Literature review on the application of ARMA model in stock price prediction

Bowen Zhang [1*]

[1] University of Toronto, Faculty of Arts & Science, Toronto, ON, Canada

Abstract: Predicting stock prices is a perennial quest in finance, yet price series are famously noisy, volatile and nonlinear. Among the many tools on offer, the autoregressive-moving-average (ARMA) model remains a surprisingly resilient workhorse. This review first sketches the logic of ARMA and the Box-Jenkins recipe for transforming raw prices into stationary returns, selecting lags and checking residuals. A broad body of evidence especially for 1 to 5 day horizons, confirms that ARMA delivers reliable, low-cost forecasts and, when paired with GARCH, can track volatility bursts with notable precision. Head-to-head studies show that on small samples or thin markets, ARMA often rivals much heavier deep-learning engines, while recent hybrids such as ARMA-LSTM and ARMA-Transformer marry linear transparency with nonlinear flexibility and shine on high-frequency data. We synthesise domestic and global findings, chart three clear trends, model fusion, finer time grids and AutoML optimisation, and flag the model's blind spots: fixed coefficients, linear assumptions and sparse use of unstructured signals. Looking ahead, regime-switching ARMA, online updating, sentiment-rich inputs and risk-band forecasts (e.g., VaR, CVaR) promise to keep this classic framework both relevant and insightful.

1 Introduction

Stock price prediction has always been an important research direction in the financial field. However, due to the high volatility, nonlinearity and high noise of stock prices, accurate prediction faces great challenges [1]. As a classic time-series analysis method, the ARMA model (autoregressive moving average model) is widely used in the modeling and prediction of financial time series.

This article will review the application of the ARMA model in stock price prediction, including: the basic principles and structure of the ARMA model, the application effect of the model in stock market prediction, comparison with other models, representative research and conclusions at home and abroad, research trends in recent years, and the limitations of current research and prospects.

2 Principles and Structure of ARMA in Financial Forecasting

The ARMA model is one of the most basic models in time-series measurement, which is composed of an autoregressive (AR) part and a moving average (MA) part [2]. Specifically, the ARMA (p, q) model uses the p-order autoregressive term to characterize the lag correlation of the series itself, and the q-order moving average term to characterize the impact of random

[*] Corresponding author: messi.zhang@mail.utoronto.ca

disturbances on the series [3]. The ARMA model assumes that the series is stationary, that is, the mean and variance are constant; for non-stationary financial time series, it is necessary to differentiate them to make them stationary before modeling. Currently, the ARMA model is expanded to an ARIMA model that includes the difference order.

The ARMA model usually follows the Box–Jenkins methodology, and the modeling steps include model identification, parameter estimation, diagnostic testing, and predictive application. First, the model orders p and q are determined through the time-series diagram and the autocorrelation function (ACF) and partial autocorrelation function (PACF). Then, the parameters are estimated using methods such as maximum likelihood, and the independence and normality of the residual sequence are tested. After the model passes the test, it can be used to predict stock prices in the future. In financial time-series analysis, the ARMA model is favored because it follows the principle of "characterizing the data-generation process with the simplest model" and can capture the main linear dependency structure with fewer parameters [4].

3 Key Domestic and International Studies and Findings

Domestic empirical work on the ARMA model focuses on the daily data of a single stock or index and tends to use a rolling window to evaluate the short-term (1–5 days) forecasting performance. Gao Yuan used the closing prices of LeTV from 2010 to 2014 as samples and applied ARMA (1, 1) to make rolling forecasts of logarithmic returns. The mean absolute error (MAE) decreased by about 12 % and was significantly better than the benchmark random-walk model in the 3-day and 5-day forecast periods; the author pointed out that the model effect is sensitive to the window length. Xie Haining and Shang Yuan examined the daily prices of CITIC Bank from 2008 to 2019, compared ARMA, exponential smoothing and support-vector machines, and found that ARMA (2, 1) had the lowest mean-square error (MSE) for 1-day rolling forecasts, and adapted well to the turbulent period of 2015-2016 after the financial crisis [5]. Peng Zijun further introduced GARCH (1, 1) to characterise conditional variance and applied ARMA (1, 1)–GARCH to the Shanghai Composite Index return. In the VaR test (99 % confidence), the number of violations of the hybrid model was reduced from 15 times for pure ARMA to 7 times, demonstrating the risk-measurement advantage of joint modelling [6].

