J. Soils and Crops 35 (2) 328-331, December 2025 ISSN 0971-2836 (Print)
ISSN 2582-2756 (Online)

A COMPARATIVE STUDY OF TIME SERIES MODELS FOR SOYBEAN PRICE FORECASTING

Ragini Chavan[1] and Omprakash Jadhav[2]

ABSTRACT

Soybean is a crucial agricultural commodity, and its price fluctuations significantly impact farmers, traders, and policymakers. This study applied time series models—ARIMA, SARIMA, Holt-Winters (additive and multiplicative) and GARCH to analyze monthly soybean price trends. The dataset, consisting of monthly soybean prices from 2010 to 2020 was divided into training and testing sets to evaluate model performance. The models were assessed based on key error metrics, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). Among the tested models, HW-Multiplicative found superior in predicting, outperforming ARIMA, SARIMA and GARCH by yielding the lowest error values. The findings suggest that HW-Multiplicative was the most suitable model for forecasting soybean prices, providing valuable insights for market participants and aiding in decision-making related to price stabilization and agricultural planning. These findings can directly contribute to better price stability, informed decision-making, and improved market efficiency in the soybean sector. Farmers can use it for optimal selling and planting decisions, while traders and wholesalers can plan inventory and procurement strategies. Policymakers can implement price stabilization measures, and agribusinesses can optimize raw material planning. Overall, HW-Multiplicative helps to improve market efficiency to reduce risks and to support informed decision-making across the soybean supply chain. Given the seasonal fluctuations in soybean prices, policymakers should implement strategies such as buffer stock mechanisms or price stabilization policies to mitigate extreme price variations

(Key words : Soybean, price prediction, correlation, time series models)

INTRODUCTION

Agriculture is the backbone of the Indian economy. It is the main source of income in India. This field is bound by the culture and heritage of India. Around 70% of the land is under cultivation. Hence India is the second-largest country in cultivable land (Barakade *et al*., 2011). Agriculture planning plays an important role in economic growth, food security, and Gross Domestic Product (GDP). Farmers are the drivers of the agriculture sector. Nowadays, the consumption of soybean has increased due to their health benefits. As soybean has a good number of proteins, this is considered an alternative source of meat for vegetarians (Merga and Haji, 2019).Price fluctuations in soybean markets are influenced by various factors, including supply-demand dynamics, weather conditions, global trade policies, and government interventions (Das *et al*., 2020).Maharashtra is one of the leading states in soybean production, so this project will empower farmers to choose soybeans as their preferred crop.The success of any Agribusiness mostly depends on judicious prices for the products. Thus, it is essential to study pricing principles, methods in the sector, and their impact on the efficiency of agribusinesses (Chavan and Jadhav, 2024). However, the current lethargic pricing process at agribusinesses argues for the lack of effective methodological and theoretical support for decision-making in the field of pricing policy which often leads to incorrect price calculations and results in significant losses.Thus, it is obligatory to study pricing principles, methods, and features in the sector and their impact on the efficiency of agribusinesses (Agbaeza *et al*., 2020).

The usage of time series analysis techniques in agriculture price prediction has garnered Consideration due to their potential to capture the innate chronological patterns and oscillations in commodity prices. Various statistical models have been applied to forecast agricultural prices, including the Autoregressive Integrated Moving Average (ARIMA), Seasonal ARIMA (SARIMA), and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models (Enders, 2014). An autoregressive integrated moving average model is a proven time series forecasting method that deals with an organized approach to understanding and predicting these price dynamics. It combines autoregressive (AR) and Moving Average (Paschke and Prokopczuk, 2010).ARIMA is a popular linear model that captures trends and autocorrelation in time series data, while SARIMA extends ARIMA by incorporating

1. Asst. Professor, Statistics Dept., Pratibha College of Commerce and Computer Studies, Chinchwad, Pune
2. Assoc. Professor, Statistics Dept., Babasaheb Ambedkar Marathwada University, Aurangabad

seasonality, making it more suitable for monthly price data (Hyndman and Athanasopoulos, 2018).GARCH models are particularly useful in capturing price volatility and conditional heteroskedasticity in agricultural markets (Engle, 1982).Meanwhile, the Holt-Winters method, available in both additive and multiplicative forms, is widely used for capturing trend and seasonality in time series data. The additive Holt-Winters model is suitable for data with constant seasonal variations, whereas the multiplicative model is preferred for data with seasonality that changes proportionally with the level of the series (Winters, 1960).

