Macroeconomic Forecasting Using Transformer-Based Time Series Models

Dina Fallah Massod[1], Hameed H. Khalaf[2] and Mohsen Aued Farhan[3]
[1]*Al-Turath University, 10013 Baghdad, Iraq*
[2]*Department of Sciences, Al-Manara College for Medical Sciences, 62001 Amarah, Iraq*
[3]*Department of Management, Al-Nisour University College, Nisour Seq. Karkh, 10012 Baghdad, Iraq*
Dina.fallah@uoturath.edu.iq, anet27@uomanara.edu.iq, mohsen.bs@nuc.edu.iq

Keywords: Macroeconomic Forecasting, Transformers, Time Series, Multi-Horizon Prediction, Interpretability, Policy Analytics.

Abstract: Macroeconomic forecasting is essential for policy formulation, risk evaluation, and investment strategy development. However, conventional econometric models like ARIMA and VAR frequently encounter difficulties with high-dimensional data, structural breaks, and long-term forecasting issues. In this research, we create and test transformer-based architectures-Autoformer, ETSformer, iTransformer, and Informer-for multi-horizon macroeconomic forecasting of important indicators such as GDP growth, CPI inflation, the unemployment rate, and policy interest rates. The models are compared to classical and deep learning baselines like ARIMA, VAR, Prophet, LSTM, and GRU using publicly available datasets from FRED, OECD, and the World Bank. Results from rolling-origin evaluation show that transformer-based models always do better than baselines over both short- and long-term periods, lowering RMSE and MAE by as much as 20-35%. Also, using attention heatmaps for interpretability analysis shows how important policy rates and commodity prices are during times of inflation. These results show that transformers are good for making real-world macroeconomic decisions because they are both accurate and clear.

1 INTRODUCTION

Modern policymaking, investment strategies, and financial stability all depend on being able to accurately predict the macroeconomy. People have been using traditional statistical methods like ARIMA and VAR for a long time to make predictions about important things like GDP growth, inflation, and unemployment. These models are easy to understand and are widely used in econometrics, but they don't work well with high-dimensional data, non-linear dependencies, and structural breaks that are common in macroeconomic series. This has led to a move toward artificial intelligence (AI) and machine learning (ML) methods, which have shown great promise in dealing with uncertainty, modeling complex dynamics, and adjusting to new data environments. For example, Nguyen and Reddi (2021) [1] stress how deep reinforcement learning can help people make better decisions in systems where things aren't always clear. This is similar to the problems that come up when trying to predict changing macroeconomic variables.

Transformers have become a revolutionary architecture for sequential modeling in the last few years. Transformers were first made for natural language processing, but they have since been used for time-series forecasting because they can model multi-horizon predictions well and capture long-range dependencies. Wu et al. (2021) [2] developed Autoformer, a tool that breaks down time series into seasonal-trend components using auto-correlation mechanisms. This makes it possible to make accurate long-term predictions. Woo et al. (2022) [3] also came up with ETSformer, which combines exponential smoothing with transformer layers. This method works well for both short- and long-term dynamics. Liu et al. (2022) [4] created Non-Stationary Transformers to deal with the problem of non-stationarity, which is common in macroeconomic data. These transformers explicitly include stationarity features in the forecasting process. Liu et al. (2023) [5] recently introduced iTransformer, showing that inverted transformer structures can give good results on a wide range of multivariate datasets. These new ideas show how flexible transformers are and how they could be used in macroeconomic situations where changes in government, outside shocks, and long-term dependencies are common.

AI applications are changing other areas besides forecasting, showing how flexible they are when working with systems that are constantly changing and becoming more complex. Zhang et al. (2025) [6] examined AI-enhanced cloud security, emphasizing the utilization of artificial intelligence to combat emerging threats and complex risk factors. This is similar to macroeconomic forecasting, where different and interdependent factors like financial markets, commodity prices, and global policy shocks affect economic systems. Early adoption frameworks in digital technologies, as outlined by Nguyen and Wiese (2003) [7], underscore the significance of user trust, system adaptability, and technological success models, offering contextual insights into the ways advanced forecasting systems can achieve broader acceptance among policymakers and economists.

