International Journal of Statistics and Applied Mathematics 2023; SP-8(4): 586-591

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2023; SP-8(4): 586-591

https://www.mathsjournal.com
Received: 17-05-2023
Accepted: 21-06-2023

Prabhat Kumar
[1] The Graduate School, ICAR-Indian Agricultural Research Institute, New Delhi, India
[2] ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India

Girish Kumar Jha
ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India

Rajeev Ranjan Kumar
ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India

Achal Lama
ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India

Corresponding Author:
Prabhat Kumar
[1] The Graduate School, ICAR-Indian Agricultural Research Institute, New Delhi, India
[2] ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India

Forecasting agricultural commodity prices using singular spectrum analysis

Prabhat Kumar, Girish Kumar Jha, Rajeev Ranjan Kumar and Achal Lama

Abstract
In recent times, the effectiveness of appropriate time series decomposition techniques has gained large acceptance. Among various time series decomposition techniques, singular spectrum analysis (SSA) is a highly promising technique. Its successful application has been demonstrated across various contexts, showcasing its efficacy in effectively separating and understanding different components within time series data. Thus, in this study, we have used SSA and its forecasting method SSA-LRF for modelling and forecasting agricultural price series, namely tomato of two markets, i.e., Delhi and Lucknow. Further, the results obtained from SSA-LRF are compared with that of the SSA-ARIMA and ARIMA models. The comparative analysis was carried out using RMSE, MAPE, and MAE criteria. We report the superiority of the SSA-LRF model over others under consideration in terms of the lowest RMSE, MAPE, and MAE values. This study has highlighted the importance of decomposition-based forecasting techniques such as SSA-LRF for agriculture price series.

Keywords: Agricultural price, decomposition, price forecasting, SSA, ARIMA

1. Introduction

Precise and reliable forecasts of agricultural commodities price are crucial for achieving a balanced supply and demand, ensuring fair remuneration for farmers and affordable prices for consumers (Sundaramoorthy *et al.*, 2014) [15]. However, forecasting vegetable prices faces significant challenges due to factors like seasonality, perishability, volatility, and non-linear characteristics. Various models have been developed to capture the complexity of price series (Xiong *et al.*, 2018) [16]. Neural networks were found ineffective at handling seasonal and trend variations without data pre-processing (Zhang and Qi, 2005) [9]. The widely recognized and commonly utilized approach for forecasting seasonal time series is the seasonal auto-regressive integrated moving average (SARIMA) model (Box *et al.*, 2015) [3]. Luo *et al.* (2013) [11] employed the SARIMA model to investigate a precise forecasting model for cucumber prices, which considers the seasonal effect. They selected $(1,0,1)(1,1,1)_{12}$ SARIMA model as it proved to be the most accurate option for short-term vegetable price prediction.

An analysis technique for observed time series may be particularly significant if it is able to expose important characteristics of the time series predictability. One such technique is known as Singular Spectrum Analysis (SSA). SSA decomposes a time series into interpretable components like trends, oscillations, and noise without strict distributional and structural assumptions (Golyandina and Zhigljavsky, 2013) [7]. SSA offers various applications, including trend identification, smoothing, seasonality extraction, and forecasting (Hassani *et al.*, 2007) [8].

The flexibility of SSA makes it suitable for non-linear and non-stationary time series analyses, as it does not rely on predefined functions like Fourier or wavelet approaches (Beneki *et al.*, 2012) [2]. SSA has shown superior performance compared to SARIMA and other state space models for signal extraction and forecasting in different domains (Yu *et al.*, 2017) [18]. Additionally, the SSA-Linear Recurrent Formulae (SSA-LRF) model has been employed successfully in hydrological forecasting and other practical applications (Zhang *et al.*, 2011) [20].

Furthermore, SSA has demonstrated its ability to accurately extract and forecast components from various time series, including annual precipitation, monthly runoff, and hourly water temperature (Marques *et al*., 2006) [12]. Combining SSA with complex seasonality methods, machine learning, and auto-regressive models have proven effective in predicting Brazil's monthly corn, soybean, and sugar spot prices (Palazzi, 2021) [14]. Overall, this study highlights the significance of SSA in addressing the challenges posed by vegetable price forecasting and emphasizes its potential to offer accurate predictions for such complex and non-linear time series.