MATERIALS AND METHODS

This study uses historical data on the average prices of soybeans in the Latur district for 13 years from January 2010 until December 2022. The data used for this research be secondary. The entire data were separated monthly.The dataset was divided into training and testing sets, with historical data used for model training and recent data for validation.In this study, we applied ARIMA, SARIMA, GARCH, and Holt-Winters (additive and multiplicative) models to analyze and forecast monthly soybean prices.

Time series models

To analyse best forecasted time series model for soybean arrivals prices, multiple time series models are employed:

1.Autoregressive Integrated Moving Average (ARIMA)

i) ARIMA (p, d, q) (p, d, q) (p, d, q) is a linear forecasting model capturing trends and autocorrelation.

2. Seasonal ARIMA (SARIMA)

i) Extends ARIMA by incorporating seasonality: SARIMA (p, d, q) (P, D, Q) SARIMA (p, d, q) (P, D, Q) SARIMA (p, d, q) (P, D, Q) .

ii) The seasonal parameters (P, D, Q) (P, D, Q) (P, D, Q) account for periodic fluctuations in soybean prices.

3. Generalized Autoregressive Conditional Heteroskedasticity (GARCH)

i) Used to model volatility in soybean prices, which can fluctuate due to market uncertainties.

ii)The optimal GARCH (p, q) (p, q) (p, q) parameters are selected based on the log- likelihood function.

4.Holt-Winters Exponential Smoothing

i) Includes both Additive and Multiplicative models for trend and seasonality.

ii) The Additive model is used when seasonal variations are constant.

iii) The Multiplicative model is used when seasonal variations grow over time.

iv)The smoothing parameters (á,â,ã\alpha, \beta, \gammaá,â,ã) are optimized using grid search.

Model evaluation metrics: The performance of each model is assessed using:

Root Mean Square Error (RMSE): Measures overall prediction error.

Mean Absolute Error (MAE): Captures average absolute forecast errors.

Mean Absolute Percentage Error (MAPE): Evaluates percentage-based forecast accuracy.

The model with the lowest RMSE, MAE, and MAPE values is considered the best-performing model for soybean price prediction.

RESULTS AND DISCUSSION

Summary of the data

To understand the market dynamics, it is essential to look at yearly arrival (tonnes) and price (q^{-1}) patterns. Figure 1 provides a trend of the arrivals and prices of soybeans during the last 13 years in the Latur district.

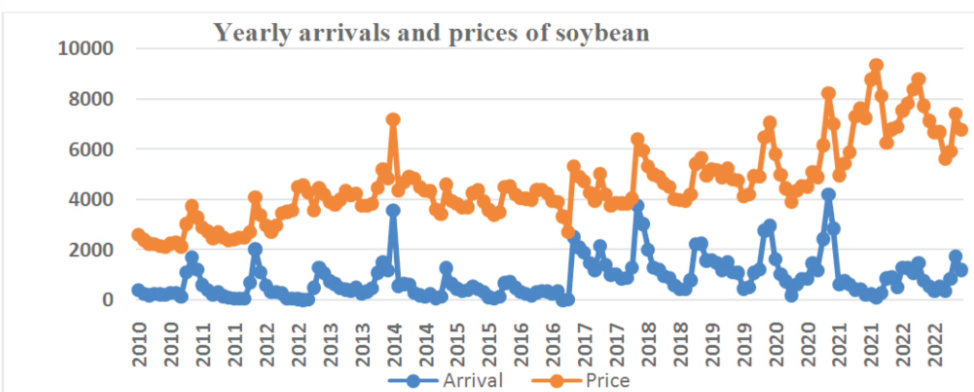

Figure1. Yearly arrivals (tonnes) and prices (q^{-1}) of soybean

The nature of the plot showed that there was no consistency in prices and arrivals of soybeans over the period. This upward movement suggests that prices was consistently rising over the span of the years.

To study the monthly behavior of soybean food grain, data were sorted month-wise. Descriptive statistics demonstrates the average, SD, and CV The standard deviation value of zero implied no volatility of the series and when its value is not equal to zero, the series indicate a volatility pattern. The productivity of soybean food grain showed a volatile pattern.

