

Forecasting Mastercard Stock Price Using Singular Spectrum Analysis with Bootstrap Prediction Intervals

M. Fariz Fadillah Mardianto^{1*}, Davin Firmansha¹, Dewi Rohayah¹

¹Department of Mathematics, Universitas Airlangga, Indonesia

m.fariz.fadillah.m@fst.unair.ac.id

ABSTRACT

Article History:

Received : 02-05-2026

Revised : 18-06-2026

Accepted : 20-06-2026

Online : 01-10-2026

Keywords:

Singular Spectrum

Analysis;

Bootstrap Prediction

Interval;

Stock Price Forecasting;

Financial Time Series;

Uncertainty

Quantification.



Stock price forecasting is essential for investment decision-making and financial risk management. However, the nonlinear and nonstationary characteristics of financial time series make accurate forecasting challenging. This quantitative study aims to forecast Mastercard Incorporated (MA) stock prices using Singular Spectrum Analysis (SSA) and bootstrap-based prediction intervals. Weekly closing price data from January 2017 to December 2025 were obtained from Investing.com. Automatic SSA was employed to automatically determine the optimal forecasting parameters using Mean Absolute Error (MAE) as the optimization criterion, while bootstrap resampling was used to estimate forecast uncertainty. The results indicate that the recurrent SSA approach outperformed vector SSA and achieved high forecasting accuracy with an out-of-sample MAPE of 3.60% and RMSE of 22.57. The optimal model was obtained with a window length of 104 and three reconstructed components, which explained 99.65% of the total variance. Furthermore, the bootstrap prediction intervals achieved a coverage probability of 98.94%, demonstrating reliable uncertainty quantification. The forecasting results suggest a continued upward trend in Mastercard stock prices over the forecast horizon, while the prediction intervals provide a realistic range of future price movements. These findings demonstrate that the integration of SSA and bootstrap prediction intervals offers an effective and interpretable framework for financial time series forecasting by combining automatic parameter optimization, accurate prediction, and uncertainty estimation.



<https://doi.org/10.31764/jtam.v10i4.39502>



This is an open access article under the [CC-BY-SA](#) license

A. INTRODUCTION

Stock price is one of the most important indicators in financial markets because it reflects the performance, financial condition, and future prospects of a company. Investors and financial analysts continuously monitor stock price movements to support investment decision-making, risk management, and portfolio optimization (Li et al., 2024). However, stock prices are known to be highly volatile and influenced by various internal and external factors such as market sentiment, economic conditions, monetary policy, and technological developments (Chun et al., 2025). As a result, stock price data often exhibit complex patterns, including trends, fluctuations, and dynamic changes over time, making accurate forecasting a challenging yet essential task in financial analysis (Feng et al., 2025).

One of the companies that attracts significant attention in the global financial market is Mastercard Incorporated, a leading payment technology company that operates an extensive electronic payment network worldwide (Chen et al., 2024). The company's stock, traded under the ticker symbol MA, is widely recognized for its large market capitalization, stable financial

performance, and high trading activity, making it an important asset in financial market studies (Sonkavde et al., 2023). These characteristics make Mastercard stock an attractive subject for financial research, particularly in studies related to stock price forecasting and risk analysis. Furthermore, the availability of comprehensive historical price data and the presence of identifiable trend and fluctuation patterns make Mastercard stock suitable for empirical time series modeling and forecasting studies (Zhang et al., 2024).

Nevertheless, forecasting financial time series remains difficult due to the nonlinearity, nonstationarity, and stochastic nature of stock price movements (Trinka et al., 2022). Traditional parametric models often require strict statistical assumptions, such as stationarity and predefined functional forms, which may not adequately capture the dynamic and complex behavior of financial data (Nguinabé et al., 2025; Oukhouya & El Himdi, 2023). Consequently, researchers have increasingly explored nonparametric and data-driven approaches that offer greater flexibility in modeling time series patterns without relying on restrictive assumptions (Feng et al., 2025).

One of the widely used nonparametric techniques for time series analysis is Singular Spectrum Analysis (SSA) (Golyandina, 2020). SSA is a decomposition-based method that transforms a time series into several interpretable components, including trend, periodic patterns, and noise, through embedding and matrix decomposition procedures (Yarmohammadi, 2024). The main advantage of SSA lies in its ability to analyze nonlinear and nonstationary data without requiring distributional assumptions. By separating signal components from noise, SSA can improve forecasting accuracy and reveal hidden structures in time series data (Radojičić et al., 2023). Therefore, SSA has become an increasingly popular method in financial forecasting applications, particularly for datasets characterized by volatility and structural complexity (Hassani & Mahmoudvand, 2018).

Several recent studies have demonstrated the effectiveness of Singular Spectrum Analysis (SSA) in forecasting financial and economic time series. A study by Zinenko (2023) applied SSA to financial time series data and showed that the method was capable of effectively analyzing and forecasting market indicators by decomposing the series into interpretable components such as trend and fluctuations, improving forecasting reliability. Furthermore, Han et al. (2022) proposed an improved SSA-based forecasting framework and demonstrated that the decomposition capability of SSA significantly enhanced prediction performance in time series forecasting applications, particularly for complex and nonlinear data. More recent research by Wang et al. (2023) developed hybrid SSA forecasting models and reported that SSA effectively separated trend, periodic components, and residuals in nonstationary time series, leading to more accurate forecasting results compared to conventional methods. Additionally, SSA has been applied for forecasting financial time series and was found to be effective in capturing underlying patterns and improving forecasting performance, highlighting its potential for accurate financial market prediction (Makshanov et al., 2022).

