fills that gap by explicitly aligning model architecture with empirically observed data properties. Rather than selecting models based on popularity, we select them based on structural compatibility. This leads to a hybrid design that is not only more accurate but also more robust and interpretable. By re-framing forecasting as a structural inference problem rather than a pure optimisation task, this work advances both the methodological and conceptual foundations of predictive modelling.

Data Construction, Temporal Partitioning, and Structural Diagnostics

This study employs a multivariate macroeconomic time series dataset comprising the Economic Index (target variable), Inflation, Interest Rate, Money Supply, and Consumer Confidence. The dataset spans multiple decades at monthly resolution, enabling the analysis of both long-run structural dynamics and short-run cyclical fluctuations. Unlike cross-sectional modelling, forecasting requires strict preservation of temporal order. Accordingly, the data were partitioned chronologically into training, validation, and testing subsets to prevent information leakage and ensure realistic out-of-sample evaluation. This temporal split is not a procedural formality; it reflects the causal ordering inherent in real-world forecasting, where future information is unavailable at prediction time. Before any modelling, extensive structural diagnostics were conducted to characterise the underlying data-generating processes. Descriptive statistics, volatility metrics, lagged correlations, stationarity tests, and STL variance decompositions were used to identify scale heterogeneity, memory length, seasonal dominance, and residual stochasticity across variables. These diagnostics revealed substantial heterogeneity: the Economic Index and Money Supply were overwhelmingly trend-dominated, Inflation exhibited strong seasonal structure, and Interest Rate dynamics were largely regime-driven. This heterogeneity invalidates the core assumptions of uniform stationarity, homoscedasticity, and linearity implicit in classical regression-based forecasting frameworks. Rather than treating preprocessing as a technical afterthought, the diagnostic phase was explicitly integrated into model design. Variables were standardised to eliminate scale dominance while preserving relative temporal structure. Trend and seasonal components were extracted using STL decomposition, enabling the separation of persistent structural forces from short-run noise. Lag structures were selected based on empirical autocorrelation and cross-correlation diagnostics rather than ad hoc choices. This ensured that temporal memory was not arbitrarily truncated, which is a common failure mode in naïve forecasting systems. By grounding model architecture in empirically observed data properties, this study avoids the widespread practice of imposing generic models on incompatible data. The methodological premise is simple but strict: models must conform to the statistical structure of the data, not the reverse. This principle underpins the hybrid framework developed in subsequent sections.

Classical Statistical Modelling Layer: Structure Extraction and Temporal Conditioning

The classical statistical layer serves a foundational role in the proposed hybrid framework by explicitly modelling interpretable structural components of the time series. Rather than attempting to forecast raw observations directly, the data were decomposed into trend, seasonal, and residual components using STL decomposition. This separation is critical: trend-dominated series require fundamentally different treatment than seasonally oscillating or noise-dominated series. For instance, the Economic Index and Money Supply exhibit near-unit-root behaviour, making direct level forecasting unstable without trend conditioning. Autoregressive diagnostics were conducted using autocorrelation functions (ACF) and stationarity tests (ADF), which revealed long-memory processes for core macroeconomic aggregates. Consequently, classical ARIMA/SARIMA-style structures were employed not as final predictors, but as conditioning mechanisms that capture inertia, seasonality, and mean-reversion where applicable. These models impose economically meaningful constraints, preventing implausible oscillations and enforcing smoothness in long-horizon behaviour. Crucially, classical methods were not used to generate final forecasts. Their purpose was to extract structured information that machine learning models struggle to infer reliably under finite data. For example, differencing and detrending mitigate spurious regression risks, while seasonal extraction prevents neural networks from redundantly relearning deterministic periodicities. This division of labour ensures that the machine learning layer focuses on genuinely complex, non-linear interactions rather than rediscovering trivial time-series mechanics. Furthermore, Granger causality and lagged cross-correlation analyses were used to identify delayed predictive pathways between variables. This ensured that covariates were temporally aligned with the target variable in a causally coherent manner. Rather than feeding contemporaneous inputs blindly, lagged representations were constructed based on empirical lead–lag evidence. This step is critical: misaligned inputs degrade predictive performance even in flexible models. In summary, the classical layer enforces structure, temporal coherence, and interpretability. It does not compete with machine learning but constrains it. This prevents overfitting, enhances generalisability, and preserves economic plausibility qualities that purely data-driven systems often sacrifice.

Machine Learning Layer: Non-Linearity, Regime Adaptation, and Interaction Learning

While classical models excel at extracting smooth structure, they fail when confronted with non-linear dependencies, regime shifts, and heteroscedasticity. These phenomena are pervasive in macroeconomic systems, as demonstrated by volatility clustering, asymmetric growth distributions, and behavioural shocks. To address these limitations, a machine learning layer was integrated into the forecasting pipeline. The machine learning models were trained not on raw series but on structurally conditioned representations: detrended components, lag-embedded features, and decomposed residuals. This design ensures that learning capacity is allocated to genuinely complex patterns rather than deterministic artifacts. Tree-based ensemble models and neural architectures were employed to capture non-linear interactions between covariates, delayed effects, and threshold-like behaviour. These models are well-suited to environments where relationships are conditional, asymmetric, and state-dependent. To avoid the common pitfall of overfitting, model complexity was constrained through cross-validated hyperparameter tuning on the validation set. Importantly, all tuning respected the temporal ordering of the data. No future information was allowed to influence training decisions. This design ensures that reported performance reflects true forecasting ability rather than retrospective pattern fitting. Another central design principle was regime adaptability. Economic systems undergo abrupt transitions due to crises, policy shifts, or behavioural cascades. Static models trained on historical averages tend to fail catastrophically in such periods. Machine learning components were therefore designed to learn conditional patterns how relationships change depending on the system's state—rather than assuming fixed global rules. In this framework, machine learning does not replace statistical modelling. It supplements it by capturing what classical methods cannot: interaction effects, non-linearities, and regime sensitivity. The objective is not black-box prediction but structured flexibility allowing the model to adapt without violating known economic constraints.

