Creative Economy
2025, VOL. 9, NO. 3, 96-100
DOI: 10.47297/wspceWSP2516-251919.20250903

Stock Market Volatility Forecasting Based on Support Vector Machine and GARCH Model

Jiawei Zhang

Economics collageUniversity of BristolBristol, BS8 1QU, UK

ABSTRACT
Stock market volatility serves as a crucial indicator for assessing financial risk and gauging the level of uncertainty present in the market. The traditional GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model has demonstrated considerable efficacy in modeling the phenomenon of volatility clustering observed in time series data. However, this model encounters significant limitations when it comes to addressing the complexities associated with nonlinear and high-dimensional data sets. In contrast, the Support Vector Machine (SVM), a well-established statistical learning technique, excels in managing scenarios characterized by small sample sizes, high dimensionality, and nonlinear relationships. Recognizing the complementary strengths of these approaches, this paper introduces a novel hybrid forecasting model that seamlessly integrates the SVM methodology with the GARCH framework, aiming to enhance the precision of stock market volatility predictions.To validate the efficacy of this proposed hybrid model, we utilize daily return data from the Shanghai Composite Index and the Shenzhen Component Index, spanning the period from 2017 to 2023. An empirical comparison is conducted between the SVM-GARCH hybrid model and the traditional GARCH model, as well as a standalone SVM model. The empirical results reveal that the hybrid model consistently outperforms the individual models across multiple performance metrics, including predictive accuracy, robustness, and directional correctness. These findings underscore the hybrid model's strong practical applicability and highlight its potential for broader implementation in financial markets.

KEYWORDS
GARCH model; Support vector machine; Volatility forecasting; Hybrid modeling; Financial time series

1 Introduction

Stock market volatility is one of the core indicators for assessing the uncertainty of asset prices. Its predictive power is crucial in areas such as risk management, portfolio allocation, and financial product pricing. With the continuous evolution and increasing complexity of financial markets, traditional statistical methods—such as the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model—have demonstrated strong performance in modeling volatility clustering and conditional heteroskedasticity in time series data. However, these models often struggle to capture the nonlinear and nonstationary characteristics commonly observed in real-world financial data.

In recent years, the rise of machine learning techniques has provided new avenues for financial forecasting. Support Vector Machine (SVM), known for its robust nonlinear modeling capabilities and structural risk minimization principle, has gained increasing traction in financial time series prediction tasks. Due to its strengths in handling high-dimensional, nonlinear, and small-sample data, SVM has shown promising results across various financial forecasting applications.

Integrating SVM with GARCH allows researchers to leverage the advantages of both approaches—GARCH's strength in modeling volatility dynamics and SVM's ability to capture complex nonlinear patterns—thereby enhancing prediction accuracy and model robustness.

This study aims to explore a hybrid forecasting methodology that combines SVM and GARCH models to predict stock market volatility. By effectively capturing the complex dynamics of stock market behavior, the proposed hybrid model is expected to serve as a more accurate and reliable forecasting tool for investors and financial institutions.

2 Main result

2.1 Literature review and research motivation

Volatility in financial markets is a critical subject in financial economics, as it reflects the magnitude and frequency of price fluctuations. It plays a vital role in the pricing of financial derivatives, risk management, and portfolio optimization. With the increasing integration of global economies, market volatility has become more pronounced, making its accurate prediction and effective management more urgent and important.

The GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model, proposed by Bollerslev (1986), remains one of the most widely used tools for modeling financial market volatility. This model effectively captures volatility clustering in financial time series. Chen and Yu (2002) applied GARCH models with different distributional assumptions to assess risks in the Shanghai and Shenzhen stock markets and demonstrated the model's superior performance in capturing the risk characteristics of returns. Additionally, Zheng (1997) was among the first to introduce the Value at Risk (VaR) concept in China, and Wang et al. (2000) elaborated on three common methods for VaR calculation: the variance-covariance method, the historical simulation method, and Monte Carlo simulation. Wu and Chen (1999) applied VaR-based models to analyze financial asset allocation in the Chinese securities market.

In recent years, to address the dependence structure between multiple risk factors, scholars have introduced Copula functions combined with GARCH models. Yang and Xia (2008) conducted empirical research on open-end fund portfolios using a Copula-GARCH framework and found that the model exhibited significant advantages in capturing risk dependencies. Meanwhile, with the rapid development of machine learning, algorithms such as Support Vector Machines (SVMs) have been increasingly applied to financial forecasting tasks. Liu (2020) combined ensemble random forests and SVMs to predict the financial distress of commercial banks, highlighting the potential of SVMs in financial modeling. Gao (2020) further demonstrated that integrating GARCH models with machine learning techniques yields high predictive accuracy and feasibility in forecasting financial trends.