Despite the promising performance of SSA in financial forecasting, several limitations remain in the existing literature (Hassani & Mahmoudvand, 2018). Previous SSA-based studies have primarily focused on broad financial indicators, such as stock market indices, exchange rates, and cryptocurrencies, or highly volatile technology stocks, with limited attention given to mature payment technology companies characterized by relatively stable growth patterns and

strong market fundamentals (Zhang et al., 2024). In addition, most prior studies have emphasized point forecasting performance while paying less attention to prediction uncertainty, despite its importance in financial risk assessment and investment decision-making (Jiang et al., 2024; Sonkavde et al., 2023). Furthermore, many existing SSA applications rely on manually selected parameters based on user judgment, although such practices are not recommended because parameter choices can substantially influence forecasting performance (Hassani & Mahmoudvand, 2018). Therefore, this study employs Automatic SSA (ASSA) to determine optimal parameters automatically. In addition, this study adopts Mean Absolute Error (MAE) as the primary loss function for model evaluation, as MAE provides a more natural and less ambiguous measure of average forecasting error compared to RMSE, which can be disproportionately influenced by large errors (Willmott & Matsuura, 2005).

This study differs from previous SSA research by applying SSA and Bootstrap SSA to Mastercard Incorporated (MA), a global payment technology company with distinctive financial time series characteristics. Unlike highly speculative or extremely volatile assets, Mastercard stock exhibits relatively stable long-term trends, high trading liquidity, and recurring fluctuation patterns associated with macroeconomic conditions and digital payment activity, making it particularly suitable for SSA-based decomposition and signal extraction (Beyaztas & Shang, 2022; Rahmani, 2014). By integrating SSA with bootstrap prediction intervals, forecasting models can not only improve prediction accuracy but also quantify uncertainty surrounding future stock price movements, thereby providing more reliable information for risk management and investment decision-making (Rahmani, 2014; Lee, 2022). This study applies SSA to Mastercard Incorporated stock prices because the company exhibits relatively stable long-term trends, high trading activity, and identifiable fluctuation patterns, making its financial time series suitable for SSA-based decomposition and forecasting. Furthermore, the incorporation of bootstrap-based prediction intervals within the SSA framework enables the estimation of a plausible range of future stock prices without relying on strict distributional assumptions, thereby producing more comprehensive forecasting results for financial risk assessment.

B. METHODS

1. Data Description

This study utilizes weekly closing price data of Mastercard Incorporated (ticker: MA) obtained from Investing.com (Investing.com, 2024), spanning from January 2017 to December 2025. The 2017-2020 period captures diverse market conditions including pre-pandemic recovery, inflationary pressures, and subsequent bull market phases.

2. Singular Spectrum Analysis

Singular Spectrum Analysis (SSA) is a nonparametric time series decomposition technique that effectively separates time series into interpretable components including trends, cyclical patterns, and noise without requiring prior knowledge of the underlying data structure (An et al., 2024). The method combines concepts from principal component analysis (PCA), multivariate statistical analysis, and signal processing to provide a robust framework for time

series analysis and forecasting (Golyandina, 2020). SSA comprises four sequential stages: embedding, singular value decomposition (SVD), eigentriple grouping, and reconstruction.

a. Embedding Stage

The embedding stage transforms the univariate time series into a trajectory matrix (also called a Hankel matrix). Given a time series $\mathbf{y} = \{y_1, y_2, \dots, y_N\}$ of length N , the embedding process constructs an $L \times K$ trajectory matrix $\mathbf{X}$ where L is the window length (embedding dimension) and $K = N - L + 1$ is the number of lag vectors. The trajectory matrix is defined as:

$$\mathbf{X} = \begin{pmatrix} y_1 & y_2 & \cdots & y_K \\ y_2 & y_3 & \cdots & y_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & \cdots & y_N \end{pmatrix} \quad (1)$$

where the window length L must satisfy $1 < L < N$. The choice of window length is critical as it determines the maximum period of oscillations that can be captured; typically, $L \leq N/2$ is recommended to ensure sufficient data points for reliable decomposition (An et al., 2024). For financial time series with weekly observations, the window length should be selected to encompass relevant periodicities, such as seasonal or annual cycles.

b. Singular Value Decomposition (SVD)

The SVD stage decomposes the trajectory matrix $\mathbf{X}$ into its constituent components. The algorithm proceeds as follows:

Compute the lag-covariance matrix $\mathbf{S} = \mathbf{X}\mathbf{X}^\top$, an $L \times L$ matrix.

Obtain the eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L > 0$ and corresponding orthonormal eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_L$ of matrix $\mathbf{S}$.

compute the singular values as $\sigma_i = \sqrt{\lambda_i}$ for $i = 1, 2, \dots, L$.

Construct the principal components (right singular vectors) as $\mathbf{v}_i = \mathbf{X}^\top \mathbf{u}_i / \sigma_i$.