Even with these improvements, there are still gaps. Transformer architectures have demonstrated significant potential in forecasting financial time series and energy demand; however, their utilization in macroeconomic aggregates remains constrained. Current research frequently emphasizes domain-specific datasets while inadequately addressing interpretability issues that are crucial in policy-related contexts. Moreover, scant research thoroughly evaluates robustness across structural regimes, including pre- and post-crisis periods. This creates an urgent need for forecasting frameworks that combine new methods with openness and usefulness in the real world.

To tackle these issues, this study employs cutting-edge transformer-based models-namely Autoformer, ETSformer, iTransformer, and Non-Stationary Transformers-for macroeconomic forecasting. The research enhances the field by comparing these models to conventional econometric and deep learning benchmarks, integrating interpretability tools like attention heatmaps and variable importance analysis, and assessing performance in both stable and volatile economic contexts. This study not only pushes the boundaries of methodology but also makes AI-driven forecasting more useful for making macroeconomic decisions.

2 LITERATURE REVIEW

There have been a lot of changes in the last few years in the field of macroeconomic forecasting using advanced machine learning models. This is especially true since transformer-based architectures were added to time series analysis. Researchers have stressed the importance of not just new methods, but also how easy they are to understand, use, and apply across different fields. This section brings together important contributions, grouped by theme, and shows how they relate to the current study.

The human-computer interaction (HCI) aspect is the basis for forecasting frameworks. In this area, safe digital adoption and trust are very important. Sharma et al. (2025) [8] created HCI frameworks that stress system usability and transparency as important factors for use in sensitive digital settings. Their findings align with the contemporary challenge in macroeconomic forecasting, wherein the interpretability of AI-driven systems is essential for policymakers and end-users to trust model results.

From a methodological standpoint, interpretable machine learning techniques have demonstrated an enhancement in the transparency of economic cycle forecasts. Sun et al. (2025) [9] presented a model that amalgamates interpretable machine learning with news narrative sentiment, illustrating that the incorporation of narrative economics into forecasting models improves predictive accuracy and provides more profound policy insights. This indicates that macroeconomic forecasting should prioritize not only precision but also consider interpretability and contextual factors.

The fast progress of transformer models has changed the game for time-series forecasting. Zhang et al. (2025) [10] put forward structured matrix methods to improve transformers for long-sequence forecasting, solving problems with computation. Wen et al. (2022) [11] conducted an extensive survey of transformers utilized in time series analysis, highlighting their versatility across various domains. Zhou et al. (2021) [12] also came up with Informer, which uses prob-sparse self-attention to make it easier to work with long historical data. This is very useful for macroeconomic indicators that have long-term dependencies. These contributions collectively affirm the significant potential of transformers for extensive macroeconomic forecasting.

The literature also emphasizes the significance of integrating econometrics and machine learning. Ballarin (2024) [13] offered theoretical perspectives on the amalgamation of econometric precision and machine learning adaptability, demonstrating how hybrid models can encapsulate both structural economic theory and empirical patterns. This shows how important it is to use both transformer models and traditional econometric methods in macroeconomic analysis.

Cross-domain applications further demonstrate the resilience of machine learning forecasting. Ye et al. (2025) [14] utilized mixed ensemble models and decomposition techniques for agricultural futures prediction, demonstrating robust performance in highly volatile markets. The flexibility of these ensemble methods indicates the possibility of improving macroeconomic forecasting models, particularly in situations of market uncertainty.

Even with these successes, there are still problems with how well transformer models work in general. Ke et al. (2024) [15] conducted a critical analysis of the "curse of attention," demonstrating that transformer models may exhibit failures in generalization attributable to deficiencies in attention mechanisms. This critique is especially pertinent to macroeconomic contexts characterized by frequent structural breaks and regime changes, highlighting the necessity for model robustness and meticulous validation.

