Proceedings of ICEMGD 2025 Symposium: Innovating in Management and Economic Development
DOI: 10.54254/2754-1169/2025.LH23829

Analysis of Short-Term Stock Price Prediction of CATL Based on the ARIMA Model and Influencing Factors

Xiaoyu Shen

College of Chemistry and Life Sciences, Beijing University of Technology, Beijing, China
XiaoyuShen@emails.bjut.edu.cn

Abstract: The new energy industry, as a core driver of the global low-carbon economic transition, the study of its stock price fluctuation patterns can offer empirical support for the application of financial time series analysis theory in non-stationary and non-linear scenarios. This research focuses on the short-term stock price fluctuations in the new energy sector. Taking the Contemporary Amperex Technology Co.,Ltd (CATL) as the object, and explores the applicability and optimization method of prediction accuracy of the ARIMA model. Based on the ARIMA model, the daily frequency closing price data of CATL for all trading days from February 28, 2023 to March 31, 2025 were selected. The optimal parameter combination was determined by integrating the ADF stationarity test, difference operation, and AIC criterion, and the Box-Ljung test was employed. Ultimately, the ARIMA model with the optimal coefficients was utilized for short-term stock price prediction. The study further incorporated exogenous influencing factors such as policies and supply chains to analyze the multi-dimensional driving mechanisms of stock price fluctuations in new energy enterprises. The empirical findings indicate that the ARIMA model demonstrates high effectiveness in the short-term stock price prediction of CATL. The model successfully captured the stock price fluctuation trend in March 2025. Nevertheless, long-term predictions require the combination of other models to address convergence limitations. It provides methodological support for investors to quantitatively assess short-term trading risks in highly volatile industries and lays a practical foundation for the model construction of stock price prediction in the new energy field.

Keywords: ARIMA, CATL, Stock Price Prediction

1. Introduction

In recent years, amid the severe global climate issues, the global demand for fossil fuels has reached its peak and begun to exhibit a downward trend [1]. In the context of the accelerating transformation of the global low-carbon economy, the new energy industry has emerged as a core domain driving the development of the capital market. Relevant studies indicate that it is projected that by 2050, renewable energy might constitute two-thirds of the global primary energy supply [2]. The current expansive development prospects in the new energy field are stimulating the increase in stock prices of related enterprises. Nevertheless, the stock prices of new energy enterprises still display significant non-stationarity and short-term autocorrelation characteristics due to multiple factors such as policy sensitivity, technological iteration, and supply chain fluctuations. This leads to challenges for traditional prediction models (e.g., linear regression, exponential smoothing) in forecasting new energy stock prices as they are unable to effectively handle the cointegration relationship of non-stationary sequences and the impact mechanism of external shocks. Hence, there is a need to explore alternative prediction methods. CATL, as a core enterprise within the new energy industrial chain, is representative of the sector and possesses rich enterprise-related data. The technological innovation capacity and extremely high market status of CATL render it a typical sample for studies in the new energy domain. The stock price fluctuations of CATL can be utilized to analyze the overall trend of the industry. Exploring the short-term stock price prediction methods of CATL not only provides quantitative support for investment decisions in high-growth industries but also holds significant value for the application and expansion of the modeling theory of non-stationary financial time series.

Current research commonly employs models such as ARIMA and LSTM for stock price prediction [3]. However, these stock price prediction models are largely concentrated in traditional industries or scenarios with stationary sequences, leaving a gap in the new energy sector. Some scholars have attempted to capture the volatility characteristics of new energy stock prices through GARCH models [4]. Nevertheless, few studies have systematically integrated the enhancement effect of exogenous policy variables and supply chain dynamics on prediction accuracy. Existing literature has certain limitations, including insufficient consideration of the distinctive policy-driven mechanism of the new energy industry; inadequate analysis of the transmission effects of raw material price cycles and geopolitical risks; and the quantification methods of market sentiment factors in short-term predictions remain immature.