Foreign research focuses more on cross-market and cross-model system comparisons. Sharma et al. conducted 30-day rolling training on ARIMA, Prophet, LSTM and FBProphet based on the 2010–2023 data of four technology stocks: AAPL, MSFT, NFLX and GOOGL. The experiment showed that on the logarithmic return after stabilization, the RMSE of the optimal ARIMA (0, 1, 1) was only 4 % higher than that of LSTM, while the calculation time was shortened by more than 90 %. Pham et al. analysed 865 articles published between 1970 and 2022 and pointed out that ARMA/ARIMA is still the most-used benchmark in deep-learning studies (63 %), often outperforming complex networks in low-liquidity markets or short-sample settings [7]. Alaminos et al. used one-minute data from the Spanish IBEX35 index, fed ARMA–GARCH residuals into a two-layer perceptron to predict 5-minute excess returns, and reported a Sharpe-ratio gain of 0.27. It shows that layering linear and nonlinear components can enhance high-frequency trading strategies. In summary, domestic and foreign studies agree that ARMA concisely captures linear short-term signals, while its weaknesses are largely mitigated when coupled with GARCH or deep networks.

4 ARMA Applications and Performance in Stock Forecasting

The application of ARMA model in the stock market generally requires converting the stock-price series into a stationary series (for example, taking the logarithmic difference of the daily closing price, or directly modelling the return series) to meet the model assumptions. Researchers usually use historical stock-price data to fit the ARMA model and use it to predict the stock-price trend in the short term to provide a reference for investment decisions [8]. Owing to the complexity of the stock market, the ARMA model is often combined with other methods: for example, for stock-return series, a common practice is to establish an ARMA model to characterise the mean structure and superimpose a GARCH model to describe volatility, thereby forming an ARMA-GARCH combination model that captures short-term price changes and reflects volatility clustering.

Many empirical studies show that the ARMA model is effective for short-term stock-price forecasts. Gao Yuan modelled the closing price of LeTV.com with an ARMA model and found the results instructive for investors; Xie Haining and Shang Yuan used CITIC Bank's stock price to demonstrate that the ARMA model can better predict short-term price fluctuations. In addition, Peng Zijun combined ARMA with GARCH for the Shanghai Composite Index, showing that the ARMA–GARCH model more accurately captures short-term volatility and aids financial risk management.

Overall, these cases indicate that the ARMA model excels at depicting short-term stock dynamics, captures autocorrelation, and yields relatively low forecast errors. When the market is stable or exhibits short technical cycles, its predictions are often reliable. However, performance declines sharply over longer horizons because many nonlinear factors drive stock prices, and linear models struggle to capture long-run trends [9]. This supports the efficient-market view that long-term price movements resemble a random walk, leaving limited predictability beyond the short term.

5 Research trends in recent years

In the past five years, stock-prediction models have evolved rapidly, showing three significant trends: integration, high frequency and automation. First, the boundaries between statistical models and deep learning are continually being broken, becoming a new hot spot for academic exploration. An increasing number of studies integrate ARMA or ARIMA with neural networks such as LSTM and Transformer: some model residual sequences, others adopt multi-scale decomposition strategies, markedly boosting the ability to fit nonlinear and non-stationary data. Second, the explosive growth of high-frequency-trading data has pushed prediction models toward ever-finer time granularity. Researchers now embed ARMA into more flexible composite architectures; for example, an ARMA–GARCH–neural-network triple structure has outperformed traditional methods in minute- or even second-level trading strategies. Third, a quiet "automation revolution" in model building and optimisation is under way. From genetic algorithms to AutoML platforms, automatic search tools are widely used for order selection, feature engineering and hyper-parameter tuning, greatly lowering the barrier to hybrid models and shortening the path from modelling to deployment. This evolution not only makes models smarter, but also better aligned with the timeliness and practicality of real-world financial decision-making.

From the above, it is not difficult to observe that future stock price prediction is moving towards multidisciplinary intersection, real-time response and high adaptability, and the ARMA series of models are constantly injecting new blood into this form, keeping up with the evolution of the times, and playing an indispensable role. How to maintain interpretability in fusion and robustness in automation will become the key issue of the next stage of research.

6 Comparative study with other models in stock price prediction

When comparing ARMA with other mainstream forecasting models, a comprehensive evaluation can be made from four dimensions: linear processing capability, adaptability to non-stationarity, variable dimension and capture of non-linear characteristics. First, ARIMA adds a differencing step to ARMA to eliminate unit roots and stabilise non-stationary series; thus, for stock-price series with pronounced trends or random-walk components, ARIMA often reflects long-term trends better than ARMA, albeit at the cost of greater model complexity and parameter uncertainty. Second, the GARCH family focuses on the conditional-variance dynamics of returns and captures the volatility clustering common in financial time series, whereas ARMA describes the mean equation; combining the two (ARMA–GARCH) delivers both price expectation and risk range within one framework, making it a staple for quantitative risk management and short-term forecasting. Third, the VAR model uses multivariate autoregression to capture dynamic linkages between stock prices and macro variables; with well-chosen factors it excels in medium- and long-term scenario analysis, but as variables and lags grow, parameter dimensionality explodes and overfitting looms, without ample data or theory-based constraints, its accuracy may lag that of univariate models. Finally, deep-learning models such as LSTM automatically extract non-linear and long-range dependencies and often achieve higher accuracy on complex or regime-shifting data, yet demand large samples, incur high training cost and offer limited interpretability [10]. Overall, thanks to its simple structure, clear parameters and small-sample friendliness, ARMA remains a lightweight benchmark for short-term stock-price forecasting; in practice it is frequently blended with GARCH or deep-learning algorithms to balance interpretability and performance.