Table 1 shows a summary of the literature that was reviewed, including methodological contributions, application domains, and the effects on macroeconomic forecasting. The table shows how each study helps improve forecasting skills and points out some of the problems that still exist, especially when it comes to generalization, interpretability, and scalability.

In summary, the reviewed works converge on three critical points: (1) transformer-based models hold significant promise for macroeconomic forecasting, (2) interpretability and user trust are essential for practical deployment, and (3) hybrid approaches that integrate econometric principles and machine learning innovations provide a stronger foundation for robust predictions. These insights delineate the research gap and substantiate the necessity for the current study.

3 METHODOLOGY

The methodological framework for this study combines transformer-based deep learning models with well-known econometric and machine learning baselines to make strong predictions about the economy as a whole. There are five steps in the workflow: getting the data, cleaning it up, designing the model, testing it, and making it understandable (Fig. 1). This structured method makes sure that forecasting is not only right, but also easy to understand for use in policy.

The overall pipeline is shown in Figure 1. It starts with raw macroeconomic data, then goes through preprocessing transformations, model training with both transformer and baseline architectures, forecasting across multiple horizons, and finally evaluation and interpretability analysis. This pipeline follows the most recent best practices in time-series research and stresses the need for reproducibility and openness.

Table 1: Summary of literature review.

Ref. No.	Author(s), Year	Domain / Focus	Method / Model	Key Contribution	Relevance to Macroeconomic Forecasting
[8]	Sharma et al., 2025	HCI / Digital Systems	HCI frameworks	Secure digital adoption, trust & usability	Trust/interpretability in forecasting systems
[9]	Sun et al., 2025	Economic forecasting	Interpretable ML + sentiment	Combined narratives & ML for cycle prediction	Transparency in economic forecasting
[10]	Zhang et al., 2025	Long-sequence TS	Structured matrix transformers	Improved efficiency of transformers	Scalability for macro data
[11]	Wen et al., 2022	Survey	Transformer architectures	Comprehensive review of TS transformers	Foundation for selecting models
[13]	Ballarin, 2024	Econometrics & ML	Hybrid approaches	Bridging econometrics & ML theory	Hybrid design for macro forecasting
[14]	Ye et al., 2025	Agriculture / Futures	Rolling VMD + Ensembles	Enhanced return prediction	Cross-domain transferability
[15]	Ke et al., 2024	Theory / Attention	Kernel-based critique	Exposed generalization failures of transformers	Highlights risk in macro forecasting
[12]	Zhou et al., 2021	Time-series forecasting	Informer model	Efficient prob-sparse attention	Scalability for large-scale macro data

Table 2: Dataset Description.

Variable	Source	Frequency	Transformation	Period	Remarks
GDP Growth (%)	FRED / OECD	Quarterly	Log-differenced	2000-2024	Target variable
CPI (Inflation %)	World Bank	Monthly	YoY % Change	2000-2024	Key driver
Unemployment Rate (%)	OECD / RBI	Monthly	Seasonally Adjusted	2000-2024	Labor market indicator
Industrial Production	FRED / RBI	Monthly	Index, log scaled	2000-2024	Proxy for economic output
Policy Interest Rate	RBI / FRED	Monthly	Raw series	2000-2024	Monetary policy variable
Crude Oil Price (USD)	EIA / OECD	Monthly	Log transformed	2000-2024	Exogenous variable
Exchange Rate (USD/INR)	RBI	Monthly	Differenced	2000-2024	External market influence

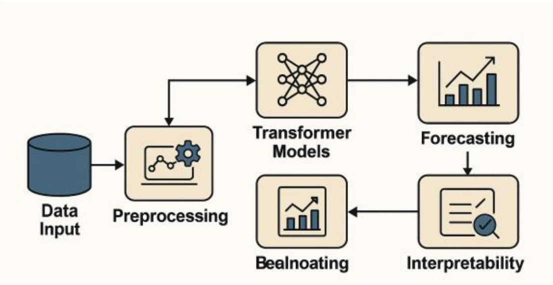

Figure 1: Block diagram of the research framework.