This paper takes CATL as the research object and constructs an analysis framework integrating the ARIMA model with multi-dimensional influencing factors. It focuses on addressing how to enhance the adaptability of the ARIMA model to non-stationary new energy stock price sequences through parameter optimization. The research is conducted through a three-stage progressive analysis. Firstly, based on the daily stock price data from 2023 to 2025, the non-stationarity issue is addressed through ADF tests and difference operations, and the ARIMA parameter combination is optimized using the AIC criterion. Secondly, the Box-Ljung test is employed to verify the white noise characteristics of the residuals, and the prediction performance is evaluated using indicators such as RMSE and MAPE. Finally, the stock price fluctuations were analyzed through incorporating relevant policies and supply chain resilience indicators, among other influencing factors. The research provides a dynamic risk warning model for short-term trading strategies in high-volatility industries based on the ARIMA prediction system and analyzes the core driving mechanisms of new energy stock price fluctuations.

2. Methodology

2.1. Basic principles of ARIMA model

Auto-Recursive Integrated Moving Average (ARIMA) model is a series analysis method based on the statistical characteristics of time series, which can be used for short-term prediction of stock prices. In the ARIMA(p,d,q) model, the parameter p denotes the autoregressive order in the AR component, q indicates the moving average term in the MA part, and d stands for the degree of differencing applied to make the non-stationary time series stationary [5].

The mathematical representation of the ARIMA(p,d,q) model can be expressed as:

$$y_t = \theta_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + \alpha_p y_{t-p} + \delta_1 \epsilon_{t-1} + \delta_2 \epsilon_{t-2} + \cdots + \delta_q \epsilon_{t-q} \quad (1)$$

y_t represents the stationary sequence of the time series y after undergoing differencing of order d. θ_0 represents the baseline level of the model. $\alpha_1, \alpha_2, \dots, \alpha_p$ are the coefficient of the autoregressive (AR) term, representing the influence of historical observations on the current value. $\delta_1, \delta_2, \dots, \delta_q$ are the coefficient of the moving average (MA) term, representing the corrective effect of historical errors

DOI: 10.54254/2754-1169/2025.LH23829

on the current value. $y_{t-1}, y_{t-2}, \ldots, y_{t-p}$ represents the historical observations lagged from period 1 to period p. $\epsilon_{t-1}, \epsilon_{t-2}, \ldots, \epsilon_{t-p}$ represents the white noise error terms lagged from period 1 to period q, with a mean of 0 and a constant variance.

The stock prices of new energy enterprises represented by CATL have fluctuated considerably as a result of adjustments in relevant policies and technological breakthroughs. Consequently, the ARIMA model is appropriate for conducting short-term stock price predictions for CATL. The difference operation contained in the ARIMA model can effectively eliminate the long-term trend interference brought about by the price cycle of battery raw materials like lithium carbonate. The price jumps triggered by market sentiment and unexpected events can also be captured by the MA term in the ARIMA model. Thereby, more precise predictions of short-term stock prices can be accomplished via the ARIMA model [6].

2.2. Data selection and source

This research has amassed the daily closing prices of CATL for all trading days from February 28, 2023 to March 31, 2025 as the entire sample dataset. The sample period encompasses a relatively complete industrial cycle. It incorporates key events such as the sharp plunge in lithium carbonate prices in the second quarter of 2023, the subsidy withdrawal in 2023, and the new energy tariff dispute in the EU in 2024. This enables adequate training of the model's adaptability to extreme market fluctuations. The data ranging from February 28, 2023 to February 28, 2025 were designated as the training set. The data from March 1, 2025 to March 31, 2025 were selected as the test set, which is utilized for comparison with the predicted data to analyze the accuracy of the prediction model. The data were sourced from the database of Investing.com. This platform enjoys high industry recognition and has a low error rate in historical data backtracking. As a listed company, CATL's historical stock prices, financial reports, R&D investments, and other data are all publicly accessible and transparent, providing data support for the impact factor analysis component of this research.

2.3. Data processing

Data processing is a fundamental step to make the dataset comply with the requirements of the stock price prediction model. The presence of non-trading days in the stock market results in the occurrence of blank records for some dates in the original dataset. To guarantee the continuity of the time series, it is necessary to handle the missing values and outliers in the data. As depicted in Figure 1, it is the time series graph of the closing price of CATL's stock after enhancing the data quality and continuity.