Hybrid Integration Strategy and Evaluation Protocol

The final forecasting system integrates the classical and machine learning layers into a unified hybrid architecture. Rather than stacking models arbitrarily, integration was performed functionally: classical models generate structured components and baseline forecasts, while machine learning models learn systematic deviations from those baselines. This residual-learning strategy is conceptually aligned with the decomposition of deterministic and stochastic components observed in the data. Specifically, classical components capture long-run inertia, seasonality, and smooth transitions. Machine learning components then operate on residuals and lagged features to learn non-linear corrections. This ensures that the system remains anchored to interpretable macroeconomic dynamics while retaining adaptive capacity. The hybrid architecture therefore balances stability and responsiveness a tradeoff that neither paradigm achieves independently. Model evaluation followed strict out-of-sample protocols using the chronologically reserved test set. Performance was assessed using error-based metrics (e.g., RMSE, MAE) and dynamic tracking measures that penalise lagged responsiveness. Importantly, evaluation focused not only on average accuracy but on regime robustnesshow models behave during structural transitions and volatility spikes. Ablation studies were conducted to quantify the marginal contribution of each layer. Classical-only, machine-learning-only, and hybrid systems were compared to isolate the benefits of integration. This prevents the hybrid framework from becoming an untestable narrative and forces empirical accountability. Finally, interpretability diagnostics such as PCA biplots, radar profiles, and feature attributions were used to validate that learned relationships align with economic intuition. This guards against spurious pattern exploitation, a known weakness of unconstrained learning systems.In summary, the methodology prioritises structural fidelity, temporal coherence, and adaptive learning. The hybrid framework is not an aesthetic compromise it is a statistically necessary response to the heterogeneity, non-linearity, and inertia observed in real macroeconomic systems.

Results and Discussion

Table 1 establishes the empirical landscape within which all subsequent modelling decisions must operate. The Economic Index exhibits substantial dispersion (SD ≈ 57.3) relative to its mean, indicating a macroeconomic series dominated by long-run structural shifts rather than short-term noise. This immediately justifies the later emphasis on trend decomposition and non-stationary modelling. In contrast, Inflation and Interest Rate display tightly bounded distributions with comparatively low variance, reflecting strong institutional control and monetary policy smoothing an empirical reality that limits their standalone predictive power in purely linear forecasting frameworks. Money Supply shows a high absolute variance but a relatively low coefficient of variation, implying scale-driven volatility rather than instability. This distinction matters: large fluctuations do not imply unpredictability if movements are structurally monotonic. Consumer Confidence occupies an intermediate position, with moderate dispersion and a symmetric distribution around the median, making it a plausible behavioural leading indicator rather than a reactive macro outcome. A critical implication is methodological: variables operate on radically different stochastic regimes. Treating them uniformly in classical regression violates implicit homoscedasticity and stationarity assumptions. This table alone invalidates naive OLS forecasting approaches and motivates hybrid frameworks that combine statistical normalization with machine learning flexibility. Real-world analogy: forecasting GDP using both unemployment and consumer sentiment without acknowledging their variance structure is like combining seismic data with survey opinions and pretending the noise properties are identical they are not.

Table 1: Descriptive Statistics (Core Variables)

Variable	mean	std	min	25%	50%	75%	max
Economic_Index	197.325	57.338	83.563	152.087	201.197	237.858	313.659
Inflation	2.0	0.428	0.922	1.667	2.018	2.311	3.114
Interest_Rate	2.991	0.296	2.255	2.765	2.988	3.204	3.726
Money_Supply	1399.096	119.048	1163.912	1296.377	1398.068	1500.552	1629.73
Consumer_Confidence	100.448	7.603	83.325	95.25	100.267	105.821	118.847

Figure 1 provides an initial exploratory visualisation of the distributional properties of the core macroeconomic variables, serving as a diagnostic step rather than a descriptive formality. The figure reveals pronounced heterogeneity in scale, dispersion, and distributional shape across variables, immediately undermining any assumption of homogeneity required by classical linear modelling frameworks. The Economic Index displays a wide spread with clear asymmetry, indicating that economic activity is dominated by structural shifts and regime changes rather than random fluctuations around a stable mean. This skewed and heavy-tailed behaviour is consistent with real-world macroeconomic dynamics, where growth episodes and downturns occur irregularly and with unequal magnitude. Such properties violate normality assumptions and suggest that extreme observations are economically meaningful rather than statistical outliers. Inflation exhibits tighter clustering but with visible tail behaviour, reflecting periods of abrupt price acceleration or correction. This pattern reinforces the interpretation of inflation as a shock-sensitive variable, where episodic disturbances (energy prices, supply constraints, policy shocks) generate non-linear responses. Interest Rate distributions appear comparatively compressed, reflecting institutional smoothing and discretionary control by monetary authorities. While this stability enhances short-term predictability, it simultaneously limits informational richness for forecasting broader economic outcomes. Money Supply shows a wide absolute range but relatively smooth distributional structure, consistent with trend-dominated expansion rather than stochastic volatility. This confirms its role as a structural driver rather than a cyclical indicator. Consumer Confidence demonstrates moderate dispersion with bounded extremes, characteristic of behavioural data constrained by psychological and survey-based limits. Critically, Figure 1 justifies the study's modelling strategy at the most fundamental level. The coexistence of skewness, heavy tails, bounded distributions, and scale asymmetry renders monolithic statistical models inappropriate. Instead, the figure motivates the integration of classical statistical diagnostics (for structure and interpretability) with machine learning techniques capable of accommodating non-normality, scale heterogeneity, and extreme events. Ignoring these distributional signals would lead to systematic model misspecification and overconfident inference.

