

IMPLEMENTASI METODE *SINGULAR SPECTRUM ANALYSIS* (SSA) PADA PERAMALAN INDEKS LQ45

Teriska Deli^{1*}, Puspita Kartikasari², Iut Tri Utami³

^{1,2,3} Departemen Statistika, Fakultas Sains dan Matematika, Universitas Diponegoro

*e-mail: teriskad@students.undip.ac.id

DOI: 10.14710/j.gauss.13.1.210-218

Article Info:

Received: 2023-05-31

Accepted: 2024-10-18

Available Online: 2024-10-25

Keywords:

Forecasting; LQ45 Index; Singular Spectrum Analysis (SSA)

Abstract: Stock market index forecasting is one of the research studies to obtain stock market index prediction using statistical and computational techniques, one of the techniques that can be used is Singular Spectrum Analysis (SSA). SSA aims to decompose a time series into a set of independent components which have nonstatic tendencies such as trends, oscillations, and noise. SSA can be applied without any prior assumptions such as stationarity, linearity, and normality which seem impossible to fulfill by the fluctuation of stock market index data. One of the stock indices listed in Indonesia Stock Exchange (IDX) is LQ45. The LQ45 index contains 45 selected stocks that have great performance in liquidity and fundamental aspects. LQ45 Index data has a nonstatic structure so it is suitable to be predicted using the SSA method. The forecasting uses the daily data of the LQ45 Index for the period of July 25, 2022 – November 14, 2022 with a window length (L) of 30. From the forecasting process, 4 groups were acquired and Mean Absolute Percentage Error (MAPE) value was 1.69% to forecast the next 20 days. These results indicate that the forecasting process using SSA for LQ45 Index of the period is accurate.

1. PENDAHULUAN

Perkembangan pasar modal Indonesia menjadi bukti bahwa pasar modal merupakan peluang investasi yang baik selain perbankan. PT Kustodian Sentral Efek Indonesia (KSEI) mencatat jumlah investor di pasar modal Indonesia pada November 2022 mencapai 10 juta investor, meningkat 33,53% dari 7,5 juta pada akhir 2021. Komoditas yang dijual di pasar modal antara lain adalah saham. Saham dapat diartikan sebagai bukti penyertaan atau kepemilikan seseorang atau badan dalam perusahaan (Tandelilin, 2001).

Pergerakan harga saham tercermin dalam indeks saham. Salah satu indeks yang ada di Bursa Efek Indonesia (BEI) ialah Indeks LQ45. Indeks LQ45 merupakan indeks yang memiliki keunggulan dibandingkan dengan indeks saham lain pada aspek likuiditas di 45 emiten terseleksi melalui beberapa kriteria pemilihan. Harga saham cenderung dinamis, tidak linier, tidak stabil dan tidak mudah diprediksi. Harga saham dapat berfluktuasi dari waktu ke waktu bergantung pada penawaran dan permintaan pasar pada hari itu (Ardyana, 2020; Vjih et al., 2020). Ketidakpastian indeks saham ini perlu diramalkan oleh para investor agar tidak terjadi kesalahan dalam mengambil keputusan investasi.

Beberapa metode statistik dan analisis runtun waktu telah dikembangkan untuk peramalan harga saham, salah satunya adalah *Singular Spectrum Analysis* (SSA) (Leles et al., 2018). SSA dapat diterapkan tanpa menggunakan berbagai asumsi seperti stasioneritas dan normalitas data secara statistik (Syukur dan Marjuni, 2020). SSA menguraikan deret waktu menjadi sekumpulan komponen yang independen dan mudah diinterpretasikan, serta memiliki kecenderungan non-statis seperti tren, osilasi, dan *noise* (Leles et al., 2018; Xiao et al., 2019).

Penelitian ini akan membahas penerapan SSA untuk memprediksi atau meramalkan indeks LQ45 menggunakan program R. Penelitian ini diharapkan dapat mempermudah investor untuk menentukan keputusan investasi pada saham atau ETF yang menggunakan *benchmark* indeks LQ45 di masa depan.