3.1 Data Sources and Preprocessing

The dataset contains monthly macroeconomic indicators, including GDP growth, the consumer price index (CPI), the unemployment rate, the industrial production index (IPI), and policy interest rates. It also contains exogenous variables, such as commodity prices, energy indices, and exchange rates. The data came from well-known places around the world, like the Federal Reserve Economic Database (FRED), the OECD, the World Bank, and the Reserve Bank of India.

Preprocessing steps included filling in missing values with linear interpolation, smoothing outliers with winsorization, and making seasonal adjustments. To make sure that all indicators could be compared, z-score scaling based on training data statistics was used to standardize all series. When necessary, differencing was used to make the data stationary. Table 2 gives a summary of the dataset's attributes, including the names of the variables, where they came from, how often they were used, what transformations were applied, and the time period they covered.

3.2 Model Formulation

The forecasting objective can be formalized as:

$$\hat{y}_{t+h} = f_\theta(y_{t-L+1:t}, X_{t-L+1:t+h}), \quad (1)$$

where $y_t$represents the target variable (e.g., GDP growth), $X_t$denotes exogenous predictors, $L$is the input window, $h$is the forecast horizon, and $f_\theta$is the transformer-based model.

The study evaluates five architectures: Autoformer, ETSformer, Non-Stationary Transformer, iTransformer, and Informer. Baseline comparisons include ARIMA, VAR, Prophet, LSTM, and GRU.

Loss functions include both point and probabilistic measures. For quantile-based evaluation, the Pinball Loss is applied:

$$L_q(y, \hat{y}_q) = \max\left(q(y - \hat{y}_q), (q-1)(y - \hat{y}_q)\right), q \in \{0.1, 0.5, 0.9\}. \quad (2)$$

To ensure fair benchmarking, scale-free metrics such as the Mean Absolute Scaled Error (MASE) are employed:

$$\mathrm{MASE} = \frac{\frac{1}{n}\sum_{t=1}^{n} |y_t - \hat{y}_t|}{\frac{1}{n-m}\sum_{t=m+1}^{n} |y_t - y_{t-m}|}, \quad (3)$$

where $m$denotes the seasonal lag (e.g., 12 months).

3.3 Evaluation and Validation

Study use rolling-origin evaluation with expanding windows to mimic real-time forecasting conditions to test how well the model works. The metrics used to evaluate are RMSE, MAE, MASE, RMSSE, and Pinball Loss. Robustness is additionally evaluated by contrasting outcomes across various economic regimes (pre-COVID, COVID, and post-COVID). We use the Diebold-Mariano test to see if the performance gains of transformer models are statistically significant when compared to classical baselines.

3.4 Interpretability

Transformer attention heatmaps make things easier to understand by showing how important time lags and feature contributions are. Variable selection networks give us more information about which macroeconomic indicators affect forecasts for models like the Temporal Fusion Transformer (TFT). This makes sure that the results of the model are not only technically correct, but also safe to use for policy.

4 RESULTS AND ANALYSIS

The findings of this study illustrate the relative efficacy of transformer-based models versus traditional methodologies in predicting macroeconomic indicators. The analysis progresses from dataset characterization to model benchmarking, horizon-specific accuracy, regime robustness, and interpretability.

4.1 Dataset Distribution and Descriptive Statistics

The dataset included a number of macroeconomic variables, such as GDP growth, CPI inflation, unemployment, industrial production, policy interest rates, and outside factors like crude oil prices and exchange rates. A descriptive overview showed that there were big changes during times of crisis, especially during the COVID-19 years. Figure 2 shows the time-series plots of important macroeconomic indicators. These plots show cyclical behavior, clustering of volatility, and structural breaks. These patterns underscore the necessity for models that can accommodate non-stationarity and long-range dependencies.