Figure 1: Time series of CATL stock closing price (photo credit: original)

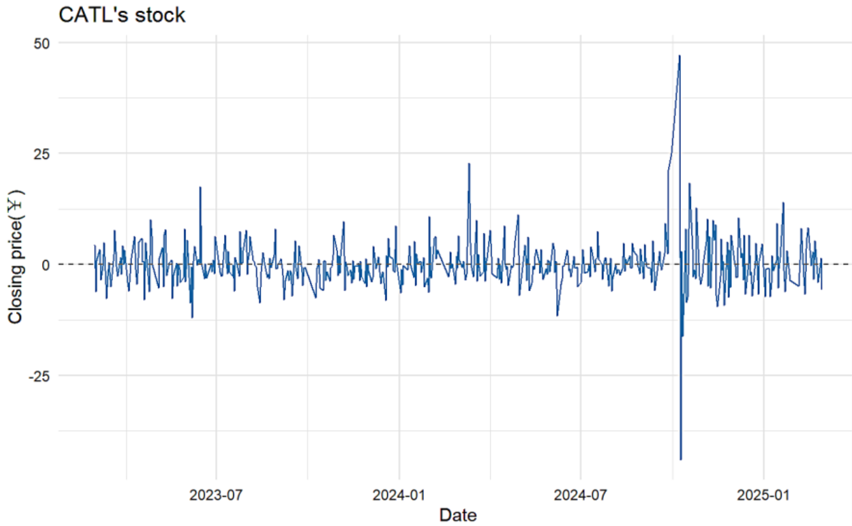

Figure 2: First difference time series (photo credit: original)

As illustrated in Figure 1, the stock price exhibits substantial irregular fluctuations over time. The original sequence demonstrates instability, a characteristic further validated by the ADF test (see Table 1). To achieve stationarity in the sequence, a differencing process is necessary. Figure 2 presents the time series graph of the first-order difference, where it can be observed that the sequence fluctuates around zero with small amplitude. The ADF test results (see Table 1) reveal that the p-values are significantly below the 5% significance threshold, thereby confirming the stationarity of the data and satisfying the fundamental requirements of the ARIMA model for the dataset.

Table 1: ADF test

Pre differential data				Differential data			
	Lag	ADF	p-value		Lag	ADF	p-value
No drift no trend	0	0.0981	0.672	No drift no trend	0	-23.29	0.01
	1	0.1029	0.674		1	-15.74	0.01
	2	0.1458	0.686		2	-14.80	0.01
	3	0.2511	0.716		3	-11.23	0.01
	4	0.1489	0.687		4	-9.76	0.01
	5	0.1742	0.694		5	-9.96	0.01
With drift no trend	0	-1.393	0.562	With drift no trend	0	-23.27	0.01
	1	-1.288	0.600		1	-15.73	0.01
	2	-1.292	0.598		2	-14.79	0.01
	3	-0.999	0.702		3	-11.22	0.01
	4	-1.186	0.636		4	-9.76	0.01
	5	-1.212	0.627		5	-9.96	0.01
With drift and trend	0	-1.65	0.727	With drift and trend	0	-23.31	0.01
	1	-1.56	0.762		1	-15.77	0.01
	2	-1.55	0.770		2	-14.85	0.01
	3	-1.29	0.880		3	-11.29	0.01
	4	-1.46	0.805		4	-9.82	0.01
	5	-1.47	0.803		5	-10.03	0.01

Following the differencing process, the Box-Ljung test was conducted on the data. The obtained p-values were all less than 0.05 (see Table 2). This result indicates that the sequence does not conform to white noise and retains predictable autocorrelation patterns. These discoveries sustain the necessity of subsequently determining the appropriate q and p values, thereby constructing a complete ARIMA model.

Table 2: Box-Ljung test

Pre differential data		Differential data	
df=6	p-value<2.2e-16	df=6	p-value=0.0018
df=12	p-value<2.2e-16	df=12	p-value=0.0049

3. Results

3.1. Determine the ARIMA model

During the process of determining the p and q values for the model, preliminary estimates of p equals 3 and q equals 2 were established by analyzing the truncation and tailing characteristics of the ACF and PACF, as illustrated in Figures 3-4.