2. TINJAUAN PUSTAKA

Indeks LQ45 adalah indeks yang mengukur performa harga saham dari 45 konstituen yang secara relatif memiliki kapitalisasi yang besar, likuiditas tinggi, dan fundamental yang baik. Likuiditas dilihat dari nilai transaksi, frekuensi transaksi, dan jumlah hari perdagangan di pasar reguler dan *free float* kapitalisasi pasar, sedangkan fundamental dilihat berdasarkan performa keuangan, ketaatan, dan sebagainya (Bursa Efek Indonesia, 2021).

Singular Spectrum Analysis (SSA) memiliki parameter yaitu *Window length* (L). Semakin besar nilai L , maka semakin rinci dekomposisi deret waktunya. Dekomposisi paling rinci ialah ketika nilai dari *window length* setengah dari panjang deret waktu (N), sehingga nilai L berada pada interval $2 < L < \frac{N}{2}$ (Golyandina et al., 2001). SSA terdiri dari dua tahap yaitu dekomposisi (*embedding* dan *singular value decomposition*) dan rekonstruksi (*grouping* dan *diagonal averaging*).

a. Embedding

Tahap ini mengubah data runtun waktu $F = F_N = f_1, \dots, f_N$ menjadi matriks lintasan (\mathbf{X}). Matriks \mathbf{X} berukuran $L \times K$, dengan L pada interval $2 < L < \frac{N}{2}$ dan $K = N - L + 1$ (Golyandina & Zhigljavsky, 2013).

$$\mathbf{X} = (x_{ij})_{i,j=1}^{L,K} = \begin{bmatrix} f_1 & f_2 & f_3 & \cdots & f_K \\ f_2 & f_3 & f_4 & \cdots & f_{K+1} \\ f_3 & f_4 & f_5 & \cdots & f_{K+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_L & f_{L+1} & f_{L+2} & \cdots & f_N \end{bmatrix} \quad (1)$$

b. Singular Value Decomposition (SVD)

Berdasarkan Golyandina et al. (2001), Nilai *singular* matriks lintasan \mathbf{X} didekomposisi merupakan faktor dari

$$\mathbf{X}_{i(L \times K)}^* = \sqrt{\lambda_i} \cdot \mathbf{u}_{i(L \times 1)} \cdot \mathbf{v}_{i(1 \times K)}^T \quad (2)$$

dengan $i = 1, 2, \dots, L$

Matriks \mathbf{X} diubah menjadi \mathbf{X}^* sebanyak L .

Tahap ini menghasilkan *eigen triple* $(\sqrt{\lambda_i}, \mathbf{u}_i, \mathbf{v}_i)$. Nilai singular dinotasikan $\sqrt{\lambda_i}$ dimana $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L > 0$ merupakan nilai eigen yang diperoleh dari perkalian matriks $\mathbf{S} = \mathbf{X}\mathbf{X}^T$. Vektor eigen dinotasikan $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_L$ dan komponen utama dihitung menggunakan persamaan berikut.

$$\mathbf{v}_{i(K \times 1)} = \frac{\mathbf{X}_{(K \times L)}^T \mathbf{u}_{i(L \times 1)}}{\sqrt{\lambda_i}} \quad (3)$$

c. Grouping

Pengelompokan dapat dilakukan dengan menganalisis plot tetapi sifatnya sangat subjektif (Khaeri et al., 2017). Golyandina & Zhigljavsky (2013) menambahkan bahwa terdapat solusi lain untuk mengatasi subjektivitas ini yaitu dengan dengan menghitung

korelasi terbobot dari setiap rata-rata diagonal matriks \mathbf{X}_i^* . Komponen yang berkorelasi tinggi terindikasi berada kelompok yang sama.