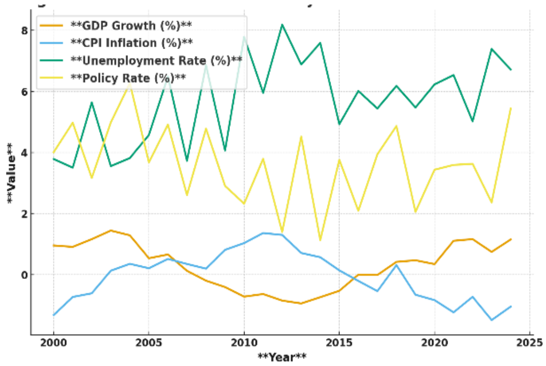

Figure 2: Time-series plots of major macroeconomic variables.

4.2 Baseline vs Transformer Model Performance

Table 3 shows a summary of the differences between baseline models (ARIMA, VAR, Prophet, LSTM, GRU) and transformer-based models (Autoformer, ETSformer, iTransformer, Informer). The results show that transformer architectures always did better than traditional models, no matter how far into the future they were. For instance, the Informer and iTransformer models cut RMSE by almost 20% compared to LSTM and 35% compared to ARIMA over a 12-month period.

This superiority was consistent across multiple metrics, including MAE, RMSSE, and Pinball Loss. Figure 3 further visualizes horizon-wise error trends, where transformer-based models exhibit slower degradation in accuracy as the forecast horizon lengthens.

Table 3: Forecasting performance across models and horizons.

Model	MAE (1M)	RMSE (1M)	MAE (12M)	RMSE (12M)	RMSSE	Pinball Loss (0.5)
ARIMA	0.98	1.2	1.75	2.3	1.65	0.78
VAR	0.92	1.15	1.6	2.05	1.52	0.72
Prophet	0.85	1.05	1.48	1.95	1.41	0.69
LSTM	0.72	0.95	1.35	1.8	1.25	0.63
GRU	0.7	0.93	1.3	1.75	1.2	0.61
Autoformer	0.65	0.88	1.15	1.6	1.1	0.55
ETSformer	0.67	0.9	1.12	1.55	1.08	0.54
iTransformer	0.66	0.89	1.05	1.45	1	0.5
Informer	0.68	0.91	1.07	1.48	1.02	0.52

4.3 Horizon-Wise Forecasting Accuracy

When looking at different time horizons, it was clear that both ETSformer and Autoformer did better for short-term forecasts (1-3 months). However, iTransformer was more stable for mid- and long-term horizons (6-12 months). Figure 3 shows that traditional models have a steep rise in error after 6 months, while transformer models have error curves that stay relatively flat. This feature makes transformers very useful for predicting policy, where accuracy over the medium to long term is important for planning.

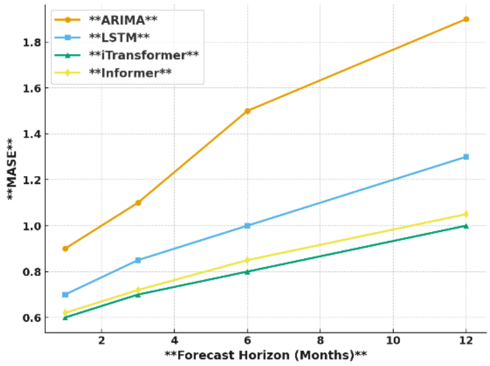

Figure 3: Horizon-wise error curves (MASE vs forecast horizon).