The SVD decomposition yields:

$$\mathbf{X} = \sum_{i=1}^L \sigma_i \mathbf{u}_i \mathbf{v}_i^\top \quad (2)$$

Each triple $(\sigma_i, \mathbf{u}_i, \mathbf{v}_i)$ is termed an eigentriple, where σ_i represents the singular value (relative importance), $\mathbf{u}_i$ is the left eigenvector, and $\mathbf{v}_i$ is the right eigenvector. The eigenvalue $\lambda_i = \sigma_i^2$ quantifies the variance explained by the i -th eigentriple, enabling ranking of components by their contribution to total variance (Nibbering & Paap, 2024).

c. Grouping Stage

The grouping stage partitions the eigentriples into m disjoint groups $I_1, I_2, \dots, I_m$ such that $I_1 \cup I_2 \cup \dots \cup I_m = \{1, 2, \dots, L\}$ and $I_i \cap I_j = \emptyset$ for $i \neq j$. This critical step determines how the time series is decomposed into interpretable components. Several automated grouping approaches exist. The W-correlation method measures the weighted

correlation between pairs of eigenvectors, grouping eigenvectors with high W-correlation together. The W-correlation between eigenvectors $\mathbf{u}_i$ and $\mathbf{u}_j$ is computed as:

$$\rho_w(i, j) = \frac{\langle \mathbf{W}(\mathbf{u}_i - \mathbf{u}_j), \mathbf{u}_i - \mathbf{u}_j \rangle}{\langle \mathbf{W}\mathbf{u}_i, \mathbf{u}_i \rangle + \langle \mathbf{W}\mathbf{u}_j, \mathbf{u}_j \rangle} \quad (3)$$

where $\mathbf{W}$ is a weight matrix typically chosen as $w_k = \min(k, L, N - k)$ for row k , reflecting the number of elements in diagonal averaging (discussed in Section 2.2.4) (An et al., 2024).

Alternatively, the periodogram-based selection method analyzes the frequency content of each eigenvector through spectral analysis. This approach identifies dominant frequencies and groups eigenvectors exhibiting similar spectral characteristics. For each eigenvector $\mathbf{v}_i$, the periodogram computes:

$$I(\omega) = \frac{1}{K} \left| \sum_{k=1}^K v_{i,k} e^{-i\omega k} \right|^2 \quad (4)$$

Eigenvectors with peaks at similar frequencies are grouped together as periodic components (P. Wang et al., 2024). Trend components are identified as those eigenvectors with eigenvalues that account for the largest proportion of total variance and exhibit slow variation. The grouping can be performed manually by visual inspection of eigenvector patterns, automatically using correlation or spectral methods, or through hybrid approaches combining multiple criteria.

d. Reconstruction Stage

The reconstruction stage recovers an approximation of the original time series by recombining selected eigentriples. For each group I_k with $k = 1, 2, \dots, m$, the partial matrix is computed as:

$$\mathbf{X}_{I_k} = \sum_{i \in I_k} \sigma_i \mathbf{u}_i \mathbf{v}_i^\top \quad (5)$$

Diagonal averaging (also called Hankelization) transforms each partial matrix $\mathbf{X}_{I_k}$ back into a time series. For a matrix element at position (i, j) , the contribution to the reconstructed series at index k is averaged over all diagonal positions where this element appears in the Hankel matrix. The reconstructed time series component corresponding to group I_k is:

$$\tilde{y}_n^{(k)} = \begin{cases} \frac{1}{n} \sum_{i=1}^n x_{i,n-i+1}^{(I_k)} & \text{if } 1 \leq n \leq L \\ \frac{1}{L} \sum_{i=1}^L x_{i,n-i+1}^{(I_k)} & \text{if } L < n \leq N - L + 1 \\ \frac{1}{N - n + 1} \sum_{i=n-N+L}^L x_{i,n-i+1}^{(I_k)} & \text{if } N - L + 2 \leq n \leq N \end{cases} \quad (6)$$

where $x_{i,j}^{(I_k)}$ denotes the element at position (i, j) in matrix $\mathbf{X}_{I_k}$ (Golyandina, 2020). The original time series is then reconstructed as:

$$\tilde{y}_n = \sum_{k=1}^m \tilde{y}_n^{(k)} \quad (7)$$

e. Forecasting with SSA

Two primary forecasting approaches exist in SSA: Recurrent Forecasting (R-forecast) and Vector Forecasting (V-forecast). Both methods utilize the eigentriples retained during reconstruction to extrapolate the time series beyond the observed period (An et al., 2024). Recurrent Forecasting (R-forecast) extends the reconstructed time series iteratively. The linear recurrent formula (LRF) establishes a linear relationship among the principal components. For forecasting M steps ahead, let the reconstructed series be $\tilde{y} = \{\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_N\}$. The coefficients $a_1, a_2, \dots, a_L$ are determined by solving a least-squares problem, typically using the eigenvectors retained in reconstruction. The R-forecast is computed as:

$$\hat{y}_{N+j} = \sum_{i=1}^r a_i \hat{y}_{N+j-i}, \quad j = 1, 2, \dots, M \quad (8)$$

where $r \leq L$ is the number of leading eigentriples used, and $\hat{y}_N, \hat{y}_{N-1}, \dots, \hat{y}_{N-r+1}$ are initialized from the reconstructed series.