The subsequent sections of this paper are organized as follows: Section 2 provides comprehensive insights into the various forecasting models utilized in this study. Section 3 presents the results and discussion. Ultimately, the paper concludes with Section 4.

2. Methodology

2.1 Singular spectrum analysis (SSA)

SSA is a non-parametric time-series analysis technique based on multivariate statistical concepts. It decomposes a given time series into a set of independent additive time series, each of which can be classified as trend, periodic or quasi-periodic component and residuals. For the implementation of SSA, four steps as follows.

First step: Embedding

The embedding procedure transfers a time series to a sequence of multidimensional lagged vectors. The original data $X = (X_1, X_N)$ is shifted to a trajectory matrix $Y = (Y_1, Y_L)$ with L dimensions, where $2 \leq L \leq N-1$. Each element of the matrix Y is defined as $Y_i = (x_i, \ldots, x_{i+L-1})$ and the matrix Y is given as follows

$$Y = \begin{pmatrix} X_1 & X_2 & \cdots & X_K \\ \vdots & & \ddots & \vdots \\ X_L & X_{L+1} & \cdots & X_N \end{pmatrix} \quad (1)$$

Where $K = N - L + 1$

Second step: Singular value decomposition (SVD): In this step, the matrix Y can be decomposed into d components, where $d = rank\ (Y)$. Through SVD, the Eigen triples (λ_i, U_i, V_i) of the matrix YY^T can be obtained in descending order by $\lambda_i (\lambda_1 \geq \cdots \geq \lambda_L \geq 0)$. U_i denotes the left eigenvector and V_i denotes the right eigenvector. Therefore, the matrix $\boldsymbol{Y}$ can be further rewritten as follows:

$$Y = Y_1 + Y_2 +. + Y_d$$

$$Y_i = \sqrt{\lambda_i} U_i V_i^T \quad (2)$$

Third step: Grouping

In this step, r out of d components are selected as the trend and seasonality components. Define $I = \{I_1,., I_r\}$ and $Y_I = Y_{I1} + Y_{I2} +. + Y_{Ir}$, then Y_I can represent the trend and seasonality component of the data, while the other $(d - r)$ components are regarded as the residuals. Where $I_1,., I_r$ is called the Eigen triple grouping.

Fourth step: Diagonal averaging

In this step, through the Hankelization procedure, the obtained group $\{Y_{I1}, Y_{I2}, \ldots, Y_{Ir}\}$ are shifted to the time series group $\{X_{I1}, X_{I2}, \ldots, X_{Ir}\}$. Then the original time series can be defined as follows:

$$X = X_{trend} + X_{seasonality} + \ldots, + X_{residuls} \quad (3)$$

2.2 Singular Spectrum Analysis-Linear Recurrent Formulae (SSA-LRF) forecasting approach

In SSA forecast, the model is expressed by a Linear Recurrent Formulae (LRF). This LRF applied to the last L -1 terms of the initial time series gives a continuation of it. The same idea can be applied to a component of the time series. Let $v^2 = \pi_1^2 + \pi_2^2,. + \pi_r^2$, where π_i is the last component of the eigenvector $U_i (i = 1,2,., r < L)$. Moreover, suppose for any vector $U \in R^L$ denoted by $U_i^{\nabla} \in R^{L-1}$ the vector consisting of the first $L - 1$ components of the vector U and r denotes the number of eigenvalues used for reconstruction, we can define the coefficient vector

$$\hat{\alpha} = (\hat{\alpha}_{L-1},., \hat{\alpha}_1)' = \frac{1}{1-v^2} \sum_{i=1}^{r} \pi_i U_i^{\nabla} \quad (4)$$

Considering the above notation, the h-step-ahead out-of-sample recurrent SSA forecasts $(\hat{X}_{N+1},., \hat{X}_{N+h})$ can be obtained as:

$$\hat{x}_i = \begin{Bmatrix} \tilde{x}_i \text{ for } n = 1,2,., N \\ \sum_{i=1}^{L-1} \hat{\alpha}_i \hat{x}_{n-i} \text{ for } n = N+1,., N+h \end{Bmatrix} \quad (5)$$

Where $\hat{x}_1,., \hat{x}_N$, are the reconstructed time series as obtained from Equation 3 and $\hat{x}_N,., \hat{x}_{N+h}$, are the estimates for the out-of-sample SSA-R forecasts.