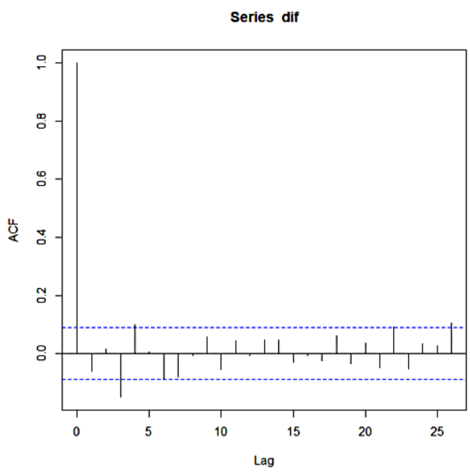

Figure 3: ACF plot (photo credit: original)

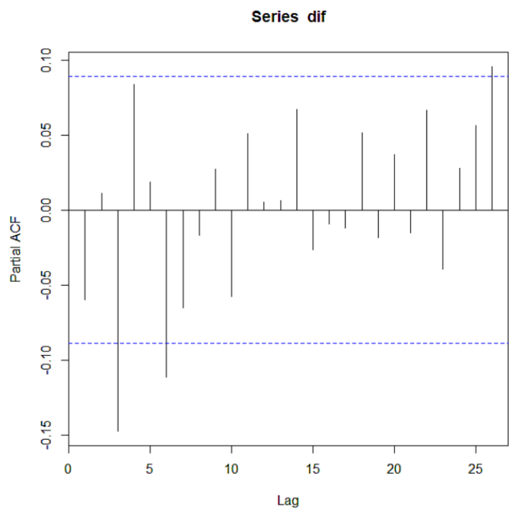

Figure 4: PACF plot (photo credit: original)

DOI: 10.54254/2754-1169/2025.LH23829

Thereafter, the AIC criterion was utilized to validate these selections. By systematically comparing the AIC function values across various orders of ARIMA models, the ARIMA(3, 1, 2) model was ultimately chosen as the optimal fitting model. The model passed the residual independence test (see Table 3), indicating that the ARIMA(3,1,2) fit to the data is reasonable. The significance of all AR and MA term coefficients (see Table 4) suggests that the model structure is reasonable.

Table 3: Box-Ljung test (ARIMA(3,1,2)

Q*	df	p-value
6.431	5	0.2665

Table 4: Parameter estimation of ARIMA(3,1,2) model

	AR(1)	AR(2)	AR(3)	MA(1)	MA(2)
	-0.5044	-0.6416	-0.1825	0.4561	0.6321
s.e.	0.1375	0.1689	0.0479	0.1358	0.1687

3.2. ARIMA model validation

The hypothesis that residuals are uncorrelated is the fundamental premise for the construction of ARIMA models. To ensure the reliability of the model, it is essential to conduct a white noise test. If the residuals fail to exhibit white noise characteristics, this suggests that the model is inadequate and requires further refinement. As shown in Figure 5, the residual time series are randomly distributed around zero, with no apparent trends or periodic patterns. The ACF graph demonstrates that the vast majority of the autocorrelation coefficients at diverse lag orders are within two standard deviations and are statistically insignificant. Additionally, the residual histogram demonstrates that the residuals approximate a normal distribution. To sum up, this sequence is a white noise sequence, and the model chosen for this project is rational.

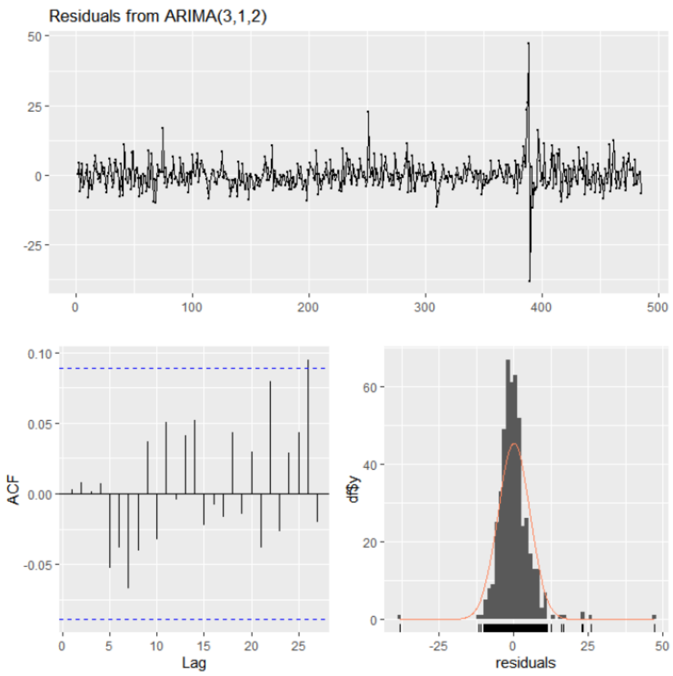

Figure 5: Residual test of the sequence (photo credit: original)