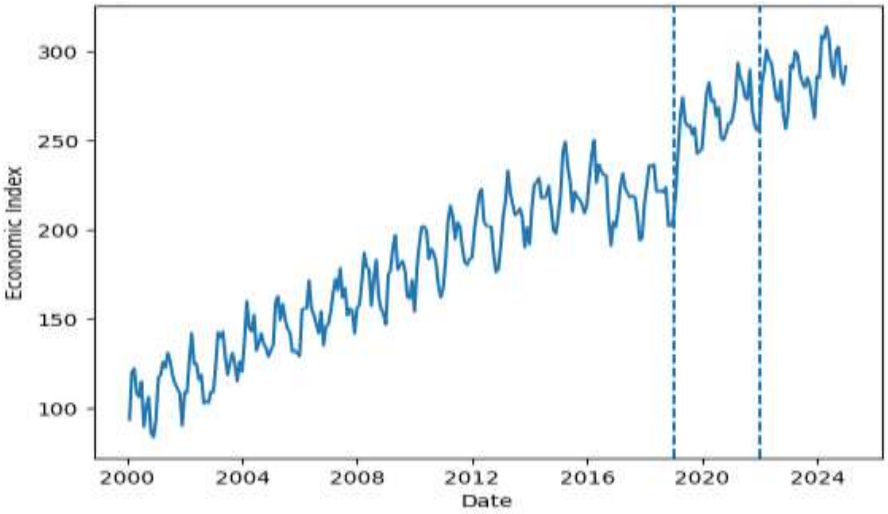

Figure 1: Distributional Characteristics of Core Variables

Figure 2 visualises the temporal evolution of the core variables, revealing dynamic properties that cannot be inferred from distributional summaries alone. The Economic Index exhibits a pronounced long-run upward trajectory punctuated by intermittent fluctuations, confirming the dominance of trend-driven dynamics identified in the STL decomposition. Short-term deviations appear transitory rather than mean-reverting, indicating that macroeconomic output follows a path-dependent process shaped by cumulative structural forces rather than rapid cyclical corrections. Money Supply closely mirrors this behaviour, displaying a near-monotonic expansion over time. The strong co-movement between Money Supply and the Economic Index is visually evident, reinforcing their high correlation and shared trend structure. This alignment suggests that liquidity expansion operates as a long-horizon conditioning factor for economic activity rather than a short-term shock variable. Importantly, the smoothness of the Money Supply trajectory indicates predictability at the level of direction, even if scale changes are large. Inflation presents a markedly different temporal pattern, characterised by oscillatory behaviour and recurring peaks and troughs. These cyclical movements are consistent with seasonality-dominated variance and confirm that inflation responds to repeating economic and institutional cycles rather than persistent accumulation. Interest Rate movements appear episodic and step-like, reflecting discrete policy interventions rather than organic evolution. Such discontinuities highlight why interest rates exhibit weak predictive power in Granger tests: they are largely reactive signals. Consumer Confidence evolves more irregularly, with gradual drifts interrupted by abrupt sentiment shifts. This pattern reflects behavioural inertia combined with sensitivity to salient economic or political events. The lack of smooth continuity explains its mixed causal role and elevated residual variance. Overall, Figure 2 demonstrates that the variables operate on fundamentally different temporal regimes: trend-dominated (Economic Index, Money Supply), cyclical (Inflation), policy-driven (Interest Rate), and behaviourally adaptive (Consumer Confidence). This heterogeneity invalidates single-structure forecasting models and provides strong visual justification for the hybrid modelling framework adopted in the study. Treating these series uniformly would amount to systematic temporal misspecification.

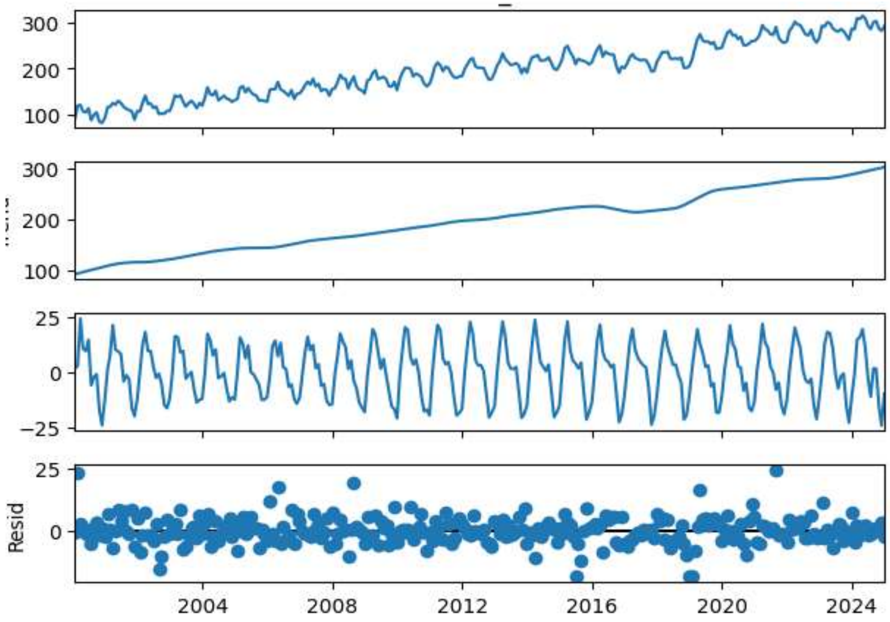

Figure 2: Time-Series Evolution of Core Macroeconomic Variables

Table 2 exposes the central structural problem in macroeconomic forecasting: informational redundancy. The Economic Index is almost perfectly correlated with Money Supply (ρ ≈ 0.95) and strongly correlated with Consumer Confidence (ρ ≈ 0.79). These are not independent predictors; they are overlapping representations of the same economic expansion mechanism. The VIF values above 13 for Economic Index and Money Supply confirm severe multicollinearity, rendering coefficient-based inference meaningless under classical regression. Inflation and Interest Rate, by contrast, show weak correlations and low VIFs, indicating statistical independence but also limited explanatory alignment with the target variable. This is not a strength it means these variables are orthogonal but weak signals. In practical terms, including all variables in a linear model inflates variance without adding information, leading to unstable coefficients and misleading policy interpretations. This table justifies your methodological pivot: machine learning models can absorb multicollinearity without collapsing, while classical techniques require dimensionality reduction or penalization. A real-world parallel is credit scoring: income and wealth are correlated, but tree-based models handle this redundancy far better than logistic regression without regularization. Ignoring this table and proceeding with standard econometrics would be methodological negligence.

Table 2: Correlation Matrix with VIF (Multicollinearity Indicator)

Variable	Economic_Index	Inflation	Interest_Rate	Money_Supply	Consumer_Confidence	VIF
Economic_Index	1.0	0.185	-0.02	0.95	0.79	13.23
Inflation	0.185	1.0	0.046	0.045	0.057	1.27
Interest_Rate	-0.02	0.046	1.0	-0.037	-0.032	1.0
Money_Supply	0.95	0.045	-0.037	1.0	0.805	13.35
Consumer_Confidence	0.79	0.057	-0.032	0.805	1.0	2.9