d. *Diagonal Averaging*

Pada tahap ini, matriks hasil dekomposisi yang telah dikelompokkan ($\mathbf{X}^* = \mathbf{X}_{I_1}^* + \dots + \mathbf{X}_{I_m}^*$) akan diubah menjadi deret baru dengan panjang N (Golyandina & Zhigljavsky, 2013). Matriks $\mathbf{X}_{I_m}^*$ adalah matriks $L \times K$ dengan elemen y_{ij} dengan $1 \leq i \leq L, 1 \leq j \leq K, L^* = \min(L, K), K^* = \max(L, K), N = L + K - 1, y_{ij}^* = y_{ij}$; untuk $L < K$, dan $y_{ij}^* = y_{ij}$; untuk lainnya.

$$g_k = \begin{cases} \frac{1}{k} \sum_{m=1}^k y_{m,k-m+1}^* & \text{untuk } 1 \leq k < L^* \\ \frac{1}{L^*} \sum_{m=1}^{L^*} y_{m,k-m+1}^* & \text{untuk } L^* \leq k \leq K^* \\ \frac{1}{N-k+1} \sum_{m=k-K^*+1}^{N-K^*+1} y_{m,k-m+1}^* & \text{untuk } K^* < k \leq N \end{cases} \quad (4)$$

Perubahan diagonal *averaging* matriks $\mathbf{X}_{I_m}^*$ menjadi deret baru yaitu $F_N^{(a)} = g_1, g_2, \dots, g_N$, dimana a adalah hasil pengelompokan dari tahap sebelumnya (Golyandina et al., 2001).

Dekomposisi SSA dikatakan berhasil ketika komponen aditif dari hasil pemisahan tidak memiliki korelasi. Persamaan (5) merupakan persamaan untuk mengukur korelasi antara komponen satu dengan yang lainnya (Hassani, 2007).

$$\rho_{ab}^{(w)} = \frac{(F_N^{(a)}, F_N^{(b)})_w}{\|F_N^{(a)}\|_w \|F_N^{(b)}\|_w} \quad (5)$$

$$\|F_N^{(a)}\|_w = \sqrt{(F_N^{(a)}, F_N^{(a)})_w}$$

$$(F_N^{(a)}, F_N^{(b)})_w = \sum_{i=1}^N w_i g_i^{(a)} g_i^{(b)}$$

$$w_k = \begin{cases} i & 1 \leq i < L^* \\ L^* & L^* \leq i \leq K^* \\ N - i + 1 & K^* < i \leq N \end{cases}$$

Linear Recurrent Formula (LRF) (r_1, r_2, \dots, r_d) dihitung menggunakan vektor eigen yang diperoleh dari tahap SVD. Dengan \mathbf{u}_i adalah vektor eigen yang bersesuaian, \mathbf{u}_i^∇ adalah $L - 1$ koordinat pertama dari \mathbf{u}_i , π_i adalah koordinat terakhir dari \mathbf{u}_i , dan $v^2 = \sum_{i=1}^{L-1} \pi_i^2$ maka koefisien LRF dapat dihitung menggunakan persamaan berikut.

$$\mathbf{r} = (r_{L-1}, \dots, r_1)^T = \frac{1}{1 - v^2} \sum_{i=1}^{L-1} \pi_i \mathbf{u}_i^\nabla \quad (6)$$

Deret baru hasil rekonstruksi yang telah diperoleh dari langkah *diagonal averaging* digunakan untuk melakukan peramalan sebanyak M titik baru sehingga akan tersusun deret hasil peramalan berdasarkan persamaan berikut.

$$\hat{g}_i = \begin{cases} g_i & \text{untuk } i = 1, \dots, N \\ \sum_{j=1}^{L-1} r_j g_{i-j} & \text{untuk } i = N + 1, \dots, N + M \end{cases} \quad (7)$$

dimana g_{N+1}, \dots, g_{N+M} adalah hasil ramalan dari SSA.

Tingkat keakuratan hasil ramalan dapat dijelaskan dengan membandingkan nilai hasil ramalan dengan nilai aktual menggunakan persamaan berikut (Hutasuhut et al., 2014).

$$MAPE = \left(\frac{1}{n} \sum_{i=1}^n \left| \frac{f_i - \hat{g}_i}{f_i} \right| \right) \times 100\% \quad (8)$$

Menurut Anggrainingsih et al. (2015), nilai MAPE dikategorikan ke dalam 4 bagian yaitu pada Tabel 1.