2.3 Autoregressive Integrated Moving Average (ARIMA) model

ARIMA models represent a widely used, straightforward approach for modelling time series data. The fundamental principle behind this modelling strategy is based on the assumption of linear correlation among the values of a time series variable. In this method, a univariate time series is expressed as a function of its own past or lagged values, along with random disturbances (Jaiswal *et al*., 2022) [9]. The ARIMA (p, d, q) model, where p and q represent the orders of autoregressive (AR) and moving average (MA) components, respectively, and d denotes the order of differencing (integration), can be expressed as follows:

$$\varphi(B)\Delta^d x_t = \theta(B)u_t \quad (6)$$

x_t represents the value of the price series at time t, while u_t is the disturbance term at time t, assumed to be randomly and identically distributed with a mean of zero and a constant variance of σ^2. The backshift operator, denoted by B, is defined as $Bx_t = x_{t-1}$. The differencing operator Δ is given by, $\Delta = (1 - B)$, which calculates the difference between consecutive values in the series. The polynomials $\varphi(B)$ and $\theta(B)$ are both functions of B with degrees p and q, respectively.

2.4 Evaluation criteria

To compare the forecasting performance of the models, three indicators including root mean square error (RMSE), mean absolute deviation (MAE) and mean absolute percentage error (MAPE) have been used.

Root mean squared error (RMSE):

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{n}(x_t-\hat{x}_t)^2}{n}} \quad (7)$$

Mean Absolute Percentage Error (MAPE):

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{x_t-\hat{x}_t}{x_t}\right| \times 100 \quad (8)$$

Mean absolute error (MAE)

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}|x_t-\hat{x}_t| \quad (9)$$

Where x_t and $\hat{x}_t$ are the actual and predicted values, respectively, and n is the size of the testing set.

2.5 Diebold-Mariano (DM) test

In addition to the various accuracy indicators, the Diebold-Mariano (DM) test (Diebold and Mariano, 2002) [4] is used to determine and compare the prediction accuracy among competing models. The null hypothesis of the test is that two predictions have the same accuracy and therefore, the loss differential, $(t) = f(e1_{(t)}) - f(e2_{(t)})$; $t = 1, 2, \ldots, n$ has zero expectation, where $(e1_{(t)})$ and $(e2_{(t)})$ are the forecast error series obtained from the two models, and $f(.)$ is a loss function. The DM test statistic is:

$z_{DM} = \frac{\bar{d}}{\sqrt{\hat{V}_{\bar{d}}/n}} \sim N(0,1)$, where n is the size of the predictions, $\bar{d} = \frac{1}{n}\sum_{t=1}^{n} d_t$, $\hat{V}(\bar{d}) = \frac{1}{n}\left[\gamma_0 + 2\sum_{k=1}^{h-1}\hat{\gamma}_k\right]$, $\hat{\gamma}_k = \frac{1}{n}\left[\sum_{t=k+1}^{n}(d_t - \bar{d})(d_{t-k} - \bar{d})\right]$ are the sample mean of the loss differential, estimate of variance of the mean using h-step forecasts and estimate of k^{th} autocovariance of d_t respectively.

3 Results and Discussion

3.1 Data description

This investigation is done on the monthly wholesale price (₹/quintal) of the tomato crops from the two major wholesale markets of India, namely Delhi and Lucknow markets. The price series of the two markets are obtained from the Directorate of Marketing and Inspection (DMI), Ministry of agriculture and farmer welfare, Government of India. (https://agmarknet.gov.in/) from July 2010 to June 2023. Table 1 presents the descriptive statistics of the two series used in the study, and Fig. 1 shows the time plots of the two series. The visualization of the time plots indicates the non-stationarity and nonlinearity of all two series, however, a statistical test for stationarity and linearity are also performed to confirm this (Tables 2 and 3).