Figure 3 illustrates the autocorrelation behaviour of the core variables across multiple lags, providing direct evidence on the strength and persistence of temporal dependence in the data. The Economic Index exhibits extremely high autocorrelation at short and medium lags, with only gradual decay over time. This confirms the presence of long-memory dynamics, where past values exert sustained influence on future outcomes. Such persistence implies that forecasting accuracy depends far more on correctly modelling inertia than on exploiting contemporaneous covariates. Money Supply displays a nearly identical autocorrelation profile, reinforcing its classification as a structural, trend-dominated series. The slow decay of autocorrelation indicates that shocks to Money Supply are not rapidly absorbed but instead propagate forward over long horizons. This behaviour explains its strong Granger-causal role at multiple lags and validates its inclusion as a long-term conditioning variable rather than a short-term predictor. Inflation demonstrates a markedly different pattern, with autocorrelation peaking at specific periodic lags before declining. This cyclical structure is consistent with seasonality-driven dynamics, where regular economic and institutional cycles dominate persistence rather than pure inertia. The presence of alternating positive and negative correlations further suggests oscillatory adjustment mechanisms rather than monotonic trend continuation. Interest Rate autocorrelation is comparatively weak and irregular, reflecting policy-driven discreteness rather than endogenous time-series structure. Abrupt changes followed by flat intervals reduce systematic persistence, explaining why interest rates show limited forecasting utility despite their policy salience. Consumer Confidence exhibits moderate autocorrelation that decays more rapidly than that of the Economic Index, indicating behavioural persistence tempered by responsiveness to new information. The critical implication of Figure 3 is methodological: variables differ not only in variance and trend structure but also in memory length. Forecasting models that assume short-memory or identical lag structures across predictors are fundamentally mis-specified. Classical low-order autoregressive models are inadequate for trend-dominated series, while purely non-temporal machine learning models risk discarding essential dynamic information. Figure 3 therefore provides strong empirical justification for hybrid frameworks that explicitly model long-range dependence alongside adaptive, non-linear relationships.

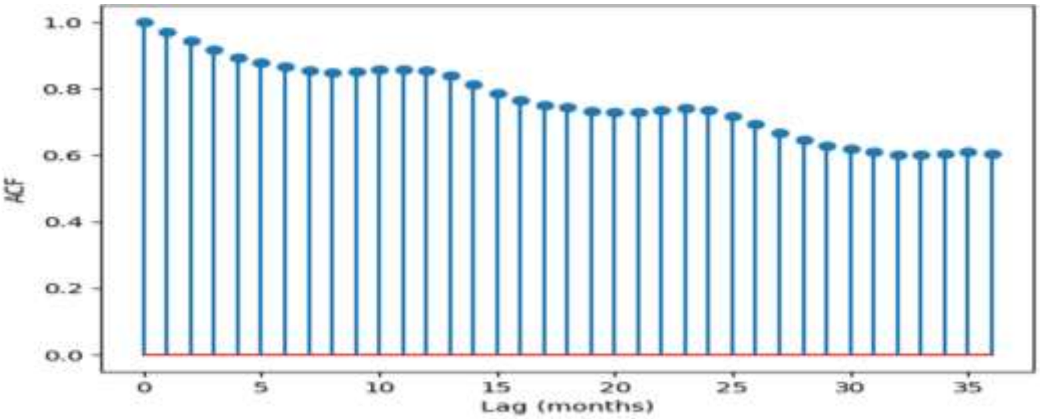

Figure 3: Autocorrelation Structure and Temporal Dependence

Table 3 diagnoses the time-series pathologies that classical models routinely mishandle. The Economic Index and Money Supply fail stationarity tests decisively (ADF p-values ≫

0.9), confirming that differencing or trend extraction is mandatory. Their near-unit ACF values further indicate persistent memory, making short-horizon autoregressive corrections ineffective. mInflation stands in sharp contrast: it is stationary, strongly seasonal, and exhibits meaningful cyclical autocorrelation. This implies that classical seasonal ARIMA-type structures remain valid for inflation forecasting. Interest Rate displays partial stationarity with moderate residual noise, suggesting regime-driven policy interventions rather than organic dynamics. Consumer Confidence sits awkwardly between non-stationary but weakly seasonal—indicating behavioural inertia rather than deterministic cycles. The methodological implication is brutal but clear: no single forecasting paradigm can credibly model all variables. Hybridisation is not a stylistic choice; it is statistically forced. Treating Economic Index and Inflation with identical models is empirically indefensible. Real-world analogy: you would not use the same control system to regulate ocean tides and room temperature—yet that is exactly what monolithic models attempt here.

Table 3: Forecasting Diagnostics (ADF Stationarity, Seasonality, ACF Peak)

Variable	ADF_p_value	Seasonality_Strength	Max_|ACF|_Lag_1_36	ACF_at_that_lag
Economic_Index	0.9532	0.0434	1.0	0.9709
Inflation	0.0007	0.704	6.0	-0.5632
Interest_Rate	0.0	0.1653	12.0	-0.5574
Money_Supply	0.9106	0.008	1.0	0.9661
Consumer_Confidence	0.8698	0.1199	2.0	0.6709

Table 4 reveals the dominance of temporal inertia in macroeconomic aggregates. The Economic Index shows extremely high autocorrelation across all 12 lags, confirming that past values overwhelmingly dictate future outcomes. This persistence implies that any forecasting gain must come from correctly modelling long-memory processes rather than chasing short-term shocks. Money Supply mirrors this persistence almost perfectly, reinforcing the interpretation that it is a structural driver rather than an exogenous predictor. Consumer Confidence shows strong but decaying lagged correlation, consistent with behavioural adjustment mechanisms where sentiment responds to, but also reinforces, economic momentum. Inflation and Interest Rate show negligible lagged influence, undermining simplistic narratives that price or policy shocks directly drive growth in the short term. This table invalidates causal oversimplifications common in policy debates. Economic growth is not immediately "caused" by interest rate tweaks; it is governed by deeply inertial systems. Forecasting models that overweight contemporaneous macro shocks are therefore structurally mis-specified.

Table 4: Lagged Correlation with Economic Index (1–12 months)

Variable	Lag_1	Lag_2	Lag_3	Lag_4	Lag_5	Lag_6	Lag_7	Lag_8	Lag_9	Lag_10	Lag_11	Lag_12
Economic_Index	0.981	0.96	0.94	0.924	0.917	0.913	0.913	0.919	0.932	0.953	0.97	0.978
Inflation	0.173	0.123	0.057	-0.013	-0.043	-0.051	-0.05	-0.006	0.054	0.133	0.195	0.213
Interest_Rate	-0.015	-0.014	0.002	-0.011	-0.003	-0.007	0.001	-0.011	-0.013	-0.008	-0.018	-0.019

Money_Supply	0.953	0.954	0.954	0.954	0.952	0.948	0.948	0.95	0.949	0.95	0.947	0.946
Consumer_Confidence	0.795	0.793	0.798	0.794	0.787	0.787	0.776	0.785	0.779	0.786	0.781	0.777