Complementary evidence shows that ARMA has been extended to liquidity-risk settings in China and remains competitive in ultra-short-term forecasting [11]. A recent machine-learning survey likewise confirms that blending ARMA with deep architectures is now a mainstream research avenue [12]. Taken together, these studies reinforce our view that linear foundations still matter, yet future progress will hinge on creative hybrid designs that couple ARMA's transparency with the representational power of modern neural networks.

7 Conclusion

Although ARMA and its extended models have been widely used in stock price prediction, the complexity of the real financial market still poses many challenges to it. First, ARMA is based on a linear framework and is often "powerless" when facing financial crises, black swan events or structural transitions (such as sudden policy changes, geopolitical conflicts, etc.), and cannot timely respond to the nonlinear characteristics of drastic market fluctuations. Second, most model parameters are set to static, ignoring the possibility of multiple states in the market, such as the alternation of bull and bear markets, volatility jumps and other dynamic structural changes. Therefore, in the future, it is urgent to introduce a state switching model (such as Markov-switching ARMA) or a dynamic update strategy based on the Bayesian method to enhance the adaptability and stability of the model.

In addition, traditional models still lack systematic use of unstructured data, such as news headlines, public opinion heat, northbound capital inflows, social media keywords, etc. This leads to blind spots in the model in capturing investor sentiment and market expectations, and it is difficult to truly touch the deep factors driving stock price changes. At the same time, although deep learning methods such as LSTM and Transformer have repeatedly achieved great results in prediction accuracy, they are often accompanied by problems of high computational overhead and poor interpretability, making it difficult to gain the trust of financial practitioners in practical applications.

Looking to the future, researchers can seek breakthroughs from multiple dimensions: building a hybrid architecture that integrates ARMA and deep neural networks (such as LSTM, Transformer), taking into account the advantages of linear structure and nonlinear expression; exploring online learning and incremental update mechanisms to achieve rapid response of models to real-time market changes; introducing multi-source heterogeneous data (such as sentiment indicators, macro signals) to improve the comprehensiveness and explanatory power of predictions; and paying attention to interval predictions and risk metrics (such as VaR, CVaR, etc.) to be closer to actual investment decision-making scenarios.

Reference

1. R. Sharma, S. Hariharan, Stock price prediction using ARIMA: A study on AAPL, MSFT, NFLX and GOOGL stocks. *Proc. Int. Conf. Comput. Intell. Comput. Appl.*, 47–50 (2024)
2. D. Alaminos, M. Salas, Partal-Ureña B., A. Hybrid ARMA-GARCH-Neural Networks for intraday strategy exploration in high-frequency trading. *Pattern Recognit.* **148**, 110139 (2024). https://doi.org/10.1016/j.patcog.2023.110139
3. G. E. P. Box, G. M. Jenkins, G. C. Reinsel, Ljung, G. M. *Time Series Analysis: Forecasting and Control*, 5th ed.; Wiley: Hoboken, 2015.
4. Y. Gao, An empirical study on LeTV stock-price prediction based on an ARMA model. *Mod. Econ. Inf.* **7**, 352–353 (2015)
5. H. Xie, Y. Shang, Stock-price prediction using the ARMA model: Evidence from China CITIC Bank. *Econ. Manag. Sci.* **12**, 122–124 (2020)
6. Z. Peng, ARMA-GARCH-based analysis of Shanghai Composite Index returns. *J. Grad. Sch. Zhongnan Univ. Econ. Law* **4**, 48–55 (2017)
7. H. V. Pham, L. H. Phu, T. H. V. Nguyen et al., A bibliometric literature review of stock price forecasting: From statistical model to deep learning approach. *Sci. Prog.* **107**, 1–23 (2024). https://doi.org/10.1177/00368504241234567
8. C. L. Li, H. Zhang, Y. A Wang, hybrid ARIMA–LSTM approach for stock price forecasting. *J. Financ. Data Sci.* **5**, 45–59 (2023). https://doi.org/10.3905/jfds.2023.1.045
9. W. Bessler, P. Lückoff, Macroeconomic variables and German stock returns: A VAR approach. *Appl. Econ.* **48**, 599–612 (2016)
10. Y. Chen, Application of an ARMA model considering liquidity risk in stock returns. *Henan Sci.* **32**, 655–659 (2014)
11. Z. X. Ma, Application of the ARMA model in short-term stock prediction. *Guangxi Qual. Superv. Bull.* **5**, 152–153 (2020)
12. R. Yan, Machine learning in stock price forecasting: A systematic review. *Financ. Innov.* **10**, 52–68 (2024). https://doi.org/10.1186/s40854-024-00452-9