Table 1: Descriptive statistics of tomato price (₹/quintal) of two major markets of India

Statistics	Delhi	Lucknow
Mean	1414.90	1630.80
Maximum	3895.40	4274.50
Minimum	225.20	403.60
Standard deviation	754.14	828.54
Skewness	1.03	0.93
Kurtosis	3.64	3.31
Jarque- Bera (p −value)	30.72(< 0.001)	23.26(< 0.001)

Table 2: Augmented–Dickey–Fuller (ADF) test results of tomato price

Price series	ADF Test		Conclusion
	t-statistics	p −value	
Delhi	-1.78	0.07	Non-stationary
Lucknow	-1.83	0.06	

Table 3: Brock-Dechert-Scheinkman (BDS) test results for tomato price

Price Series	Epsilon	Embedding dimension				Conclusion
		2		3		
		Statistics	p − value	Statistics	p −value	
Delhi	0.5σ	14.40	< 0.001	16.64	< 0.001	Non-linear
	1.0 σ	11.10	< 0.001	11.70	< 0.001	
	1.5σ	8.34	< 0.001	7.97	< 0.001	
	2.0 σ	5.27	< 0.001	4.78	< 0.001	
Lucknow	0.5σ	21.62	< 0.001	12.30	< 0.001	Non-linear
	1.0 σ	18.16	< 0.001	9.72	< 0.001	
	1.5σ	24.62	< 0.001	12.13	< 0.001	
	2.0 σ	18.94	< 0.001	9.01	< 0.001	

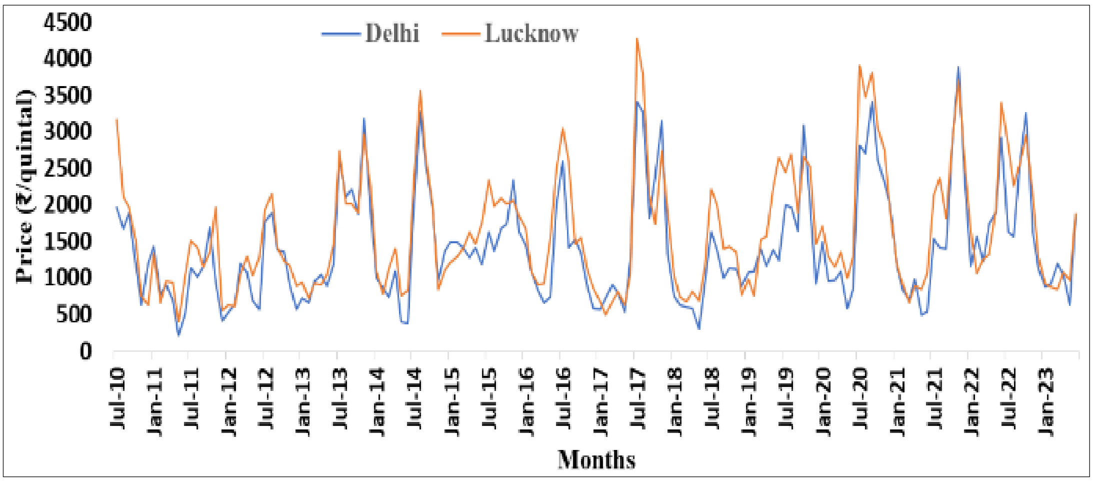


Fig 1: Time plot for monthly tomato price (₹/quintal) of Delhi and Lucknow market

The price series for the markets in Delhi and Lucknow varies from ₹ 225.20/quintal to ₹ 3895.40/quintal, and ₹ 403.60/quintal to ₹ 4274.50/quintal, respectively. Using the standard deviation as a crude measure of volatility, it was found that the price series for the Delhi and Lucknow markets are volatile, with standard deviations of 754.14 and 828.54, respectively (Table 1). The price series of two markets are leptokurtic and positively skewed, indicating that they are not normal, and the Jarque-Bera test further supports this (Jarque and Bera, 1980) [10]. Each price series have 156 observations, which are split 144 data points into training and 12 points data into testing sets respectively. While the testing set is preserved for post-sample prediction, the training set is used for model building and in-sample prediction. All the model building and statistical analysis were conducted in this study using R statistical software, version 4.1.1.