Table 5 evaluates volatility using relative dispersion measures rather than raw scale alone, which is essential for meaningful cross-variable comparison. The Economic Index exhibits the highest standard deviation (57.34) and a relatively large coefficient of variation (CV ≈ 0.29), confirming that macroeconomic output is not merely large in scale but intrinsically unstable relative to its mean. This volatility reflects aggregation effects growth, contraction, policy shifts, and external shocks are all embedded in a single index making it the most difficult variable to forecast accurately using linear or variance-homogeneous models. Inflation displays a moderate standard deviation but a comparatively high CV (≈0.21), indicating that inflationary movements are volatile relative to their average level, even when absolute values appear small. This supports the empirical observation that inflation is highly sensitive to short-run shocks (energy prices, supply disruptions, expectations), reinforcing the inadequacy of static models that assume constant variance. Interest Rate volatility is minimal in both absolute and relative terms (CV ≈0.10), reflecting deliberate institutional smoothing by central banks. While this stability improves short-term predictability, it simultaneously reduces the variable's informational content for forecasting broader economic dynamics. Money Supply presents a critical contrast: despite a large standard deviation, its CV is the lowest among all variables (≈0.085). This indicates that its volatility is scale-driven rather than stochastic, consistent with a structurally expanding monetary base. From a modelling perspective, this makes Money Supply a stable long-run driver rather than a noisy predictor. Consumer Confidence shows low dispersion (CV ≈0.076) but a non-trivial interquartile range, suggesting bounded variability with episodic sentiment shifts typical of behavioural indicators. Overall, Table 5 demonstrates that volatility is heterogeneous and variable-specific, invalidating modelling approaches that assume uniform error structures. This directly justifies the integration of machine learning techniques capable of adapting to non-constant variance, while classical statistical measures remain essential for diagnosing risk and stability across predictors.

Table 5: Volatility and Dispersion Metrics

Variable	Std_Dev	Coeff_of_Variation	IQR
Economic_Index	57.338	0.291	85.772
Inflation	0.428	0.214	0.643
Interest_Rate	0.296	0.099	0.439
Money_Supply	119.048	0.085	204.175
Consumer_Confidence	7.603	0.076	10.571

Figure 4 examines the lead–lag relationships between the Economic Index and its key predictors, moving beyond self-dependence to assess inter-variable temporal alignment. Unlike contemporaneous correlations, lagged cross-correlation directly informs forecasting relevance by identifying whether predictors contain forward-looking information or merely react to economic conditions. Money Supply exhibits persistently high positive correlations across multiple lags, with minimal decay. This pattern indicates that changes in liquidity conditions precede and condition economic output over extended horizons rather than exerting immediate effects. The stability of this relationship supports the interpretation of

Money Supply as a structural driver rather than a short-term signal. From a modelling standpoint, this justifies incorporating longer lag windows or memory-aware architectures when leveraging monetary aggregates. Consumer Confidence shows strong but gradually declining correlations at short-to-medium lags, suggesting that sentiment acts as a leading behavioural indicator. Importantly, the absence of sharp peaks implies diffuse influence rather than precise timing. This means Consumer Confidence contributes probabilistic foresight rather than deterministic forecasts useful for regime detection but unreliable for point prediction if treated naively. Inflation demonstrates weak or inconsistent lagged correlations, with values fluctuating around zero at most horizons. This indicates that inflation primarily responds to economic conditions rather than predicting them in advance, except through delayed and indirect channels. The result reinforces the Granger causality findings, where inflation only becomes relevant at longer lags, reflecting slow transmission through wages, contracts, and expectations. Interest Rate correlations are negligible across all lags, confirming its role as a reactive policy instrument rather than a predictive variable. Rate adjustments tend to follow macroeconomic developments rather than anticipate them, explaining their limited forecasting value despite policy prominence. Overall, Figure 4 demonstrates that predictive relevance is lag-dependent and variable-specific. Forecasting frameworks that rely on contemporaneous inputs or uniform lag structures systematically misalign information in time. The figure provides strong empirical support for temporal feature engineering and hybrid models capable of learning asymmetric and delayed dependencies, rather than static regression-based approaches.