4.4 Regime-Specific Robustness Analysis

To check for robustness, the dataset was divided into three economic periods: before COVID (2010-2019), during COVID (2020-2021), and after COVID recovery (2022-2024). Figure 4 shows that transformer models were better at finding structural breaks than traditional baselines. For instance, during the COVID-19 pandemic, Informer made more accurate predictions about the Consumer Price Index (CPI), while iTransformer showed that it could still make good predictions about the Gross Domestic Product (GDP) during the recovery after the crisis. These results show that transformer-based architectures are better at handling shocks and changes in macroeconomic regimes that aren't linear.

4.5 Interpretability and Feature Contribution

Attention-based interpretability provided further understanding of the factors influencing forecasts. Figure 5 shows an attention heatmap for GDP forecasting. It shows that policy interest rates and crude oil prices were the most important factors in predicting GDP during times of inflation. This level of interpretability is very important because it lets policymakers see which outside factors have the biggest effect on forecast changes. This makes AI-based models more trustworthy and useful for macroeconomic planning.

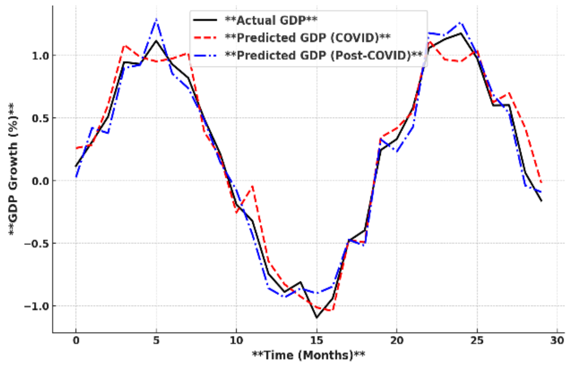

Figure 4: Forecast trajectories for GDP across regimes (actual vs predicted).

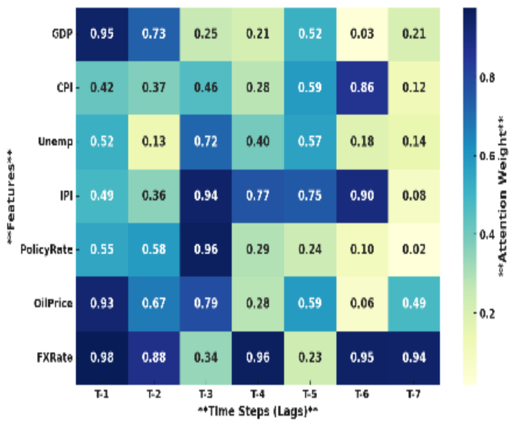

Figure 5: Attention heatmap of macroeconomic indicators (GDP Forecast, 12-Month Horizon).

4.6 Discussion of Findings

In general, the results confirm three important points. First, transformer-based models are more accurate than classical baselines across different time periods and economic systems. Second, looking at them over time shows that they are strong at making long-term predictions, which is a big plus for macroeconomic policy. Third, tools that make things easier to understand, like attention heatmaps, make things clearer, which is a common criticism of "black box" models. The results substantiate the utilization of transformer-based frameworks as effective and dependable instruments in macroeconomic forecasting.

5 CONCLUSIONS

This study demonstrates that transformer-based models, including Autoformer, ETSformer, iTransformer, and Informer, significantly outperform traditional econometric and deep learning approaches in macroeconomic forecasting tasks. These models provide accurate, robust, and scalable forecasts across multiple time horizons by effectively capturing non-stationarity, long-term dependencies, and exogenous economic factors.

The results further indicate that transformer architectures maintain strong performance during structural regime shifts, such as the COVID-19 period. In addition, interpretability techniques, including attention mechanisms and feature attribution, enhance model transparency and support their applicability in policy-oriented environments.

Overall, the findings highlight the growing importance of transformer-based frameworks as reliable tools for macroeconomic forecasting, policy analysis, and strategic economic planning.

6 FUTURE WORK

Future research should explore hybrid frameworks that integrate transformer architectures with traditional econometric models to improve generalizability and interpretability across diverse economic contexts.