Vector Forecasting (V-forecast) treats the trajectory matrix components directly. For each selected eigentriple, the future principal component values are extrapolated using PCA reconstruction principles. The V-forecast involves computing future projections of eigenvectors and aggregating them through diagonal averaging. V-forecast often produces more accurate predictions for financial time series exhibiting trend and cyclical patterns, as it preserves the multivariate structure of the trajectory matrix during forecasting (Golyandina, 2020).

f. SSA Bootstrap Prediction Interval

The bootstrap prediction interval construction begins with model fitting to the training data and residual computation. After fitting the model, the residuals are computed as:

$$e_t = y_t - \hat{y}_t \quad (9)$$

for every, $t = 1, 2, \dots, N$. Where y_t is the actual observation, $\hat{y}_t$ is the fitted value, and N is the number of training observations. These residuals form the basis for the bootstrap resampling process. For each bootstrap replication $b = 1, 2, \dots, B$, a bootstrap sample is generated by resampling residuals with replacement. The bootstrap sample is constructed by adding randomly selected residuals to the fitted values:

$$y_t^{(b)} = \hat{y}_t + e_t^{(b)}, t = 1, 2, \dots, N \quad (10)$$

where $e_t^{(b)}$ is a residual randomly drawn with replacement from the set $\{e_1, e_2, \dots, e_n\}$ on b replication and. This procedure preserves the temporal structure and correlation patterns in the original residuals while introducing the necessary randomness for inference. For each bootstrap sample, the model is refitted to obtain bootstrap parameter (Beyaztas & Shang, 2022). The bootstrap forecast for horizon h on b replication is then computed as $\hat{y}_{N+h}^{(b)}$.

After B bootstrap replications, the $(1 - \alpha) \times 100\%$ prediction interval is constructed using the empirical quantiles of the bootstrap forecast distribution below (Golyandina, 2020):

$$\hat{L}_{N+h}^{(b)} = Q_{\frac{\alpha}{2}}(\hat{y}_{N+h}^{(b)}) + Q_{\frac{\alpha}{2}}(e_t) \quad (11)$$

$$\hat{U}_{N+h}^{(b)} = Q_{1-\frac{\alpha}{2}}(\hat{y}_{N+h}^{(b)}) + Q_{1-\frac{\alpha}{2}}(e_t) \quad (12)$$

where $Q_{\alpha/2}$ and $Q_{1-\alpha/2}$ denote the $\alpha/2$ and $1 - \alpha/2$ quantiles of the bootstrap forecast distribution, respectively. The inclusion of this additional residual term is crucial for capturing the forecast uncertainty that arises from both model estimation error and future shocks (Beyaztas & Shang, 2022). This quantile-based approach is attractive because it adapts automatically to the shape of the forecast distribution without requiring symmetry assumptions. In this study, some bootstrap replications were used to construct 95% prediction intervals ($\alpha = 0.05$) as $\{100; 500; 1,000; 2,000; 5,000; 10,000\}$. This numbers of replications was chosen to see the computational efficiency and sufficient stability of the quantile estimates, following recommendations in the literature for practical forecasting applications (Beyaztas & Shang, 2022).

g. Automatic SSA Algorithms

SSA parameter selection in this research is likely based on the Automated SSA algorithms as in Hassani & Mahmoudvand (2018) document. The algorithms of this research method is as follows.

- 1) Consider a real-valued nonzero univariate time series.
- 2) Devide the time series into two part. Which in this research, the first part is training and validating, and the other part is testing. Note that the training on the full dataset (which 80:20 training testing) isn't the samewith the training on this Automatic SSA, otherly the testing data is the same.
- 3) Use the training data to construct the trajectory matrix $\mathbf{X}$.

- 4) Obtain the SVD of $\mathbf{X}$.
- 5) Evaluate all possible combinations of r ($1 \leq r \leq L - 1$) singular values (step by step) for the selected L and split the elementary matrices X_i ($i = 1, \dots, L$) into several groups and sum the matrices within each group.
- 6) Perform diagonal averaging to transform the matrix with the selected r singular values into a Hankel matrix.
- 7) Depending on the forecasting approach one wishes to use, select either the RSSA approach or the VSSA approach
- 8) Define a loss function $\mathcal{L}$
- 9) When forecasting a series Y_N h -step ahead, the forecast error is minimized by setting $\mathcal{L}$ to be minimum where the vector $\hat{y}_{K+h}$ contains the h -step ahead forecasts obtained using the RSSA or VSSA forecasting algorithm
- 10) Find the combination of L and r which minimized $\mathcal{L}$
- 11) Use the optimal L to decompose the series comprising of the training set, and then select r singular values for reconstructing time series, and use this newly reconstructed series for forecasting the remaining observations (in this research, the remaining observation is testing data).

In this study, the validation set used for loss calculation consists of 54 observations, while the forecasting horizon is extended to 10 steps ahead to evaluate long-term forecasting capability. This validation scheme follows an expanding window approach, in which the training window is progressively enlarged as the forecasting process moves forward. The structure is described in Table 1 below.

Table 1. Automatic SSA expanding window design

Iteration	Series for Automatic SSA training	Series for Validation
1	1-312	322
2	1-313	323
3	1-314	324
$\vdots$	$\vdots$	$\vdots$
54	1-364	374

Unlike the referenced document, which employs the Root Mean Square Error (RMSE) loss function, this study uses the Mean Absolute Error (MAE) loss function since findings indicate that MAE is a more natural measure of average error and (unlike RMSE) is unambiguous (Willmott & Matsuura, 2005). The MAE loss function is defined below.

$$\mathcal{L}(y_t - \hat{y}_t) = \frac{1}{N} \sum_{t=1}^N |y_t - \hat{y}_t| \quad (13)$$

3. Model Evaluation Criteria

Model performance was assessed using both point forecast accuracy and prediction interval quality metrics.

a. Point Forecast Accuracy

Two metrics were employed to evaluate point forecast accuracy.