Tabel 1. Kategori Nilai MAPE

Nilai MAPE	Akurasi Peramalan
$MAPE \leq 10\%$	Tinggi
$10\% < MAPE \leq 20\%$	Baik
$20\% < MAPE \leq 50\%$	Reasonable
$MAPE > 50\%$	Rendah

3. METODE PENELITIAN

Penelitian ini menggunakan data yang bersumber dari laman resmi Bursa Efek Indonesia (<https://www.idx.co.id/id/data-pasar/ringkasan-perdagangan/ringkasan-indeks>) sebanyak 80 data harian nilai indeks LQ45 pada periode 25 Juli 2022 – 14 November 2022. Data *in sample* terdiri dari 60 data (25 Juli 2022 – 18 Oktober 2022) dan data *out sample* sebanyak 20 data (19 Oktober 2022 – 14 November 2022).

Perangkat lunak yang mendukung penelitian ini yaitu Microsoft Excel dan RStudio. Berikut tahapan analisis secara urut.

1. Mengamati plot data
2. Memulai analisis dengan tahap dekomposisi yaitu *embedding* dan SVD
3. Saat proses *embedding*, data dikonversi menjadi matriks \mathbf{X} berukuran $L \times K$ dimana L merupakan pendekatan dari setengah jumlah data dengan nilai $K = N - L + 1$
4. Pada langkah SVD, dihasilkan *eigen triple* yaitu diantaranya *singular value* ($\sqrt{\lambda_i}$), *eigen vector* (\mathbf{u}_i), dan komponen utama (\mathbf{v}_i) sebanyak L . Ketika *eigen triple* telah diperoleh, maka akan terbentuk matriks baru \mathbf{X}_i^* menggunakan persamaan $\mathbf{X}_i^* = \sqrt{\lambda_i} \mathbf{u}_i \mathbf{v}_i^T$
5. Melakukan proses rekonstruksi yang terdiri dari *grouping* dan *diagonal averaging*
6. Komponen *noise* dipisahkan berdasarkan *weighted correlation* dari vektor eigen pada tahap *grouping*
7. Matriks \mathbf{X}_i^* yang diperoleh pada tahap SVD dikelompokkan berdasarkan kategori kelompok yang telah ditentukan saat tahap *grouping*
8. Memperoleh deret baru dari matriks \mathbf{X}_i^* pada proses *diagonal averaging*
9. Melakukan penaksiran koefisien *Linear Recurrent Formula*
10. Peramalan data *out sample* dan mengukur nilai MAPE
11. Melakukan peramalan pada seluruh data

4. HASIL DAN PEMBAHASAN

a. Embedding

Tahap ini membentuk matriks lintasan berdimensi $L \times K$, dengan L ditentukan menggunakan pedekatan $L \sim \frac{N}{2}$. Jika panjang $N = 60$, maka nilai L adalah $L = 30$ sehingga $K = 60 - 30 + 1 = 31$.

$$\mathbf{X}_{30 \times 31} = \begin{bmatrix} 967,131 & 968,482 & \vdots & 1027,013 \\ 968,482 & 974,462 & \vdots & 1019,666 \\ \vdots & \vdots & \ddots & \vdots \\ 1029,296 & 1027,013 & \dots & 972,673 \end{bmatrix}$$

b. Singular Value Decomposition (SVD)

1) Eigen value dan singular value

Tabel 2. Nilai Eigen dan Nilai Singular

Nilai Eigen				
966706525,05	72284,90	6825,85	...	6,68
Nilai Singular				
31091,90	268,86	82,62	...	2,59