Table 1. Monthly behavior of arrival (tonnes) and prices (q^{-1}) of soybean

Months	Mean	SD	CV	Mean	SD	CV
	arrival (tonnes)			price (q^{-1})		
January	1223.94	892.349	72.9078	3472.908	1025.67	29.5335
February	790.119	456.853	57.8208	3554.146	1114.74	31.3646
March	664.572	386.85	58.2105	3689.301	1318.58	35.7406
April	739.172	593.754	80.3269	3969.747	1472.71	37.0983
May	564.338	377.666	66.9221	3989.405	1478.3	37.0556
June	477.566	328.337	68.7522	3862.107	1401.6	36.291
July	361.225	283.92	78.5994	4008.111	1655.54	41.3046
August	418.842	378.426	90.3504	4036.202	1778.96	44.075
September	435.176	407.991	93.7532	3780.249	1397.69	36.9736
October	1020.52	674.656	66.109	3377.582	974.822	28.8615
November	2064.24	997.444	48.3201	3545.928	1114.09	31.419
December	1543.72	875.892	56.7389	3647.014	1167.2	32.0041

The average monthly arrivals and associated variability measured in terms of the coefficient of variation of soybean in Table 1 discerns that the highest number of arrivals was registered in November month, followed by December month. From May onwards it started decreasing. In short, the study revealed that the average monthly arrivals were noted in the season of harvesting or post-harvesting and arrival was lowest in the time of sowing and pre-sowing period which was June –July was the sowing period for soybean (Gajic *et al*., 2018).The average monthly soybean prices along with the associated volatility displayed in Table 1 revealed that the average soybean price was the highest in July and August, which is the sowing period for soybean. Prices were more stable across the months. The reason for these lower fluctuation rates across the months and markets can be attributed to the overall uniform demand for soybeans throughout the year.

To study the strength and direction of the relationship between soybean prices and arrivals correlation coefficient was determined. Interpreting the correlation coefficient provides insight into how changes in soybean arrivals influence price movement and vice-versa, helping stakeholders predict and respond to market fluctuations. The correlation value (-0.6541) indicates that as the number of arrivals increases prices go on decreasing and vice-versa (Hassani *et al*., 2020).

Stationarity testing

To test the stationarity of the soybean prices , the Augmented Dickey-Fuller (ADF) test was conducted. The null hypothesis (H0) assumes that the given data was non-stationary, while the alternative hypothesis suggested stationarity. The p-value obtained from the ADF test was compared with the significance level (alpha). In this case, the p-value was greater than 0.05, indicating that the series was non-stationary. Entire time series analysis was done in R software. The plot of ACF and PACF were examined.

Augmented Dickey-Fuller Test

data: ts_data

Dickey-Fuller = -2.3478, Lag order = 5, p-value = 0.4314

alternative hypothesis: stationary

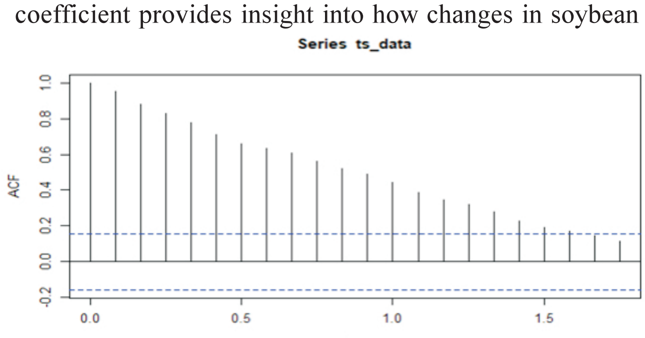

Figure 2. ACF plot of price data

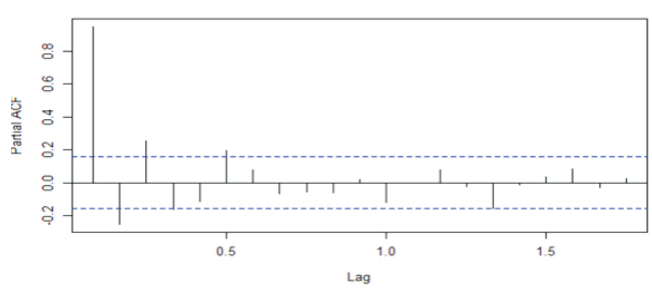

Figure 3. PACF plot of price data

Differencing and model identification

To achieve stationarity, differencing was performed on the series. After differencing, the ADF test was conducted again, resulting in a p-value less than 0.05, confirming stationarity. The plot of the different series was also examined.

Comparison of all three models

To evaluate the performance of the forecasting models, various error metrics were computed, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). These metrics provide insights into the accuracy and reliability of each model by quantifying the deviations between the predicted and actual values. Table 2 is showing all findings of error terms for different models.

Table 2.Comparison of error values

Model	RMSE	MAE	MAPE
ARIMA	2538.3718	2259.309432	1.072694977
SARIMA	2545.36634	2256.728832	33.13515596
HW Additive	2035.00196	1720.877318	24.84463706
HW Multiplicative	1943.88518	1593.367552	22.78096984
GARCH	492821.004	471577.1569	7625.096203

Table 2 shows comparative analysis of ARIMA, SARIMA, GARCH, and Holt-Winters (Additive and Multiplicative) models. Data revealed that Holt-Winters (Multiplicative) demonstrated superior forecasting accuracy, as it effectively captures both the trend and seasonality inherent in soybean prices and arrivals. The lower RMSE, MAE, and MAPE values for Holt-Winters (Multiplicative) indicated that it minimized forecasting errors better than the other time series models.