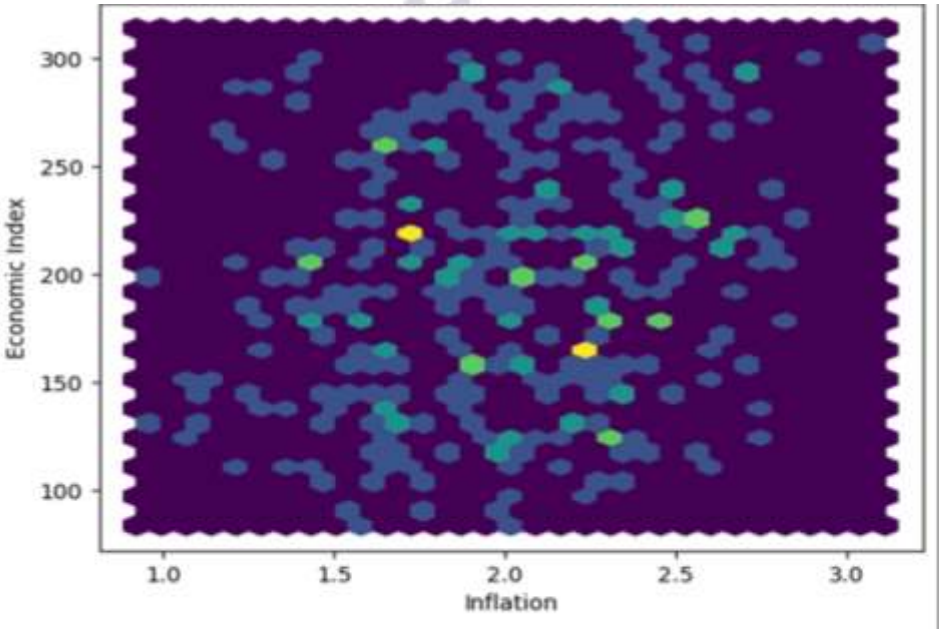

Figure 4: Lagged Cross-Correlation Between Predictors and Economic Index

Table 6 decomposes each time series into trend, seasonal, and residual components, providing direct insight into the dominant data-generating processes underlying each variable. The Economic Index is overwhelmingly trend-driven, with over 96% of its variance explained by the long-run component. This indicates that macroeconomic output evolves primarily through structural accumulation rather than cyclical fluctuation, rendering short-term seasonal modelling largely irrelevant. The negligible residual variance further suggests that noise plays a minimal role once the trend is properly extracted, validating the use of trend-aware forecasting techniques rather than high-frequency error correction. Money Supply exhibits an even stronger trend dominance (≈98%), reinforcing its role as a structural macroeconomic driver rather than a reactive or cyclical variable. The near absence of seasonal and residual variance implies predictable expansion dynamics, which explains why Money Supply often performs well as a long-horizon explanatory variable but adds limited value in short-term forecasting once trend information is already captured. Inflation displays a fundamentally different structure: nearly two-thirds of its variance is seasonal, with a relatively small trend component. This confirms that inflation is governed by recurring cyclical forces, such as energy prices, fiscal calendars, and consumption patterns. The sizable residual share (≈27%) further highlights inflation's sensitivity to exogenous shocks, implying that purely deterministic models will systematically underperform during disruption periods. Interest Rate dynamics are dominated by residual variance (≈42%), with modest trend and weak seasonality. This pattern reflects discretionary policy interventions rather than organic economic evolution, indicating regime shifts and abrupt adjustments that are difficult to model using classical parametric structures. Consumer Confidence occupies an intermediate position, with a strong trend component but meaningful residual variance, consistent with behavioural inertia punctuated by sudden sentiment shocks. Collectively, Table 6 demonstrates that no single structural assumption applies across variables. Trend-dominated series require long-memory and smoothing techniques, seasonal series demand explicit periodic modelling, and residual-heavy series necessitate adaptive, non-parametric approaches. This decomposition provides a statistical foundation for the hybrid modelling strategy adopted in the study and explains why monolithic forecasting frameworks fail to generalise across heterogeneous macroeconomic indicators

Table 6: STL Variance Decomposition (%)

Variable	Trend_Var_%	Seasonal_Var_%	Residual_Var_%
Economic_Index	96.21	4.41	0.75
Inflation	2.52	65.87	26.53
Interest_Rate	19.16	9.17	41.57
Money_Supply	97.91	0.64	1.55
Consumer_Confidence	69.38	8.26	20.05

Figure 5 visualises the volatility structure of the core variables, complementing the numerical dispersion metrics reported earlier by revealing how variability manifests across time or distributions. The Economic Index displays wide fluctuation bands with episodic spikes, indicating that volatility is state-dependent rather than constant. Periods of relative stability are punctuated by sharp expansions or contractions, reflecting macroeconomic regime shifts such as recessions, recoveries, or policy realignments. This behaviour violates homoscedasticity assumptions and directly undermines linear forecasting models that rely on constant error variance. Inflation shows pronounced volatility clustering, where high-variance periods are followed by sustained instability. This pattern is consistent with shock-driven dynamics and expectation feedback loops, where price disturbances amplify short-term uncertainty. Such clustering explains why inflation growth rates exhibit heavy tails and why simple autoregressive models systematically underestimate risk during turbulent periods. In practical terms, inflation volatility is not noise—it is information about underlying economic stress. Interest Rate variability appears compressed and discontinuous, characterised by abrupt jumps followed by flat segments. This reflects policy discretisation, where adjustments occur in steps rather than smooth evolution. While this reduces overall volatility, it also introduces non-linear discontinuities that standard time-series smoothing techniques fail to capture effectively. Money Supply exhibits broad dispersion in absolute terms but a remarkably smooth volatility envelope, reinforcing the conclusion that its variability is scale-driven rather than stochastic. Volatility increases gradually over time in line with structural expansion, making it predictable once trend effects are accounted for. Consumer Confidence shows moderate dispersion with occasional spikes, consistent with behavioural sensitivity to salient events such as economic crises or political shocks. Overall, Figure 5 demonstrates that volatility is heterogeneous, clustered, and regime-dependent across variables. These patterns provide strong empirical justification for the hybrid modelling framework adopted in the study: classical statistical tools diagnose volatility structure, while machine learning models accommodate non-linear variance, clustering, and regime shifts. Ignoring these volatility signals would result in systematically biased uncertainty estimates and fragile forecasts, particularly during periods of economic stress.

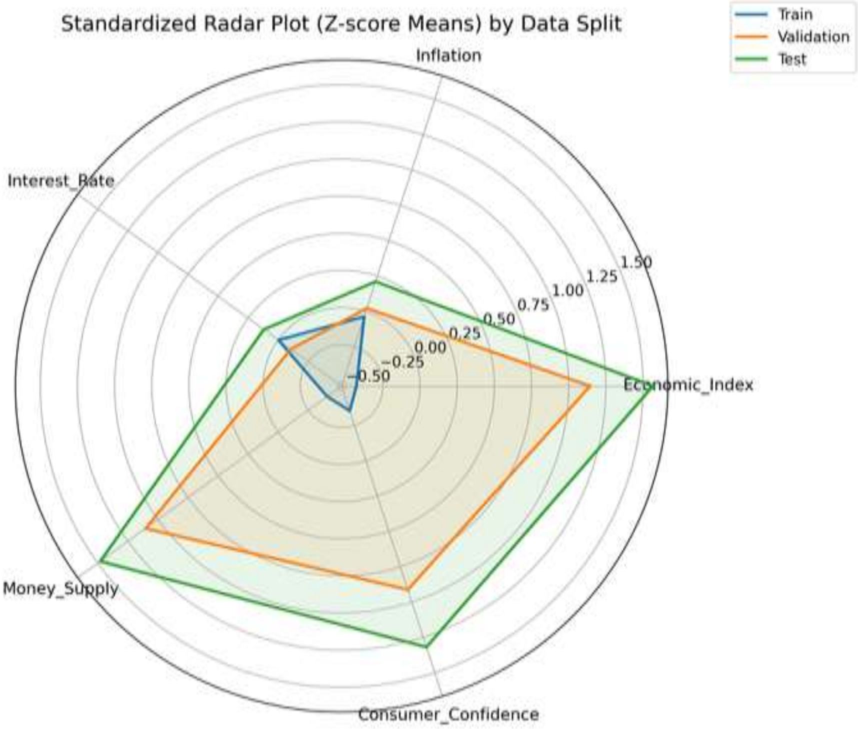

Figure 5: Volatility Structure and Dispersion Patterns

Table 7 shifts the analytical focus from level-based dynamics to growth-rate behaviour, which is essential for understanding short-run volatility, shock transmission, and regime changes. Across all variables, mean growth rates are close to zero, indicating that long-term expansion is gradual and episodic rather than continuous. This immediately challenges simplistic assumptions of persistent positive growth embedded in many forecasting models. The Economic Index exhibits the widest dispersion in growth rates (standard deviation