2) Vektor eigen

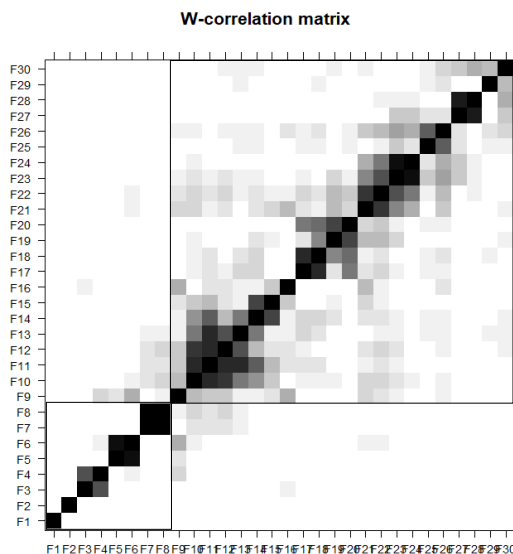
$$\mathbf{U}_{30 \times 30} = \begin{bmatrix} -0,18042 & -0,27586 & \vdots & 0,43551 \\ -0,18072 & -0,24208 & \vdots & -0,21014 \\ \vdots & \vdots & \ddots & \vdots \\ -0,18164 & 0,40095 & \dots & 0,20076 \end{bmatrix}$$

3) Komponen utama

$$\mathbf{V}_{31 \times 30} = \begin{bmatrix} -0,17737 & 0,27265 & \vdots & -0,19810 \\ -0,17772 & 0,25121 & \vdots & 0,50736 \\ \vdots & \vdots & \ddots & \vdots \\ -0,17860 & -0,39619 & \dots & 0,12864 \end{bmatrix}$$

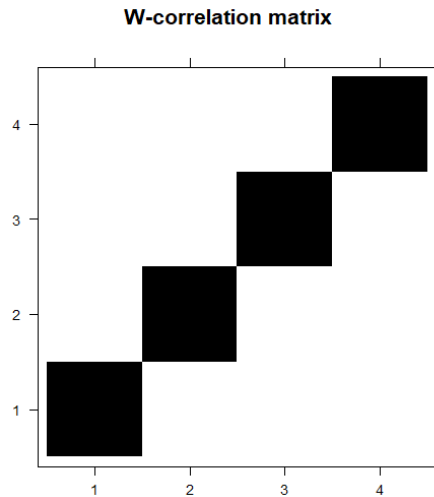
c. Grouping

Pada tahap ini dilakukan perhitungan *weighted correlation* pada vektor eigen. Berikut plot matriks *weighted correlation* yang terlihat pada Gambar 2.



Gambar 2. Matriks *W-Correlation*

Golyandina & Zhigljavsky (2013) mengungkapkan bahwa matriks *w-correlation* dapat membantu menentukan jumlah komponen yang akan diidentifikasi. Pada Gambar 2, matriks *w-correlation* dibagi menjadi dua dengan kotak besar yang gemerlap mengindikasikan komponen *noise* sehingga komponen yang akan digunakan untuk proses *grouping* yaitu 8 komponen pertama. Dengan bantuan program R, diperoleh *output* pengelompokan yaitu (1, 2, (3, 4, 7, 8), dan (5,6))



Gambar 3. Matriks *W-Correlation* Hasil Pengelompokan

Berdasarkan Gambar 3, tidak terdapat korelasi pada masing-masing kelompok satu terhadap yang lainnya dibuktikan dengan tidak adanya bayangan warna lain di setiap kombinasi kelompok.

d. Diagonal Averaging

Langkah *diagonal averaging* memperoleh deret baru yaitu dapat dilihat di Tabel 4.

Tabel 3. Hasil *Diagonal Averaging*

Data ke-	Kel. 1	Kel. 2	Kel 3.	Kel. 4
1	994,9293	-20,2223	-7,6181	-0,7907
2	996,7486	-18,1888	-6,3156	-0,6309
3	998,5470	-16,0392	-6,4263	-0,4723
4	1000,3445	-14,4279	-8,4872	-0,5939
5	1002,0984	-13,2506	-6,9334	-0,7528
⋮	⋮	⋮	⋮	⋮
60	1008,6693	-42,7087	0,8387	5,2226

Berdasarkan hasil rekonstruksi diperoleh MAPE *in sample* sebesar 0,26%. Nilai MAPE yang diperoleh termasuk ke dalam kategori yang memiliki akurasi tinggi berdasarkan Tabel 1.