In summary, prices of soybean increased from 2010-2022. This period observed a consistently increasing demand for soybean. It is a really multipurpose crop, it has several uses as Edible oils, protein, flour, soy sauce, and soy paneer (Tiwari, 2017). It is observed that arrival and prices were negatively correlated. Consumers can take advantage of lower prices during periods of high arrivals. They might choose to buy larger quantities of soybean or soy-based products when prices are more favorable. Also, traders can adjust their inventory levels based on the correlation. They might reduce their stockpile during periods of high arrivals and lower prices to minimize losses due to price fluctuations. The increasing trend of soybean prices suggests a continuation of the prevailing upward trend observed in the coming years. Soybean prices return the ongoing demand for soybean in various sectors, like food, animal feed, industrial application, and biofuel. This study applied multiple time series models—ARIMA, SARIMA, GARCH, and Holt-Winters (Additive and Multiplicative)—to forecast monthly soybean arrivals and prices. Overall, Holt-Winters (Multiplicative) emerges as the most reliable model for predicting soybean prices, making it a valuable tool for stakeholders, including farmers, traders, and policymakers. Accurate price forecasting can aid in better decision-making regarding production planning, market interventions, and risk management in the agricultural sector. In short, this analysis helps as a valuable tool for market participants, policymakers, and other. Stakeholders seeking to make informed decisions in the context of soybean related industries (Heman, 2014). As with any market projection, ongoing monitoring and adaptability to changing conditions will be crucial for effectively navigating the dynamic landscape of soybean prices in the year ahead.

REFERENCES

Agbaeza, Emmanuel, Maureen Chiemeke Ngozichukwu, Anna Ogbo and I Ukpere Wilfred, 2020. Impact of Pricing Practice Management on Performance and Sustainability of Supermarkets in the Urban Area of Enugu State,Nigeria.**12**(15):1-23.

Barakade, A. J., L. B. Tonape and T. N. Lokahnde , 2011.Agricultural Land use Pattern In Satara District of Maharashtra. IRRJ. **11**(25) :54-57.

Chavan, R. S. and O. S. Jadhav, 2024. Agriculture Productivity and Land Use Efficiency of Tehsils in Chhatrapati Sambhaji Nagar District of Maharashtra.J.Soils and Crops.**33**(2):399-404.

Das, P., P. K. Mishra, and R. Singh ,2020. Analysing Price Volatility in Agricultural Markets: A Case Study of Oilseeds.J. Agric. Econ.**45**(3):215-232.

Enders, W. 2014.Applied Econometric Time Series. John Wiley & Sons.**45**(3):444-496.

Engle, R. F. 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica.**50**(4):987-1007.

Gajiæ, Boško ,Ranka Kresoviæ, Angelina Tapanarova , Ljubomir Životiæ and Mladen Todoroviæ , 2018. Effect of irrigation regime on yield, harvest index and water productivity of soybean grown under different precipitation conditions in a temperate environment. Agric.Water Manag.**210**:224-231.

Hassani, Hossein, Mansi Ghodsi, Xu Huang and Emmanuel Silva Sirimal, 2020. Is there a causal relationship between oil prices and tourist arrivals?J.Appl.Stat.**48**(1):191-202.

Heman, M. Eliot, 2014.Soybean seed proteome rebalancing. Frontiers in PlantScience.**5**(437):1-8.

Hyndman, R. J. and G. Athanasopoulos, 2018.Forecasting: Principles and Practice.Texts.

Merga, Bulti and Haji Jema,2011.Economic importance of chickpea: Production, value, and worldtrade. Cogent Food Agric. **5**(1615718):1-12.

Paschke, Raphale and Marcel Prokopczuk, 2010.Commodity derivatives valuation with autoregressive and moving average components in the price dynamics.JBF.**34(**11):2742-2752.

Tiwari, Siddharth Paul, 2022.Emerging trends in soybean industry.Gen Econs. **15**(1):1-17. Winters P. R. 1960. Forecasting Sales by Exponentially Weighted Moving Averages.MS.**6**(3):324-342.

Winter, P.R. 1960. Forecasting sales by exponentially weighted moving averages.MS.**6**(3):231-362.

Rec. on 26.02.2025 & Acc. on 06.03.2025