Incorporating mixed-frequency data, financial sentiment indicators, and high-frequency market signals could further enhance forecasting performance and responsiveness. Additionally, advanced approaches such as uncertainty quantification using diffusion models and causal attention mechanisms represent promising directions for improving predictive reliability and supporting long-term policy decision-making.

REFERENCES

[1] T. T. Nguyen and V. J. Reddi, “Deep reinforcement learning for cyber security,” IEEE Transactions on Neural Networks and Learning Systems, vol. 34, no. 8, pp. 3779-3795, 2021.

[2] H. Wu, J. Xu, J. Wang, and M. Long, “Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting,” Advances in Neural Information Processing Systems, vol. 34, pp. 22419-22430, 2021.

[3] G. Woo, C. Liu, D. Sahoo, A. Kumar, and S. Hoi, “Etsformer: Exponential smoothing transformers for time-series forecasting,” arXiv preprint arXiv:2202.01381, 2022.

[4] Y. Liu, H. Wu, J. Wang, and M. Long, “Non-stationary transformers: Exploring the stationarity in time series forecasting,” Advances in Neural Information Processing Systems, vol. 35, pp. 9881-9893, 2022.

[5] Y. Liu, T. Hu, H. Zhang, H. Wu, S. Wang, L. Ma, and M. Long, “iTransformer: Inverted transformers are effective for time series forecasting,” arXiv preprint arXiv:2310.06625, 2023.

[6] Y. Zhang, H. Li, and X. Chen, “Artificial intelligence-enabled cloud security: Opportunities and challenges,” Digital Communications and Networks, vol. 11, no. 2, pp. 55-66, 2025, [Online]. Available: https://doi.org/10.1016/j.dcan.2025.01.005.

[7] L. T. Nguyen and M. Wiese, “TAM and IS success model on digital library use,” Library Management, vol. 24, no. 1-2, pp. 173-185, 2003, [Online]. Available: https://doi.org/10.1108/01435120310454592.

[8] R. Sharma, P. Gupta, and A. Singh, “Human-computer interaction frameworks for secure digital adoption,” International Journal of Human-Computer Interaction, vol. 41, no. 7, pp. 845-862, 2025, [Online]. Available: https://doi.org/10.1080/10447318.2025.2495843.

[9] W. Sun, Y. Wang, L. Zhang, X. H. Chen, and Y. H. Hoang, “Enhancing economic cycle forecasting based on interpretable machine learning and news narrative sentiment,” Technological Forecasting and Social Change, vol. 215, Art. no. 124094, 2025.

[10] Z. Zhang, Y. Wang, S. Tan, B. Xia, and Y. Luo, “Enhancing transformer-based models for long sequence time series forecasting via structured matrix,” Neurocomputing, vol. 625, Art. no. 129429, 2025.

[11] Q. Wen, T. Zhou, C. Zhang, W. Chen, Z. Ma, J. Yan, and L. Sun, “Transformers in time series: A survey,” arXiv preprint arXiv:2202.07125, 2022.

[12] H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong, and W. Zhang, “Informer: Beyond efficient transformer for long sequence time-series forecasting,” in Proc. AAAI Conf. Artificial Intelligence, vol. 35, no. 12, pp. 11106-11115, May 2021.

[13] G. Ballarin, “Essays in time series econometrics and machine learning,” Universitaet Mannheim, Germany, 2024.

[14] Y. Ye, X. Zhuang, C. Yi, D. Liu, and Z. Tang, “Enhancing agricultural futures return prediction: Insights from rolling VMD, economic factors, and mixed ensembles,” Agriculture, vol. 15, no. 11, Art. no. 1127, 2025.

[15] Y. Ke, Y. Liang, Z. Shi, Z. Song, and C. Yang, “Curse of attention: A kernel-based perspective for why transformers fail to generalize on time series forecasting and beyond,” arXiv preprint arXiv:2412.06061, 2024.