Tahap peramalan data *out sample* terlebih dahulu menghitung koefisien LRF yaitu terlihat di Tabel 4.

Tabel 4. Koefisien LRF

LRF	Kel. 1	Kel. 2	Kel. 3	Kel. 4
r_{29}	0,03417	0,17229	0,02033	0,04902
r_{28}	0,03423	0,14373	0,01396	-0,00600
r_{27}	0,03427	0,12286	0,02535	-0,07920
r_{26}	0,03431	0,10360	0,02690	-0,11372
r_{25}	0,03434	0,08229	-0,00531	-0,08407
⋮	⋮	⋮	⋮	⋮
r_1	0,03389	-0,13179	-0,04052	-0,02116

Hasil peramalan data *out sample* diperoleh dengan hasil pada Tabel 5.

Tabel 5. Data Ramalan dan Data Aktual

Data ke-	Kel. 1	Kel. 2	Kel. 3.	Kel. 4	Data Aktual	Data Ramalan
61	1014,7328	-31,5363	1,7279	3,5203	971,350	988,445
62	1014,5063	-31,7979	0,7175	-0,3583	974,072	983,068
63	1014,2643	-31,8484	-0,2887	-4,0874	993,682	978,040
64	1013,9919	-31,6128	0,0142	-5,4592	1000,357	976,934
65	1013,6912	-31,0871	0,1882	-3,6295	1006,502	979,163
⋮	⋮	⋮	⋮	⋮	⋮	⋮
80	1007,4704	27,1564	0,7334	-6,3470	1001,020	1029,013

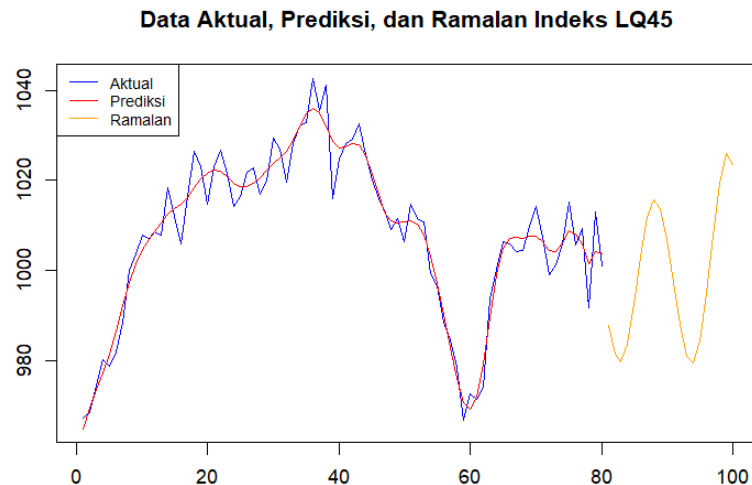
Dari hasil ramalan yang telah diperoleh, diketahui nilai MAPE sebesar 1,69%. Berdasarkan kategori keakuratan pada Tabel 1, MAPE yang diperoleh termasuk pada kategori dengan akurasi yang tinggi.

Peramalan akhir menghasilkan ramalan sebanyak 20 periode ke depan dengan hasil peramalan dapat dilihat di Tabel 6.