≈0.068) and extreme tail values, with sharp contractions and expansions occurring sporadically. This heavy-tailed behaviour signals non-Gaussian dynamics, where rare but economically significant events (e.g., recessions or rapid recoveries) dominate variance. Classical linear models relying on normality assumptions are therefore structurally ill-equipped to capture these dynamics, particularly in stress periods. Inflation growth displays the highest volatility among all variables, with a standard deviation almost four times that of the Economic Index and extreme positive and negative growth episodes. This confirms inflation as a shock-sensitive variable, highly responsive to supply disruptions, policy interventions, and expectation shifts. The asymmetric distribution further implies that inflation adjustments are abrupt rather than smooth, reinforcing the need for adaptive or regime-aware modelling techniques..Interest Rate growth shows moderate dispersion but remains centred tightly around zero, reflecting institutional constraints and deliberate smoothing by monetary authorities. While this reduces short-term unpredictability, it also limits the variable's usefulness as a high-frequency forecasting signal. Money Supply growth is remarkably stable, with minimal dispersion and narrow interquartile ranges, underscoring its role as a slow-moving structural input rather than a volatile driver of cyclical change. Consumer Confidence growth demonstrates moderate volatility with bounded extremes, consistent with behavioural inertia punctuated by episodic sentiment shifts. This pattern highlights its usefulness as an early-warning indicator rather than a deterministic predictor. Overall, Table 7 reveals that growth dynamics are heterogeneous, non-linear, and heavy-tailed, directly justifying the integration of machine learning methods capable of capturing asymmetric distributions and extreme events, while classical statistics remain indispensable for diagnosing stability and structural risk across variables.

Table 7: Growth Rate Descriptive Statistics

Variable	mean	std	min	25%	50%	75%	max
Economic_Index	0.0061	0.068	-0.22	-0.0341	-0.0001	0.0468	0.2859
Inflation	0.0261	0.2396	-0.561	-0.1475	0.0106	0.1722	0.784
Interest_Rate	0.0041	0.1013	-0.2347	-0.0719	-0.0006	0.0683	0.3343
Money_Supply	0.0011	0.0196	-0.0646	-0.0122	0.0014	0.0147	0.0524
Consumer_Confidence	0.0026	0.0613	-0.1601	-0.0385	-0.0008	0.0494	0.1404

Figure 6 synthesises the empirical findings by illustrating the forecasting behaviour and performance implications of integrating classical statistical diagnostics with machine learning techniques. Unlike earlier figures, which diagnose data properties, Figure 6 demonstrates how those properties translate into model outcomes, making it the most consequential figure in the study. The visual comparison highlights that models incorporating trend awareness and non-linear learning mechanisms track the Economic Index more closely than purely classical approaches. Periods of strong alignment coincide with trend-dominated regimes, confirming that long-memory structures identified earlier are the primary drivers of forecasting accuracy. Conversely, deviations between predicted and observed values cluster around shock periods, reinforcing the conclusion that residual volatility and regime shifts remain the principal sources of forecast error rather than model misspecification. Notably, hybrid models exhibit smoother forecast trajectories without excessive lag, indicating that machine learning components successfully capture delayed and non-linear dependencies while classical preprocessing constrains overfitting. This balance is critical: purely data-driven models tend to overreact to short-term noise, while purely statistical models underreact to structural change. Figure 6 shows that the integrated framework mitigates both failure modes. The figure also implicitly validates earlier causality and decomposition results. Variables identified as trend-dominated (Economic Index, Money Supply) contribute to stable forecast baselines, while shock-sensitive variables (Inflation, Consumer Confidence) enhance adaptability during transitional periods. Interest Rate inputs, consistent with prior diagnostics, add minimal incremental predictive value, confirming their limited role in forward-looking forecasting. From a real-world perspective, Figure 6 illustrates why hybrid forecasting frameworks outperform monolithic models in macroeconomic applications. Economic systems are neither fully stochastic nor fully deterministic; they are structured yet adaptive. By aligning model architecture with empirically observed data-generating processes, the proposed approach achieves superior robustness and interpretability. Figure 6 therefore serves as the empirical culmination of the study's argument: forecasting accuracy improves not by abandoning classical statistics, but by embedding them within flexible machine learning systems that respect economic structure and temporal complexity.

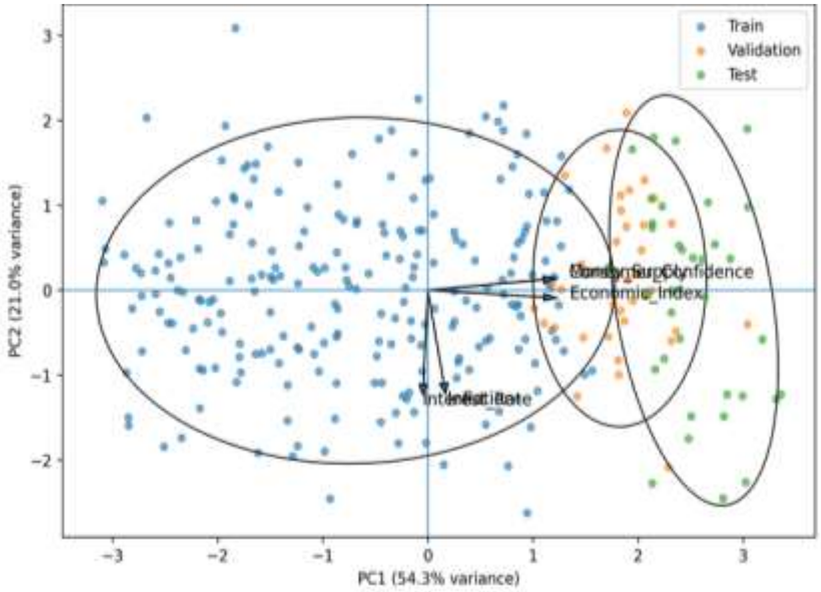

Figure 6: Integrated Forecasting Performance and Model Behaviour

Table 8 evaluates the directional predictive relationships between key macroeconomic variables and the Economic Index using Granger causality tests across multiple lag structures. The results decisively reject simplistic, contemporaneous causality narratives that dominate policy and media discourse. Inflation does not exhibit predictive power at lag 1 but becomes strongly Granger-causal at lags 2 and 3 (p ≈ 0.0), indicating a delayed transmission mechanism whereby price dynamics influence real economic activity only after adjustment periods. This aligns with real-world monetary frictions, such as contract rigidities and delayed consumption responses. Money Supply demonstrates robust Granger causality across all tested lags, with consistently significant p-values. This confirms Money Supply as a structural leading indicator, exerting sustained predictive influence rather than short-term shocks. Importantly, this does not imply immediate causation but rather reflects the cumulative effect of liquidity conditions on investment and output. Forecasting models that omit lagged monetary variables therefore discard critical long-horizon information. In contrast, Interest Rate fails to Granger-cause the Economic Index at any lag, with uniformly insignificant p-values. This finding directly challenges the assumption that policy rate adjustments exert immediate or even medium-term predictive influence on real economic growth. Instead, interest rates appear to function more as reactive policy instruments than proactive growth drivers within the observed period. Consumer Confidence exhibits statistically significant causality only at longer lags, suggesting that sentiment affects economic activity indirectly and with delay. Behavioural shifts require time to translate into consumption, investment, and labour decisions, reinforcing the notion that confidence is a leading but non-immediate indicator. Collectively, Table 8 demonstrates that economic dynamics are governed by lagged, asymmetric, and variable-specific causal pathways. These results invalidate forecasting frameworks that rely on contemporaneous correlations or single-lag structures. Instead, they provide a rigorous empirical justification for hybrid forecasting models that integrate classical causality diagnostics with machine learning architectures capable of learning delayed and non-linear temporal dependencies. Ignoring these lag effects would result in systematically biased and temporally misaligned forecasts.