DOI: 10.54254/2754-1169/2025.LH23829

3.3. Model fitting and evaluation

In this study, the daily closing price data for all trading days of CATL from February 28, 2023, to February 28, 2025, were selected as the training set. The ARIMA (3,1,2) model was employed to fit the training set. Subsequently, the Ljung-Box test was applied to analyze the residual series of the fitted model. The results indicated that the p-values for all lag orders exceeded twice the standard deviation (see Figure 6), which demonstrates that the residual series of the fitted model can be considered a white noise series. Therefore, it is concluded that the constructed ARIMA (3,1,2) model exhibits significant effectiveness and fits the training set data well.

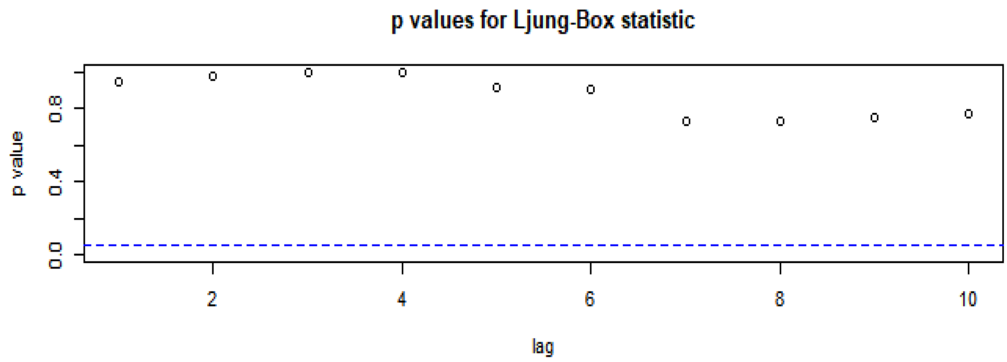

Figure 6: Fitting model of residual error sequence Ljung - box test (photo credit: original)

To further assess the predictive capacity of the ARIMA (3,1,2) model, in this study, the model was employed to predict the daily closing prices of CATL for all trading days throughout March 2025. A detailed comparison between the prediction results and the actual closing prices is presented (see Figure 7), where the red dotted line represents the actual values, and the black solid line represents the predicted values, providing an intuitive visualization of the model's prediction performance. It can be observed from the time series plot that the predicted values eventually converge to a constant value.

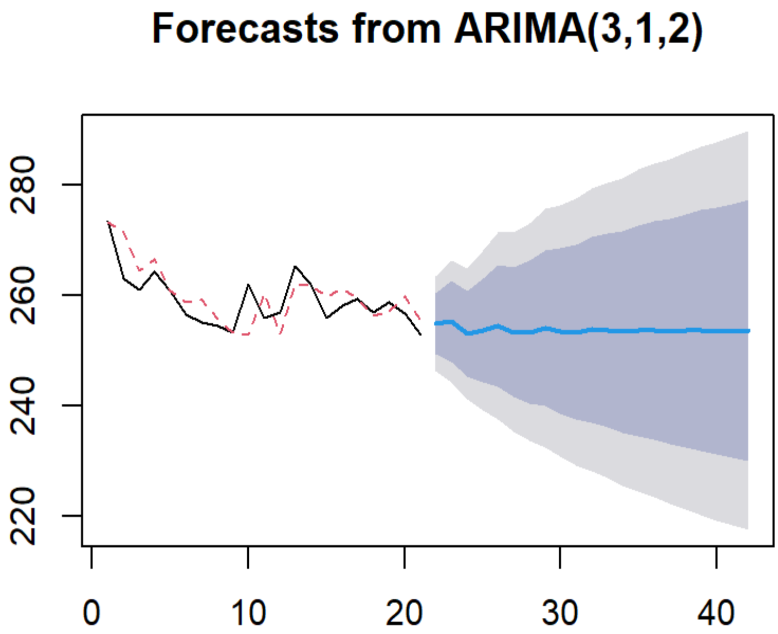

Figure 7: Comparison of stock price forecast and actual value in March 2025 (photo credit: original)