Tabel 6. Hasil Peramalan

Data ke-	Tanggal	Peramalan	Data ke-	Tanggal	Peramalan
81	15-11-22	987,855	91	29-11-22	997,399
82	16-11-22	981,463	92	30-11-22	987,937
83	17-11-22	979,627	93	01-12-22	981,018
84	18-11-22	983,604	94	02-12-22	979,390
85	21-11-22	992,423	95	05-12-22	984,439
86	22-11-22	1002,919	96	06-12-22	994,929
87	23-11-22	1011,701	97	07-12-22	1007,797
88	24-11-22	1015,667	98	08-12-22	1019,424
89	25-11-22	1013,684	99	09-12-22	1025,927
90	28-11-22	1006,798	100	12-12-22	1023,473

Berdasarkan hasil peramalan, indeks LQ45 dalam 20 hari ke depan (15 November 2022 – 12 Desember 2022) mengalami kenaikan dan penurunan dengan prediksi indeks LQ45 tertinggi ialah pada tanggal 12 Desember 2022 yaitu sebesar 1023,473 sedangkan prediksi indeks LQ45 terendah ialah pada 1 Desember 2022 yaitu sebesar 981,018.



Gambar 4. Plot Data Aktual, Prediksi, dan Ramalan

Gambar 10 merupakan plot data aktual, prediksi, dan hasil ramalan. Terlihat bahwa plot data prediksi mengikuti plot data aktual dengan baik sehingga peramalan indeks LQ45 untuk periode 25 Juli 2022 – 14 November 2022 menggunakan metode SSA sudah sesuai.

5. KESIMPULAN

Proses peramalan indeks LQ45 dengan mengimplementasikan metode *Singular Spectrum Analysis* (SSA) menggunakan *window length* (L) sebesar 30 serta pembentukan empat kelompok menghasilkan nilai MAPE sebesar 1,69%. Berdasarkan ketentuan kategori MAPE, hasil peramalan tersebut masuk ke dalam kategori akurasi peramalan tinggi. Berdasarkan hasil tersebut, data indeks LQ45 periode 25 Juli 2022 – 14 November 2022 cocok untuk diramalkan menggunakan metode SSA.

DAFTAR PUSTAKA

- Anggrainingsih, R., Aprianto, G. R., & Sihwi, S. W. 2015. *Time Series Forecasting Using Exponential Smoothing To Predict The Number Of Website Visitor Of Sebelas Maret University*. ICITACEE, 14–19.
- Ardyana, I. M. 2020. *Manajemen Investasi dan Portofolio*. Lembaga Penerbitan Universitas Nasional.
- [BEI] Bursa Efek Indonesia. 2021. *IDX Stock Index Handbook v1.2*. Jakarta: BEI.
- Golyandina, N., Nekrutkin, V., & Zhigljavsky, A. 2001. *Analysis of Time Series Structure: SSA and Related Techniques*. United States of America: Chapman and Hall/CRC.
- Golyandina, N., & Zhigljavsky, A. 2013. *Singular Spectrum Analysis for Time Series*. Berlin: Springer.
- Hassani, H. 2007. *Singular Spectrum Analysis: Methodology and Comparison*. In *Journal of Data Science*, 5, 239–257.
- Husnan, S. 1996. *Manajemen Keuangan Teori Dan Penerapannya*. Yogyakarta: BPFE
- Hutasuhut, A. H., Anggraeni, W., & Tyasnurita, R. 2014. *Pembuatan Aplikasi Pendukung Keputusan untuk Peramalan Persediaan BahanBaku Produksi Plastik Blowing dan Inject Menggunakan Metode ARIMA(Autoregressive Integrated Moving Average) di CV. Asia*. *Jurnal Teknik Pomits*, 1-6.

- Leles, M., Mozelli, L., Nascimento Jr, C., Sbruzzi, E., *et al.* 2018. *Study on Singular Spectrum Analysis as a New Technical Oscillator for Trading Rules Design*. *Fluctuation and Noise Letters*, 17, 1–21.
- Tandelilin, E. 2001. *Analisis Investasi dan Manajemen Risiko (Pertama)*. Yogyakarta: BPFE.
- Vijh, M., Chandola, D., Tikkiwal, V. A., & Kumar, A. 2020. *Stock Closing Price Prediction using Machine Learning Techniques*. *Procedia Computer Science*, 167.
- Xiao, J. *et al.* 2018. *A New Approach for Stock Price Analysis and Prediction Based on SSA and SVM*. *International Journal of Information Technology & Decision Making*, 18, 287–310.