1) Mean Absolute Percentage Error (MAPE):

$$\text{MAPE} = \frac{100\%}{N} \sum_{t=1}^N \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (14)$$

where y_t is the actual value, $\hat{y}_t$ is the forecasted value, and N is the number of observations. MAPE provides a scale-independent measure of forecast accuracy, with lower values indicating better performance (Beyaztas & Shang, 2022).

2) Root Mean Square Error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t)^2} \quad (15)$$

RMSE penalizes large errors more heavily than MAPE due to the squaring operation, making it sensitive to outliers. Lower RMSE values indicate superior forecast accuracy (An et al., 2024).

b. Prediction Interval Quality

Four complementary metrics were used to evaluate prediction interval performance at the $(1 - \alpha)$ confidence level:

1) Prediction Interval Coverage Probability (PICP)

$$\text{PICP} = \frac{1}{N} \sum_{t=1}^N c_t \quad (16)$$

where $c_t = 1$ if $y_t \in [\hat{L}_t, \hat{U}_t]$ and $c_t = 0$ otherwise, with $\hat{L}_t$ and $\hat{U}_t$ denoting the lower and upper bounds of the prediction interval. Ideally, PICP should be close to the nominal coverage probability $(1 - \alpha)$ (Beyaztas & Shang, 2022).

2) Mean Prediction Interval Width (MPIW)

$$\text{MPIW} = \frac{1}{N} \sum_{t=1}^N (\hat{U}_t - \hat{L}_t) \quad (17)$$

MPIW measures the average width of prediction intervals. Narrower intervals are preferred, provided adequate coverage is maintained (Nguinabé et al., 2025).

3) Winkler Score

$$WS = \frac{1}{N} \sum_{t=1}^N \left[(\hat{U}_t - \hat{L}_t) + \frac{2}{\alpha} (\hat{L}_t - y_t) \mathbb{1}_{y_t < \hat{L}_t} + \frac{2}{\alpha} (y_t - \hat{U}_t) \mathbb{1}_{y_t > \hat{U}_t} \right] \quad (18)$$

where $\mathbb{1}$ is the indicator function. The Winkler Score balances interval width and coverage by imposing penalties for observations falling outside the prediction interval. Lower Winkler Scores indicate better overall prediction interval performance (Kamlangdee, 2024).

4) Average Coverage Error (ACE)

$$ACE = PICP - (1 - \alpha) \quad (19)$$

ACE is used to evaluate the quality and reliability of prediction intervals by measuring the average deviation between the actual coverage probability and the target confidence level.

C. RESULT AND DISCUSSION

This section presents the empirical findings of a quantitative forecasting study using univariate time series data on Mastercard stock prices. The analysis includes data identification, model estimation, prediction results, and forecast evaluation to determine the superior SSA-based forecasting approach.

1. Data Identification

This study utilizes weekly Mastercard stock price data consisting of 470 observations. The dataset is split into train data (376 observations, 80%) and test data (94 observations, 20%). The time series plot is presented in Figure 1 below.



Figure 1. Weekly Mastercard Stock Price Series Plot

Figure 1 shows that the Mastercard stock price has a clear upward trend. In the pre-COVID-19 period (2017 to early 2020), the Mastercard stock price showed a stable positive trend from a value range of 100 to around 350. Then, during the COVID-19 period, the trend was clearly disrupted, causing the data to become more volatile and the trend to be less steep than the previous period. After COVID-19 subsided (2023 to the end of 2025), the stock value resumed

a strong positive trend, consistently climbing upward from 350 to almost 600 by the end of the period.

2. Singular Spectrum Analysis Result

The first step of the SSA analysis using Automated SSA parameters selection is obtain the parameters (r, L) using the Automatic SSA algorithm in section (2.7). The validation set consists of 54 observations, and the forecast horizon is extended to 10 steps ahead in order to assess the model's ability to perform long-term forecasting. The window lengths are selected from grid search of chosen candidate 104, and 156 based on the periodicity of the weekly data, allowing for a more optimal component extraction process and separability as Hassani & Mahmoudvand (2018) research that higher L makes component more separable. The optimal parameter selection is presented in Table 2 as follows.

Table 2. Automated SSA optimal parameters

Forecast Approach	Optimal L	Optimal r	$\mathcal{L}$
Recurrent	104	3	30.32891
Vector	104	3	20.1597

Based on the table, the optimal automated SSA parameters are obtained using the recurrent forecasting approach, with a window length of $L = 104$ and number of first grouping components $r = 3$ (means the group components are first 3 components $\{1,2,3\}$), which produce the minimum loss value. Following the selection of the optimal SSA parameters, the 3D surface plot is presented on Figure 2 to visualize the distribution of MAE values across all evaluated combinations of L and r , allowing the analysis to assess parameter sensitivity and confirm whether the selected combination provides the lowest forecasting error.