DOI: 10.54254/2754-1169/2025.LH23829

After computation, the root mean square error (RMSE) of the model's prediction results was approximately 7.08, indicating a relatively small deviation between the predicted and actual values. The Mean Absolute Percentage Error (MAPE) is approximately 2.12%. This implies that when the actual stock price is approximately 260 yuan, the prediction deviation is around 5.51. It demonstrates that the deviation between the predicted price and the actual price is relatively small in relation to the stock price itself, and this model possesses a relatively high prediction accuracy (see Table 5). Based on the above analysis, it has been demonstrated that the model exhibits a good forecasting performance.

Table 5: Results of model accuracy evaluation

MSE	RMSE	MAE	MAPE (%)
50.0787	7.0766	5.5690	2.1192

4. Discussion

In the context of the rapidly evolving global energy structure, the new energy sector has become a pivotal area that spurs investment in the capital market. Under the trend of a low-carbon economy, which emphasizes greater environmental protection and lower consumption of new energy resources, the industry holds very promising development prospects. The rise of new energy vehicles has significantly promoted the increase in share prices of new energy companies. However, the factors influencing stock price fluctuations in new energy enterprises are highly complex. Essentially, the volatility of new energy enterprise share prices is the result of multiple factors acting together, including policy environment, technology iteration, supply chain resilience, and market sentiment.

4.1. Analysis of core influencing factors of stock price volatility

The new energy industry is highly sensitive to policy intervention. Key variables such as policy subsidies, carbon tariffs, and technical standards can cause fluctuations in share prices by altering corporate earnings expectations. Changes in international trade policies may raise market concerns about supply chain stability, leading to short-term stock price volatility. Furthermore, differences in carbon neutrality and energy transformation policies can affect market access and competitive patterns for enterprises, thereby impacting the share prices of new energy companies.

Favorable policy support typically propels the virtuous cyclic development of new energy enterprises [7]. The more substantial the policy support is, the more funds’ enterprises have for technological research and development [8]. Technology iteration is another core variable driving stock price growth. New technological achievements will further facilitate the advancement of enterprises. Consequently, the level of enterprise valuation is positively correlated with the intensity of R&D investment. New energy enterprises with leading technologies generally receive higher market expectations, further enhancing their stock price performance. However, some research indicates that there exists an inverted U-shaped relationship between new energy subsidies and enterprises' innovation investment [9]. Namely, subsidies can effectively stimulate enterprises' innovation within a certain scope. However, beyond the critical point, the marginal effect of subsidies will turn negative and restrain innovation. Hence, a reasonable policy subsidy range is more beneficial for the development of new energy enterprises and the enhancement of their stock prices.

The cost-price fluctuation mechanism of raw materials also affects corporate profits through the supply chain. Supply chain stability determines a firm's ability to adapt to market changes [10]. For example, fluctuations in the prices of key mineral resources such as lithium and cobalt can impact enterprise profit margins. The arrangement of supply chain regionalization is also an important

DOI: 10.54254/2754-1169/2025.LH23829

consideration. For example, local manufacturing can decrease geopolitical risks and strengthen a company's ability to withstand risks.

Additionally, market sentiment is an unstable factor affecting new energy companies [11]. Short-term stock price volatility is highly correlated with social media public opinion. Retail investor behavior may amplify market volatility, creating short-term bubbles or excessive declines in stock prices. Since market sentiment is often driven by unexpected events, conventional forecasting models struggle to capture this factor, potentially leading to reduced forecasting performance.

4.2. Investor strategy suggestions

Given the complexity and high volatility of stock prices in the new energy industry, it is recommended that investors adopt a combination of long-term configuration and short-term trading strategies to ensure a balance between risk and return. Investors should prioritize enterprises with a high proportion of R&D investment, a strong patent portfolio, stable upstream raw material supplies, and reliable downstream market demand. At the same time, it is advisable to select enterprises that have comparatively lower geopolitical risks. In short-term stock trading, investors must pay close attention to national policies and anticipate short-term market sentiment changes to identify trading opportunities.