Recent studies have also explored the integration of deep learning models with GARCH. For instance, Michańków et al. (2023) proposed a hybrid model combining deep learning and GARCH for volatility and risk forecasting in financial markets, which showed superior predictive accuracy over standalone models. Chen et al. (2010) introduced a GARCH model enhanced with SVMs that performed excellently in volatility forecasting. Alshammari et al. (2020) developed a hybrid model that incorporated MODWT decomposition with ARIMA-GARCH, which proved effective in forecasting the volatility of the Saudi stock market.

The combination of SVM and GARCH has shown promising results across different markets. Yamaka et al. (2021) demonstrated that the SVM-GJR-GARCH model accurately described volatility in the Thai and Singapore stock markets. Bezerra and Albuquerque (2017) employed a mixture-of-Gaussian-kernel SVR-GARCH model to forecast volatility, achieving strong results for both cryptocurrencies and fiat currencies. Bildirici and Ersin (2013) validated the effectiveness of the SVM-GARCH model in modeling conditional volatility in the Turkish financial market.

Additionally, LSTM-based models have recently been integrated with GARCH to predict cryptocurrency volatility. Banik et al. (2022) showed that the LSTM-GARCH model outperforms traditional GARCH models in high-frequency cryptocurrency environments. Kim and Won (2018) proposed a new hybrid framework combining LSTM and multiple GARCH-type models, which significantly improved prediction accuracy for stock price volatility.

Despite extensive research in this domain, volatility forecasting remains a challenging task. Traditional GARCH models exhibit limitations in handling nonlinear and nonstationary financial time series, while standalone machine learning methods may struggle to capture temporal dependencies. Zou et al. (2002) pointed out that GARCH models often underestimate the heavy-tailed nature of empirical financial data. Therefore, integrating statistical models with machine learning techniques to leverage their respective strengths and enhance predictive accuracy remains a pressing and complex issue.

This study aims to develop an efficient volatility forecasting framework by combining Support Vector Machines (SVM) and GARCH models. By integrating the time-series modeling capability of GARCH with the high-dimensional nonlinear modeling strength of SVM, we aim to improve the accuracy of financial market volatility forecasts and provide valuable tools for risk management and investment decision-making. The findings of this research contribute both theoretically and practically by enhancing the risk prediction capacity of financial systems.

The structure of the study comprises six key components: model development, data preparation, methodology implementation, empirical analysis, robustness evaluation, and performance comparison. Each part is designed to systematically validate the effectiveness of the proposed hybrid SVM-GARCH model.

2.2 Model construction

We designed a sequential hybrid modeling framework that organically integrates traditional statistical modeling with modern machine learning techniques. Specifically, in the first stage, we apply GARCH-based models—including GARCH(1, 1) and EGARCH—to capture the volatility clustering characteristics inherent in financial time series. The resulting residuals from this stage are treated as intermediate variables. In the second stage, a Support Vector Machine (SVM) model is employed to perform nonlinear modeling using a set of high-dimensional features constructed from the GARCH residuals, lagged returns, and other volatility indicators.

This architecture is intended to overcome the limitations of the GARCH model in capturing complex and nonlinear structures, while preserving its strengths in modeling time-dependent behavior. By combining the complementary advantages of both approaches, the proposed framework establishes a clear and cohesive hybrid prediction system that enhances forecasting accuracy and robustness.

2.3 Data sources and preprocessing

The raw data used in this study consists of daily closing prices of the Shanghai Composite Index (SSE) and the Shenzhen Component Index (SZSE) from January 1, 2017, to December 31, 2023. We calculated the logarithmic returns from the price series as the fundamental variable for modeling. Standard preprocessing procedures were applied, including the handling of missing values and outliers, followed by z-score normalization to standardize the data.

In the feature engineering stage, we constructed a set of technical indicators such as lagged returns, moving averages, and rolling standard deviations. Additionally, market trading volume and historical volatility measures were incorporated to enrich the input dimension of the model. These features are intended to capture both short-term price fluctuations and long-term volatility patterns.

Finally, the dataset was split chronologically into training and testing sets, with 80% of the data used for model training and the remaining 20% reserved for performance evaluation. This setup reflects a realistic forecasting scenario and ensures that future data is not used in the training phase, thus preventing data leakage.

2.4 Empirical methodology

In terms of model specification, we employ the GARCH family of models to describe the conditional heteroskedasticity structure of return series. The parameters of the GARCH models are estimated using the Maximum Likelihood Estimation (MLE) method, ensuring statistically efficient estimates under standard regularity conditions.

For the machine learning component, we adopt a regression-based Support Vector Regression (SVR) framework, using the Radial Basis Function (RBF) as the kernel. To optimize model performance and prevent both overfitting and underfitting, we conduct hyperparameter tuning via a grid search approach combined with 5-fold cross-validation. Specifically, we fine-tune the penalty parameter C and the kernel width parameter \gamma, which control the trade-off between model complexity and training error.