3.2 Choice of parameters and separability

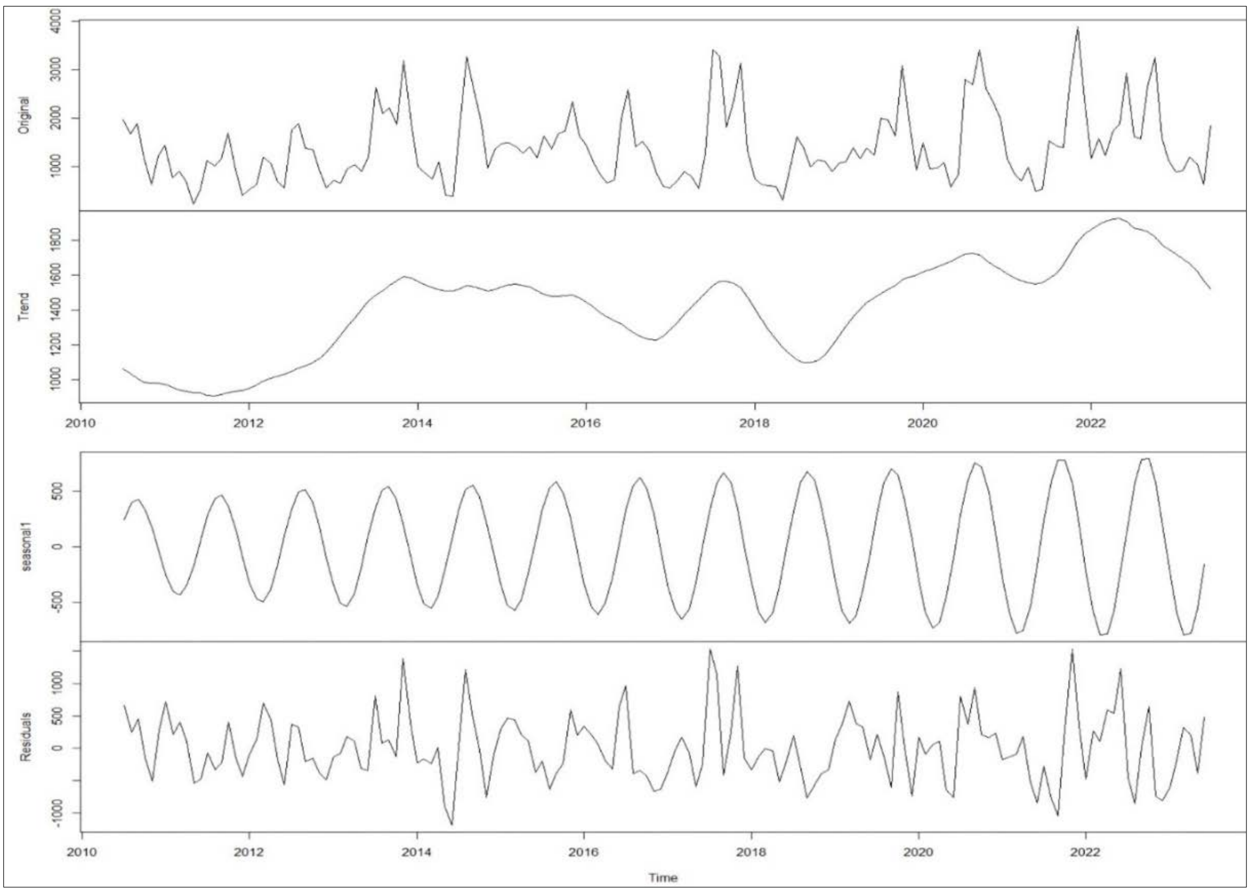
SSA is a powerful method for decomposing intricate data into a few independent and easily understandable components, such as trends, oscillations, and residuals. However, its success depends heavily on two crucial parameters: the window length (L) and the number of eigentriples (r). In particular, the appropriate selection of L is crucial to construct the trajectory matrix of the time series data accurately. If L is too small, closely spaced frequencies may not be distinguishable, while an excessively large L value can reduce the statistical significance of detected periodicities. Many researchers recommend choosing a reasonably large window length for time series decomposition to increase decomposition detail with longer windows. However, the window length should be smaller than half the series length. In this investigation, the focus is on analyzing two-price series, which exhibit non-stationary, non-linear, and complex characteristics. Due to the intricate nature of these series, attempting to fully decompose them in a single step is difficult. Initially, our attention was directed towards extracting the trend component. Given the dynamic nature of the trend, this extraction process bears a resemblance to smoothing. The parameter r in SSA determines the number of eigentriples retained for reconstructing the original time series. Selection of r is important as it affects the level of smoothing and noise reduction in the reconstructed time series. If r is too small, significant information may be lost, resulting in an under-smoothed and noisy reconstruction. Conversely, an overly large r can lead to an over-smoothed reconstruction, causing the loss of important details and features in the time series. The parameter r decides the numbers of trend, periodic and residuals components. After successfully extracting the trend, the remaining data constituted the residuals from the first step. The subsequent step involved extracting the seasonality component from these residuals. To achieve better separability, we determined the maximal window length L = 72, ensuring that $L \leq N/2$ (N = 156) and L is divisible by 12 (Golyandina *et al*., 2001). The selection of seasonal components relied on the weighted correlation matrix, incorporating various sine waves with distinct periods, contingent upon the nature of the price series. To accurately identify the desired sine waves, we employed scatterplots of eigenvectors and the w-correlation matrix of the elementary components. In this study, the initial step encompassed the selection of trend components from two markets. In the second step, both series have one seasonal component and the remaining components in both markets were treated as residuals.