At the same time, it is advisable to select enterprises that have comparatively lower geopolitical risks. In short-term stock trading, investors must pay close attention to national policies and anticipate short-term market sentiment changes to identify trading opportunities. Investors can leverage ARIMA forecasts to identify optimal timing for buying and selling. In specific forecasting, exogenous variables such as policy dummy variables and raw material prices can be incorporated to improve model prediction accuracy.

5. Conclusion

This paper conducts predictions on the short-term price fluctuations of CATL stocks by employing the ARIMA model and investigates its applicability and optimization approaches in the new energy sector in conjunction with external factors like policies and supply chains. The model's prediction outcomes reveal that it exhibits an excellent short-term predictive performance and can effectively capture the stock price fluctuation trend in March 2025. This paper conducts an analysis of how factors such as policies, technological iterations, supply chain stability, and market sentiment form the core elements influencing the stock price fluctuations of new energy enterprises. Consequently, it is proposed that investors integrate long-term allocation and short-term trading strategies, focus on new energy enterprises with strong R&D capabilities, stable supply chains, and low geopolitical risks, and utilize the ARIMA model to aid in decision-making. This study offers methodological support for stock predictions in the new energy domain and provides a dynamic risk warning model for investors.

From the experimental results, it is observable that the predicted values of the ARIMA model will ultimately converge to a constant. This is attributed to the fact that the model is based on the assumption of stationarity, and the predicted values will gradually lose their sensitivity to short-term fluctuations and eventually tend toward the long-term mean of the sequence. This characteristic renders the ARIMA model more suitable for short-term predictions. However, for long-term predictions, it is requisite to combine with other methods to enhance prediction accuracy. For instance, by integrating with other models such as LSTM or GARCH. Through the introduction of exogenous variables, for example, taking policy variables and raw material prices as input variables, the accuracy of the prediction model can be enhanced.

DOI: 10.54254/2754-1169/2025.LH23829

References

[1] *Raimi, D., Zhu, Y., Newell, R. G., & Prest, B. C. (2024). Global Energy Outlook 2024: Peaks or Plateaus. Resources for the future.*

[2] *Hassan, Q., Viktor, P., Al-Musawi, T. J., Ali, B. M., Algburi, S., Alzoubi, H. M., ... & Jaszczur, M. (2024). The renewable energy role in the global energy Transformations. Renewable Energy Focus, 48, 100545.*

[3] *Su, S. (2023). Stock price prediction model comparison based on LSTM and ARIMA (Master's thesis, Chongqing University). https://link.cnki.net/doi/10.27670/d.cnki.gcqdu.2023.003380.*

[4] *Xu, W. & Zhou, Y. (2021). New energy stock price research based on VAR and GARCH models. China Economic and Trade Herald (Middle), (01), 98-100.*

[5] *Cao, L. & Zhang, R. (2024). Stock price prediction based on ARIMA model: A case study of Muyuan Foods. Wealth Times, (02), 54-56.*

[6] *Huang, S. (2022). Stock price analysis and prediction based on ARIMA model: A case study of China Merchants Bank. Small and Medium Enterprise Management and Technology, (11), 184-187.*

[7] *Jiang, Z., & Xu, C. (2023). Policy incentives, government subsidies, and technological innovation in new energy vehicle enterprises: Evidence from China. Energy Policy, 177, 113527.*

[8] *Zhang, Z., Luo, C., Zhang, G., Shu, Y., & Shao, S. (2024). New energy policy and green technology innovation of new energy enterprises: Evidence from China. Energy Economics, 136, 107743.*

[9] *Wu, Z., Fan, X., Zhu, B., Xia, J., Zhang, L., & Wang, P. (2022). Do government subsidies improve innovation investment for new energy firms: a quasi-natural experiment of China's listed companies. Technological Forecasting and Social Change, 175, 121418.*

[10] *Yang, X. & Zhang, R. (2024). How supply chain finance affects stock price synchronicity. Modern Financial Research, 29(12), 70-80.*

[11] *Cai, J. & Wang, A. (2025). Stock price prediction model based on investor sentiment and stacked autoencoder. Journal of Harbin University of Commerce (Natural Science Edition), 41(01), 120-128.*