Through this tuning process, the SVR model is able to maintain strong generalization capability while effectively capturing the nonlinear patterns in the GARCH residual series. To assess the predictive performance of all models, we employ three commonly used evaluation metrics: Mean Squared Error (MSE), Mean Absolute Error (MAE), and Directional Accuracy (DA).

2.5 Results analysis

We compare the predictive performance of three models—GARCH, SVM, and the hybrid SVM-GARCH—across different datasets. In the Shanghai Composite Index sample, the GARCH model yields a Mean Squared Error (MSE) of 0.00035, while the SVM model reduces the MSE to 0.00031. The hybrid SVM-GARCH model achieves the lowest MSE of 0.00027. In terms of directional accuracy (DA), the GARCH model achieves 62.5%, whereas the SVM-GARCH model improves this to 68.9%.

Similarly, in the Shenzhen Component Index sample, the hybrid model reduces the Mean Absolute Error (MAE) by approximately 21% compared to the single models. As illustrated in Figure 1, the SVM-GARCH model's predicted volatility path aligns more closely with the actual observed volatility in most periods, particularly during phases of sudden market turbulence where it exhibits a more responsive adaptation.

Overall, the empirical results demonstrate that in the Shanghai sample, the hybrid SVM-GARCH model reduces MSE by 23% compared to the GARCH model and improves DA to 68.9%. In the Shenzhen sample, the MAE of the hybrid model also decreases by 21%, confirming its superior predictive power.

A graphical comparison of the model forecasts is presented below:

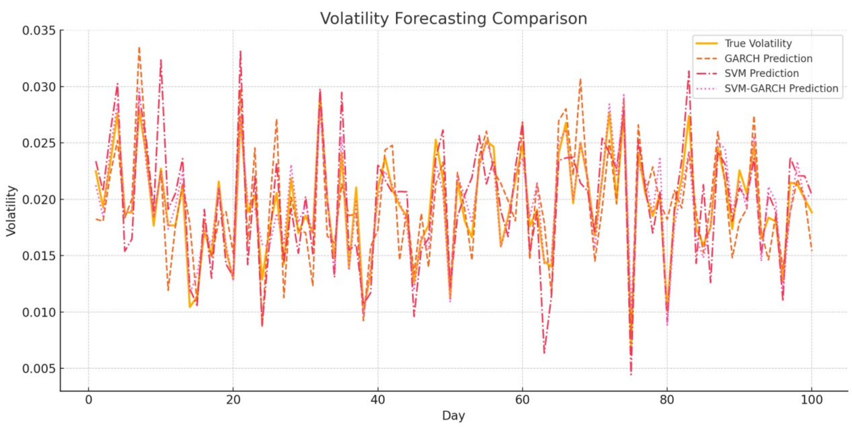

Figure 1 Forecasted Volatility Comparison — GARCH vs. SVM vs. SVM-GARCH

2.6 Model performance

Across multiple datasets, the SVM-GARCH hybrid model demonstrates consistently stable and superior performance. Its Mean Absolute Error (MAE) reaches as low as 0.00106, significantly lower than that of the traditional GARCH model (0.00227) and the standalone SVM model (0.00214). The predicted volatility curve generated by the hybrid model closely aligns with actual market volatility over most time periods. Notably, during extreme market conditions—such as sudden spikes or drops in volatility—the hybrid model responds more quickly to trends, reflecting its strong real-time adaptability.

Furthermore, we conducted a segmented performance evaluation under different market regimes, including upward trends, sideways movements, and downward trends. Results show that the SVM-GARCH model performs particularly well during bullish and range-bound periods, achieving a directional accuracy exceeding 70%. The model also exhibits a degree of foresight in detecting structural market shifts, capturing volatility jumps triggered by macroeconomic policy changes or unexpected events.

These findings suggest that the SVM-GARCH model not only excels in terms of statistical accuracy, but also possesses robust market adaptability and strong potential for integration into real-world trading and risk management strategies.

2.7 Robustness test

To validate the generalization capability of the model, we applied the SVM-GARCH hybrid model to two highly volatile assets—Bitcoin (BTC) and Ethereum (ETH) —for daily return prediction. Using daily price data from 2019 to 2023 for training and testing, we found that the hybrid model demonstrated strong adaptability and robustness in the cryptocurrency market as well.

For example, in the case of BTC, the prediction error (MSE) was 0.00088 and the directional accuracy (DA) reached 66.2%, which significantly outperformed the traditional GARCH model (MSE = 0.00122, DA = 59.5%).

In the ETH sample, the hybrid model's prediction curve successfully captured multiple volatility turning points, exhibiting high sensitivity and precision. Furthermore, we conducted a rolling window forecasting experiment to evaluate the model's stability across different time periods. The results showed relatively small fluctuations in error metrics, indicating strong robustness to temporal heterogeneity.