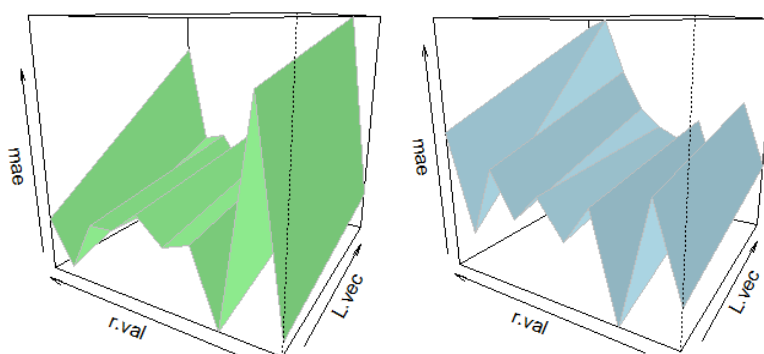


Figure 2. Sensitivity Analysis of MAE of forecast

The MAE surface plots illustrate how different combinations of parameters affect the forecasting performance of the RSSA and VSSA models, with r limited to a maximum of 10. The left plot shows that the MAE surface of RSSA fluctuates more strongly, indicating that its forecasting accuracy is more sensitive to changes in L and r . In contrast, the right plot shows that VSSA has a relatively little bit more structured MAE pattern, suggesting a more stable response to parameter variation. Therefore, the selected parameters correspond to the parameter combination that produces the lowest prediction error based on the MAE criterion.

Following the automatic selection of L and r based on the minimum MAE, Table 3 presents the top three selected eigenvalues that contribute most strongly to the reconstruction process and the final forecasting performance.

Table 3. Eigenvalues and Variance Contribution of SSA Components

Comp.	Singular Value	Eigen Value	Variance Explained	Cumulative Variance
1	50855.5883	2586290864	99.16%	99.16%
2	3251.6751	10573391.02	0.41%	99.56%
3	1548.463	2397737.716	0.09%	99.65%

Based on the Table 3, the selected eigentriples show that the reconstructed series is dominated by the first component, which captures the main structure of the stock price movement. The remaining components add minor fluctuations that help refine the reconstruction without introducing excessive noise. The W-correlation matrix was then examined to assess the separability of the reconstructed components and determine whether the selected components represent distinct signal structures, as shown in Figure 3.

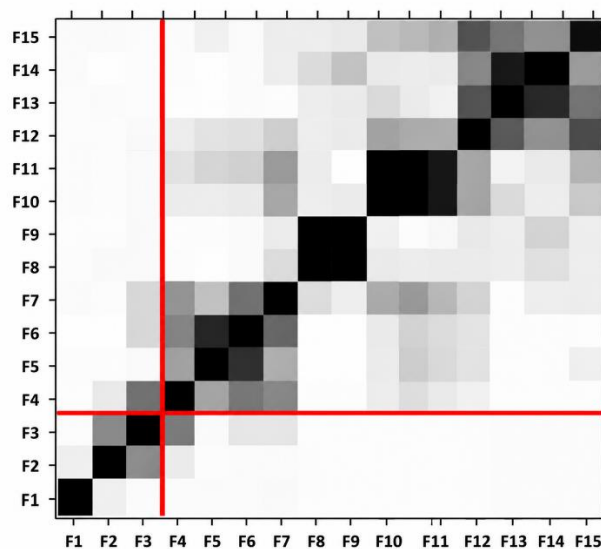


Figure 3. W-correlation matrix

The W-Correlation matrix visualization shows that the components are empirically separable. Component 1 is almost perfectly separated from the other components, characterized by an isolated deep black square and clean white cross-sections against all other components. Meanwhile, components 2 and 3 show minimal interaction with subsequent components, confirming that the trend signal and primary fluctuations in the Mastercard stock data have been extracted purely without being mixed with noise. The following is the best SSA evaluation matrix, as shown in Table 4.

Table 4. Best SSA Evaluation Matrix

Forecast	Signal Groups	Training MAPE	Testing MAPE	Training RMSE	Testing RMSE
RSSA	{1,2,3}	3.9923%	3.6034%	15.9912	22.5693
VSSA	{1,2,3}	3.9923%	5.7927%	15.9912	37.6882

Recurrent SSA forecast with selected best parameter produces minimum testing MAPE well below 10% indicating very high forecasting accuracy. The relatively small differences between the training and testing MAPE and RMSE values indicate that the model has good generalization ability and does not show signs of overfitting. Overall, the model demonstrates reliable performance with good generalization capability. Beyond point forecast accuracy, prediction intervals were constructed to quantify forecast uncertainty. Bootstrap method with some replications was employed to generate 95% SSA prediction intervals. Table 5 illustrates the SSA interval forecast evaluation matrix with its varying bootstrap replication below.

Table 5. Prediction Interval Evaluation Metrics

B	Covered	PICP	MPIW	Winkler Score	ACE	Computation Time
100	93/94	98.9362	127.2572	127.9708	3.9362	10.83 seconds
500	93/94	98.9362	127.0630	127.7167	3.9362	24.05 seconds
1000	93/94	98.9362	129.6876	129.9842	3.9362	42.56 seconds
2000	93/94	98.9362	127.4811	127.7519	3.9362	80.81 seconds
5000	93/94	98.9362	126.7203	127.1113	3.9362	176.40 seconds
10000	93/94	98.9362	126.5356	127.0689	3.9362	290.26 seconds

Based on the table, all scenarios of B result in a consistent PICP of 98.9362%, indicating that the prediction intervals provide very high coverage exceeding the specified confidence level. The stable ACE value of 3.9362 indicates consistent interval reliability in capturing actual data regardless of the number of bootstrap iterations used. While increasing the value of B significantly extends computation time, it leads to smaller and more stable MPIW and Winkler Scores, with the lowest values achieved at B=10000 (126.5356 and 127.0589, respectively). This suggests that increasing the bootstrap count produces narrower and more precise prediction intervals for Mastercard's stock price without sacrificing coverage performance.