Table 4: Forecasting performance of different models of (a) Delhi and (b) Lucknow market

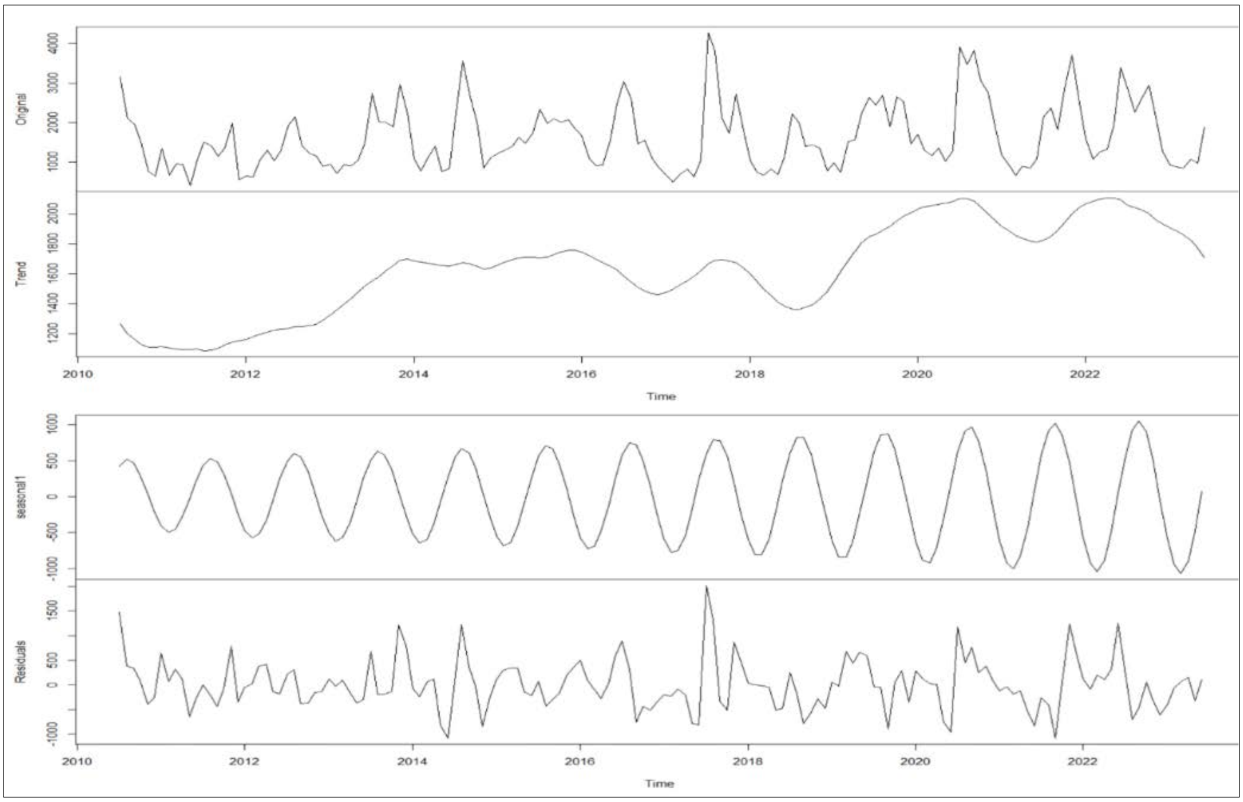
Models	RMSE	MAPE (%)	MAE
ARIMA	908.37	63.30	845.82
SSA-ARIMA	688.01	49.56	598.44
SSA-LRF	473.41	27.95	365.76