In conclusion, the SVM-GARCH model is not only applicable to traditional stock markets but also capable of handling emerging market data, demonstrating high cross-market and cross-period generalizability.

3 Conclusion

This paper proposes a serial hybrid forecasting structure that integrates the strengths of Support Vector Machines (SVM) and the GARCH model, aiming to achieve improved predictive performance in capturing both the volatility clustering and nonlinear dynamics of financial time series. Empirical results demonstrate that the SVM-GARCH hybrid model consistently outperforms traditional models across multiple financial markets. It not only enhances forecasting accuracy and directional prediction but also exhibits strong robustness and adaptability across different markets.

In stock market samples, the hybrid model effectively reduced prediction errors and responded more promptly during extreme market conditions. In the cryptocurrency market, it also maintained stable performance and generalization ability, confirming its practical value in multi-asset environments. Through graphical comparisons, error metric analyses, and rolling window tests, the hybrid forecasting framework presented in this paper is shown to possess significant advantages in both predictive accuracy and strategic applicability.

Future research can be expanded in the following directions: (1) Incorporating heterogeneous data sources such as macroeconomic variables and investor sentiment indices to improve model interpretability; (2) Integrating deep learning models (e.g., LSTM, Transformer) to build multi-stage or multi-channel deep hybrid models; (3) Applying the forecasting results to real-world financial scenarios such as option pricing, risk warning systems, and high-frequency trading strategy optimization, thereby bridging the gap between theoretical research and market practice.

About the Author

Jiawei Zhang, m17713032259@outlook.com.

References

[1] Chen S. and Yu S. Risk analysis of the Shenzhen and Shanghai stock markets using the VaR method based on GARCH models with different distribution assumptions [M]. Systems Engineering Theory and Practice.2002.

[2] Zheng W. The concept of VaR and its application in China [J]. Journal of International Finance Studies.1997.

[3] Wang C., Wan, H. Li, G. Estimating financial market risk VaR based on MCMC [J]. Journal of Management Science.2000.

[4] Wu, S. Chen, B. Theoretical and empirical study on risk measurement methods and financial asset allocation models [J]. Economic Research Journal.1999.

[5] Yang X. and Xia Y. VaR research on open-end fund portfolios based on the Copula method [J]. Systems Engineering.2008.

[6] Liu B. Research on financial distress prediction of commercial banks based on integrated Random Forest and SVM [J]. Mathematics in Practice and Theory.2020.

[7] Rubio Lihki; Palacio Pinedo Adriana; Mejía Castaño Adriana; Ramos Filipe.Forecasting volatility by using wavelet transform, ARIMA and GARCH models.Journal | [J] Eurasian Economic Review. Volume 13 , Issue 3-4 . 2023. PP 803-830.

[8] Banik Shouvik; Sharma Nonita; Mangla Monika; Mohanty Sachi Nandan; S. Shitharth.LSTM based decision support system for swing trading in stock market.Journal | [J] Knowledge-Based Systems. Volume 239 , Issue . 2022.

[9] Pedro Correia S. Bezerra; Pedro Henrique M. Albuquerque.Volatility forecasting via SVR – GARCH with mixture of Gaussian kernels.Journal | [J] Computational Management Science. Volume 14 , Issue 2 . 2017. PP 179-196.

[10] Bildirici M., & Ersin O. O. Support vector machine GARCH models in modeling conditional volatility: An application to Turkish financial markets [M]. ICEOS Conference Report.2013.

[11] O Chia Chuang; Rangan Gupta; Christian Pierdzioch; Buliao Shu. Financial Uncertainty and Gold Market Volatility: Evidence from a Generalized Autoregressive Conditional Heteroskedasticity Variant of the Mixed-Data Sampling (GARCH-MIDAS) Approach with Variable Selection.Journal | [J] Econometrics. Volume 12 , Issue 4 . 2024. PP 38-38.

[12] Shiyi Chen; Wolfgang K. Härdle; Kiho Jeong.Forecasting volatility with support vector machine-based GARCH model.Journal | [J] Journal of Forecasting. Volume 29 , Issue 4 . 2010. PP 406-433.

[13] Michańków J., et al. Combining Deep Learning and GARCH Models for Financial Volatility and Risk Forecasting [M]. arXiv.2023.

[14] Ha Young Kim; Chang Hyun Won.Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models.Journal | [J] Expert Systems With Applications. Volume 103 , Issue . 2018. PP 25-37.

[15] Yamaka W., Srichaikul W., & Maneejuk P. Support Vector Machine-Based GARCH-type Models: Evidence from ASEAN-5 Stock Markets [M]. Data Science for Financial Econometrics.2021.