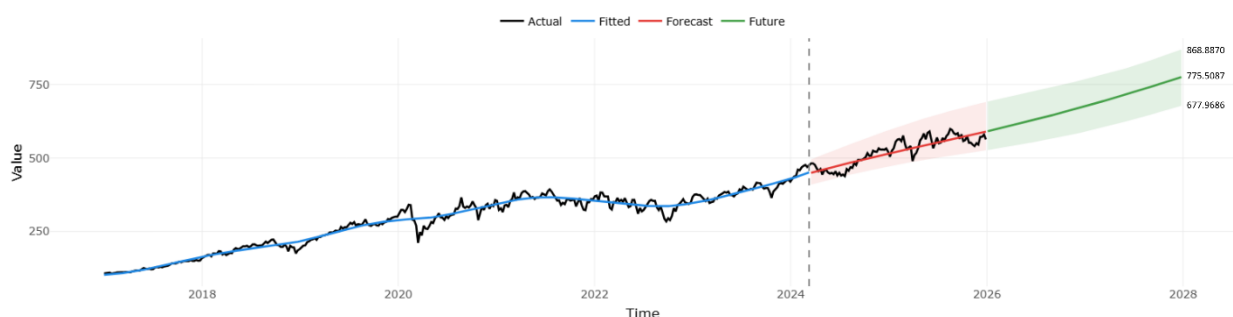
**Figure 4.** Best SSA Full Forecast with 95% Prediction Interval

Figure 4 displays the RSSA forecast with a 95% prediction interval using B=10,000, selected based on the previously determined optimal prediction evaluation matrix. The resulting forecast tends to show an upward trend through the end of 2027, with a prediction interval

width that remains relatively constant and does not exhibit excessive widening. Based on the plot, SSA reliably handles data affected by the COVID-19 intervention, as evidenced by the excellent and robust evaluation metrics and prediction intervals. The prediction interval indicates that Mastercard stock price is projected to increase over the forecast horizon. However, the widening interval shows that uncertainty becomes greater in longer-term forecasts. For investors, the central forecast can support hold or gradual accumulation decisions. Meanwhile, the lower and upper bounds can help assess downside risk and upside potential before determining investment allocation.

Compared to Zinenko (2023) findings, this study achieved a MAPE of 3.60%, outperforming SSA 17%, Fourier Series 3.80%, and RNN 86.55%, although it remained below ARIMA 1.40%. This comparison should be interpreted carefully because this study used 94 testing data points, which is 9.4 times longer than Zinenko (2023), which used only 10 data points. This study also adds prediction intervals to represent uncertainty and future data variation, which was not presented in the reference study. In addition, our findings that the use of MAE as the loss function can be considered reliable because the SSA model produced good forecasting performance, with an out-sample MAPE of 3.60%. Therefore, MAE can serve as a relevant alternative to RMSE as in Hassani & Mahmoudvand (2018), particularly for evaluating the average absolute prediction error of the model.

D. CONCLUSION

This study demonstrates that the integration of Automatic SSA and bootstrap-based prediction intervals provides a comprehensive framework for forecasting Mastercard stock prices. By automating the selection of SSA parameters, the proposed approach reduces subjectivity in model construction and improves the reproducibility of the forecasting process. The study further shows that a limited number of dominant SSA components can effectively capture the underlying trend and major fluctuations of Mastercard stock prices while separating noise from meaningful signals. Moreover, the incorporation of bootstrap prediction intervals extends the capability of conventional SSA by quantifying forecast uncertainty, thereby providing more informative insights for financial risk assessment and investment decision-making. The use of Mean Absolute Error (MAE) as the optimization criterion also demonstrates the feasibility of employing an interpretable and transparent loss function for automatic parameter selection. Overall, this research contributes to the advancement of SSA-based financial forecasting by combining automatic parameter optimization, uncertainty quantification, and long-term forecasting within a unified methodological framework.

ACKNOWLEDGEMENT

This work was independently conducted and self-funded by the author. Data for this research were retrieved from the Investing.com Indonesia database regarding Mastercard Inc. historical prices.