(b)

Models	RMSE	MAPE (%)	MAE
ARIMA	1845.11	144.64	1673.02
SSA-ARIMA	757.17	55.51	669.97
SSA-LRF	467.03	27.72	391.15



(a)



(b)

Fig 2: SSA decomposition plots of tomato price series for (a) Delhi, and (b) Lucknow markets

Table 5: Results of Diebold-Mariano (DM) test of tomato price series of Delhi market

Hypothesis	DM value	p −value	Remarks
H_0: The accuracy of both SSA-LRF and SSA- ARIMA is the same. H_1: The accuracy of SSA-LRF is superior to SSA-ARIMA.	-1.88	0.043	The accuracy of SSA-LRF is superior to SSA-ARIMA.
H_0: The accuracy of both SSA-LRF and ARIMA is the same. H_1: The accuracy of SSA-LRF is superior to ARIMA.	-2.67	0.010	The accuracy of SSA-LRF is superior to ARIMA.
H_0: The accuracy of both SSA-ARIMA and ARIMA is the same. H_1: The accuracy of SSA-ARIMA is superior to ARIMA.	-1.90	0.041	The accuracy of SSA-ARIMA is superior to ARIMA.

Table 6: Results of Diebold-Mariano (DM) test of tomato price series of Lucknow market

Hypothesis	DM value	p −value	Remarks
H_0: The accuracy of both SSA-LRF and SSA- ARIMA is the same. H_1: The accuracy of SSA-LRF is superior to SSA-ARIMA.	-2.32	0.019	The accuracy of SSA-LRF is superior to SSA-ARIMA.
H_0: The accuracy of both SSA-LRF and ARIMA is the same. H_1: The accuracy of SSA-LRF is superior to ARIMA.	-4.28	<0.001	The accuracy of SSA-LRF is superior to ARIMA.
H_0: The accuracy of both SSA-ARIMA and ARIMA is the same. H_1: The accuracy of SSA-ARIMA is superior to ARIMA.	-4.23	<0.001	The accuracy of SSA-ARIMA is superior to ARIMA.

3.3 Model Implementation and Evaluation

This study presents the forecasting experiments conducted on the tomato price series using the research design described earlier. Three accuracy measures and the Diebold-Mariano test were used to evaluate the predictive performance of different models. Figure 2 displays the outcomes of applying the Singular Spectrum Analysis (SSA) technique to the decomposition of the two market tomato price series. Each original tomato price series was decomposed into trend, seasonals, and residual components using SSA. Notably, the seasonal component for all tomato price series exhibits a consistent periodicity with a 12-month cycle. Subsequently, the extracted trend, seasonals, and residual components were forecasted using the SSA-LRF method, as discussed in Section 2.2. The forecasts for these components were then aggregated to generate the final output. The forecasting results for 12-month forecasts of RMSE, MAPE, and MAE for three models (SSA-LRF, SSA-ARIMA, and ARIMA) are presented in Table 4.

The superior performance of SSA-LRF and SSA-ARIMA over ARIMA is attributed to their ability to directly capture nonlinearity, high volatility, complexity, and seasonality present in the data after the decomposition step. In contrast, ARIMA could not directly handle these factors, making data pre-processing, such as decomposition via SSA, crucial in developing a more accurate forecaster, as demonstrated in the proposed SSA-LRF approach. To assess the statistical significance of the mentioned models, the DM statistic was employed. The results of the DM test (Tables 5 and 6) provided evidence that SSA-LRF was the best forecasting model among the ones considered in the study.

4. Conclusion

The results of this study demonstrate the effectiveness of the SSA-LRF technique as a forecasting algorithm in the case of tomato price series. The application of SSA successfully decomposed the tomato price series into trend, seasonal, and residual components. SSA-LRF and SSA-ARIMA forecasting models outperformed the traditional ARIMA model due to their ability to handle nonlinearity, complexity, and seasonality after decomposition. Among these models, SSA-LRF consistently showed superior predictive performance in terms of RMSE, MAPE, and MAE for forecasting tomato prices.