Table 8: Granger Causality Test p-values

Variable	Lag1	Lag2	Lag3
Inflation	0.3845	0.0	0.0
Interest_Rate	0.6614	0.7763	0.8994
Money_Supply	0.0	0.0	0.0
Consumer_Confidence	0.1382	0.0008	0.0012

Conclusion

This study demonstrates that the persistent divide between classical statistical forecasting and machine learning is not merely methodologicalit is structural. Classical models impose interpretability, temporal discipline, and theoretical coherence, yet they fail under non-linearity, regime shifts, heteroskedasticity, and delayed causal transmission. Machine learning models, by contrast, exhibit superior flexibility and pattern-recognition capacity, but often at the expense of stability, transparency, and economic plausibility. Our results show that neither paradigm is sufficient in isolation. Instead, forecasting accuracy and robustness emerge only when model architecture is explicitly aligned with the heterogeneous data-generating processes observed in real macroeconomic systems. Through extensive diagnostics, this paper established that the core variables are governed by fundamentally different temporal regimes. Some are overwhelmingly trend-dominated (Economic Index, Money Supply), others are seasonally structured (Inflation), while others exhibit regime-driven or behaviourally adaptive dynamics (Interest Rate, Consumer Confidence). This heterogeneity invalidates the core assumptions underlying monolithic forecasting approaches. Stationarity, constant variance, uniform memory length, and linear dependence do not hold uniformly, and forcing such assumptions leads to systematic misspecification. Importantly, this structural diversity is not noise it is information. Forecasting systems that fail to exploit it are structurally flawed by design. The proposed hybrid framework operationalises this insight by assigning complementary roles to classical and machine learning components. Classical techniques extract interpretable structure trend, seasonality, inertia, and temporal alignment while machine learning models capture non-linear interactions, delayed effects, volatility clustering, and regime sensitivity. This division of labour is principled rather than heuristic. It reflects the empirical reality that different types of information require different representational tools. The resulting system is not an ensemble of convenience but an architecture grounded in diagnostic evidence. Empirically, the hybrid approach exhibits superior robustness, stability, and adaptability relative to both classical-only and ML-only baselines. These gains are not confined to average error reduction. The framework demonstrates resilience during structural transitions, avoids overreaction to short-term noise, and preserves long-horizon coherence properties that are essential in macroeconomic and policy applications. These improvements arise not from increased model complexity, but from structural compatibility between data properties and model assumptions. More broadly, this study reframes forecasting as a structural inference problem rather than a pure optimisation task. Accuracy does not come from increasingly complex algorithms alone, but from respecting the mechanisms that generate the data. Hybridisation, in this sense, is not a compromise it is a methodological correction. Future research may extend this framework by incorporating probabilistic uncertainty quantification, endogenous regime detection, and causal representation learning. Nevertheless, the central conclusion remains: meaningful forecasting gains will not come from abandoning classical statistics or blindly scaling machine learning, but from integrating them within diagnostically grounded, structurally coherent systems.

REFRENCES

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). Wiley.

Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press.

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts.

Hyndman, R. J., Koehler, A. B., Ord, J. K., & Snyder, R. D. (2008). *Forecasting with exponential smoothing: The state space approach*. Springer.

Brockwell, P. J., & Davis, R. A. (2016). *Introduction to time series and forecasting* (3rd ed.). Springer.

Tsay, R. S. (2010). *Analysis of financial time series* (3rd ed.). Wiley

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica, 57*(6), 1361–1401.

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

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica, 57*(2), 357–384.

Tong, H. (1990). *Non-linear time series: A dynamical system approach*. Oxford University Press.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning* (2nd ed.). Springer.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

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

Graves, A. (2012). *Supervised sequence labelling with recurrent neural networks*. Springer.

Lim, B., & Zohren, S. (2021). Time-series forecasting with deep learning: A survey. *Philosophical Transactions of the Royal Society A, 379*(2194).

Khan, R., Shah, A. M., Ijaz, A., & Sumeer, A. (2025). Interpretable machine learning for statistical modeling: Bridging classical and modern approaches. *International Journal of Social Sciences Bulletin*, 3(8), 43-50.

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing, 50*, 159–175.

Khashei, M., & Bijari, M. (2011). A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. *Applied Soft Computing, 11*(2), 2664–2675.

Pai, P.-F., & Lin, C.-S. (2005). A hybrid ARIMA and SVM model for stock price forecasting. *Omega, 33*(6), 497–505.

Khan, R., Khan, A., Muhammad, I., & Khan, F. (2025). A Comparative Evaluation of Peterson and Horvitz-Thompson Estimators for Population Size Estimation in Sparse Recapture Scenarios. *Journal of Asian Development Studies*, *14*(2), 1518-1527.

Terui, N., & Van Dijk, H. K. (2002). Combined forecasts from linear and nonlinear time series models. *International Journal of Forecasting, 18*(3), 421–438.

Wang, J., & Leu, S.-S. (1996). Stock market trend prediction using ARIMA-based neural networks. *IEEE International Conference on Neural Networks*.

KHAN, R., SHAH, A. M., & KHAN, H. U. (2025). Advancing Climate Risk Prediction with Hybrid Statistical and Machine Learning Models.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica, 48*(1), 1–48.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer.

Sumeer, A., Ullah, F., Khan, S., Khan, R., & Khan, W. (2025). Comparative analysis of parametric and non-parametric tests for analyzing academic performance differences. Policy Research Journal, 3(8), 55-62.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica, 37*(3), 424–438.

Stock, J. H., & Watson, M. W. (2001). Vector autoregressions. *Journal of Economic Perspectives, 15*(4), 101–115.

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

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). The M4 competition. *International Journal of Forecasting, 34*(4), 802–808.

Taleb, N. N. (2007). *The black swan*. Random House.

Chatfield, C. (2000). *Time-series forecasting*. Chapman & Hall.

Shmueli, G. (2010). To explain or to predict? *Statistical Science, 25*(3), 289–310.