REFERENCES

- An, W., Gao, B., Liu, J., Ni, J., & Liu, J. (2024). Predicting hourly heating load in residential buildings using a hybrid SSA-CNN-SVM approach. *Case Studies in Thermal Engineering*, 59, 104516. <https://doi.org/10.1016/j.csite.2024.104516>
- Beyaztas, U., & Shang, H. L. (2022). Robust bootstrap prediction intervals for univariate and multivariate autoregressive time series models. *Journal of Applied Statistics*, 49(5), 1179–1202. <https://doi.org/10.1080/02664763.2020.1856351>
- Chen, W., Hussain, W., Cauteruccio, F., & Zhang, X. (2024). Deep Learning for Financial Time Series Prediction: A State-of-the-Art Review of Standalone and Hybrid Models. *Computer Modeling in Engineering & Sciences*, 139(1), 187–224. <https://doi.org/10.32604/cmescs.2023.031388>
- Chun, D., Cho, H., & Ryu, D. (2025). Volatility forecasting and volatility-timing strategies: A machine learning approach. *Research in International Business and Finance*, 75, 102723. <https://doi.org/10.1016/j.ribaf.2024.102723>
- Feng, L., Zheng, Y., Wang, X., Guo, C., & Xue, R. (2025). Global stock market forecasting: Insights from series and parallel combination of machine learning models. *Pacific-Basin Finance Journal*, 93, 102905. <https://doi.org/10.1016/j.pacfin.2025.102905>
- Golyandina, N. (2020). Particularities and commonalities of singular spectrum analysis as a method of time series analysis and signal processing. *WIREs Computational Statistics*, 12(4). <https://doi.org/10.1002/wics.1487>
- Han, M., Zhong, J., Sang, P., Liao, H., & Tan, A. (2022). A Combined Model Incorporating Improved SSA and LSTM Algorithms for Short-Term Load Forecasting. *Electronics*, 11(12), 1835. <https://doi.org/10.3390/electronics11121835>
- Hassani, H., & Mahmoudvand, R. (2018). *Singular Spectrum Analysis*. Palgrave Macmillan UK. <https://doi.org/10.1057/978-1-137-40951-5>
- Investing.com. (2024). *Mastercard Incorporated (MA) Historical Stock Price Data*. <https://www.investing.com/equities/mastercard-cl-a-historical-data>
- Jiang, Y., Liu, X., & Lu, Z. (2024). Financial uncertainty and stock market volatility. *European Financial Management*, 30(3), 1618–1667. <https://doi.org/10.1111/eufm.12456>
- Kamlangdee, P. (2024). Circular bootstrap on residuals for interval forecasting in time series. *Engineering and Applied Science Research*. <https://doi.org/10.14456/easr.2024.xx>
- Li, Z.-C., Xie, C., Wang, G.-J., Zhu, Y., Zeng, Z.-J., & Gong, J. (2024). Forecasting global stock market volatilities: A shrinkage heterogeneous autoregressive (HAR) model with a large cross-market predictor set. *International Review of Economics & Finance*, 93, 673–711. <https://doi.org/10.1016/j.iref.2024.05.008>
- Makshanov, A., Musaev, A., & Grigoriev, D. (2022). Analyzing and forecasting financial series with singular spectral analysis. *Dependence Modeling*, 10(1), 215–224. <https://doi.org/10.1515/demo-2022-0112>
- Nguinabé, J., & others. (2025). Bootstrap prediction intervals for the long short term memory network. *Expert Systems with Applications*, 267, 127728. <https://doi.org/10.1016/j.eswa.2025.127728>
- Nguinabé, J., Rockefeller, R., Khan, N. M., & Rodrigues, P. C. (2025). Bootstrap prediction intervals for the long short term memory (LSTM) recurrent neural network. *Expert Systems with Applications*, 284, 127728. <https://doi.org/10.1016/j.eswa.2025.127728>
- Nibbering, D., & Paap, R. (2024). Forecasting carbon emissions using asymmetric grouping. *Journal of Forecasting*, 43(6), 2228–2256. <https://doi.org/10.1002/for.3124>
- Oukhouya, H., & El Himdi, K. (2023). Comparing Machine Learning Methods—SVR, XGBoost, LSTM, and MLP— For Forecasting the Moroccan Stock Market. *IOCMA 2023*, 39. <https://doi.org/10.3390/IOCMA2023-14409>
- Radojičić, U., Nordhausen, K., & Taskinen, S. (2022). Singular Spectrum Analysis. In *Encyclopedia of Mathematical Geosciences* (pp. 1-5). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-85040-1_294
- Rahmani, D. (2014). A forecasting algorithm for Singular Spectrum Analysis based on bootstrap Linear Recurrent Formula coefficients. *International Journal of Energy and Statistics*, 02(04), 287–299. <https://doi.org/10.1142/S2335680414500197>

- Sonkavde, G., Dharrao, D. S., Bongale, A. M., Deokate, S. T., Doreswamy, D., & Bhat, S. K. (2023). Forecasting Stock Market Prices Using Machine Learning and Deep Learning Models: A Systematic Review, Performance Analysis and Discussion of Implications. *International Journal of Financial Studies*, 11(3), 94. <https://doi.org/10.3390/ijfs11030094>
- Trinka, J., Haghbin, H., & Maadooliat, M. (2022). Multivariate functional singular spectrum analysis: A nonparametric approach for analyzing multivariate functional time series. In *Innovations in multivariate statistical modeling: Navigating theoretical and multidisciplinary domains* (pp. 187-221). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-13971-0_9
- Wang, J., Peng, X., Wu, J., Ding, Y., Ali, B., Luo, Y., Hu, Y., & Zhang, K. (2023). Singular spectrum analysis (SSA) based hybrid models for emergency ambulance demand (EAD) time series forecasting. *IMA Journal of Management Mathematics*, 35(1), 45-64. <https://doi.org/10.1093/imaman/dpad019>
- Wang, P., Dong, J., Wang, L., & Qiao, S. (2024). Signal Denoising Method Based on EEMD and SSA Processing for MEMS Vector Hydrophones. *Micromachines*, 15(10), 1183. <https://doi.org/10.3390/mi15101183>
- Willmott, C., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30, 79-82. <https://doi.org/10.3354/cr030079>
- Yarmohammadi, M. (2024). Singular Spectrum Analysis. In *Flexible Nonparametric Curve Estimation* (pp. 175-196). Springer International Publishing. https://doi.org/10.1007/978-3-031-66501-1_8
- Zhang, C., Sjarif, N. N. A., & Ibrahim, R. (2024). Deep learning models for price forecasting of financial time series: A review of recent advancements: 2020-2022. *WIREs Data Mining and Knowledge Discovery*, 14(1). <https://doi.org/10.1002/widm.1519>
- Zinenko, A. V. (2023). Forecasting financial time series using singular spectrum analysis. *Business Informatics*, 17(3), 87-100. <https://doi.org/10.17323/2587-814X.2023.3.87.100>