Furthermore, a statistical analysis using the DM test confirmed SSA-LRF as the most effective forecasting model among the considered model. These findings highlight the importance of proper data preprocessing, such as SSA-based decomposition, in achieving more accurate predictions for tomato prices over the next 12 months. Therefore, the study suggests that SSA can be a valuable addition to time series analysis and forecasting methodologies for agricultural price data, improving the accuracy of predictions and aiding decision-making in the agricultural sector.

5. Acknowledgements: The first author is grateful to PG School, ICAR-Indian Agricultural Research Institute, New Delhi for providing the requisite facilities to carry out this study.

6. Availability of data and materials: The paper uses public domain data.

7. Competing interests: The authors have no potential conflicts to report that are important to the article's content.

8. References

1. Alharbi N, Hassani H. A new approach for selecting the number of the eigenvalues in singular spectrum analysis. Journal of the Franklin Institute. 2016;353(1):1-16.
2. Beneki C, Eeckels B, Leon C. Signal extraction and forecasting of the UK tourism income time series: A singular spectrum analysis approach. Journal of Forecasting. 2012;31(5):391-400.
3. Box GEP, Reinsel GC, Jenkins M, Ljung GM. Time series analysis: forecasting and control. Wiley, Hoboken, 5th edn. Wiley, Hoboken; c2015.
4. Diebold FX, Mariano RS. Comparing Predictive Accuracy. Journal of Business and economic statistics; 2002;20:134-144.
5. Goh C, Law R. Modeling and forecasting tourism demand for arrivals with stochastic non-stationary seasonality and intervention. Tourism management. 2002;23(5):499-510.
6. Golyandina N, Nekrutkin V, Zhigljavsky AA. Analysis of time series structure: SSA and related techniques. CRC press; c2001.
7. Golyandina N, Korobeynikov A. Basic singular spectrum analysis and forecasting with R. Computational Statistics and Data Analysis. 2013;71:934-954.
8. Hassani H. Singular Spectrum Analysis: Methodology and Comparison. Journal of Data Science. 2007;5(2):239-257.
9. Jaiswal R, Jha GK, Choudhary K, Kumar RR. STL decomposition-based LSTM model for seasonal agricultural price forecasting. Preprint; c2022. https://doi.org/10.21203/rs.3.rs-1350423/v1
10. Jarque CM, Bera AK. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. Economics Letters. 1980;6:255-259.
11. Luo CS, Zhou LY, Wei QF. Application of SARIMA modeling in cucumber price forecast. Applied Mechanics and Materials; c2013. p. 373-375, 1686-1690.
12. Marques CAF, Ferreira JA, Rocha A, Castanheira JM, Melo-Gonçalves P, Vaz N, *et al*. Singular spectrum analysis and forecasting of hydrological time series. Physics and Chemistry of the Earth. 2006;31(18):1172-1179.
13. Mahmoudvand R, Najari N, Zokaei M. On the Optimal Parameters for Reconstructionand Forecasting in Singular Spectrum Analysis, Communications in Statistics-Simulation and Computation. 2013;42:860-870.
14. Palazzi RB, Maçaira P, Meira E, Klotzle M. Forecasting Commodity Prices in Brazil Through Hybrid SSA-Complex Seasonality Models (No 6652). Easy Chair; c2021.
15. Sundaramoorthy C, Jha GK, Suresh P, Mathur VC. Market integration and volatility in edible oil sector in India. Journal of the Indian Society of Agricultural Statistics. 2014;68(1):67-76.
16. Xiong T, Li C, Bao Y. Seasonal forecasting of agricultural commodity price using a hybrid STL and ELM method: Evidence from the vegetable market in China. Neurocomputing. 2018;275:2831-2844.
17. Wang R, Ma HG, Liu GQ, Zuo DG. Selection of window length for singular spectrum analysis. Journal of the Franklin Institute. 2015;352(4):1541-1560.
18. Yu C, Li Y, Zhang M. An improved wavelet transforms using singular spectrum analysis for wind speed forecasting based on Elman neural network. Energy Conversion and Management. 2017;148:895-904.
19. Zhang GP, Qi M. Neural network forecasting for seasonal and trend time series. European Journal of Operational Research. 2005;160(2):501-514.
20. Zhang Q, Wang BD, He B, Peng Y, Ren ML. Singular spectrum analysis and ARIMA hybrid model for annual runoff forecasting. Water resources management. 2011;25:2683-2703.