www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277-7695
ISSN (P): 2349-8242
NAAS Rating: 5.23
TPI 2023; 12(2): 1624-1628

www.thepharmajournal.com
Received: 07-11-2022
Accepted: 16-12-2022

Aishwarya V
Department of Agricultural Statistics, Bidhan Chandra Krishi Viswavidyalaya, Mohanpur, West Bengal, India

Banjul Bhattacharyya
Department of Agricultural Statistics, Bidhan Chandra Krishi Viswavidyalaya, Mohanpur, West Bengal, India

Nevashini N
Department of Agricultural Statistics, Bidhan Chandra Krishi Viswavidyalaya, Mohanpur, West Bengal, India

Gowthaman T
Department of Agricultural Statistics, Bidhan Chandra Krishi Viswavidyalaya, Mohanpur, West Bengal, India

Corresponding Author:
Gowthaman T
Department of Agricultural Statistics, Bidhan Chandra Krishi Viswavidyalaya, Mohanpur, West Bengal, India

Forecasting of area and production of onion and tomato by employing exponential smoothing models

Aishwarya V, Banjul Bhattacharyya, Nevashini N and Gowthaman T

Abstract
India is currently the world's second-largest producer of fruits and vegetables. Due to higher profitable price factors, horticultural crop demand and export opportunities have been steadily rising. In this context, an attempt has been made to forecast the area and production of Onion and Tomato. Presence of auto correlation in the data, time series forecasting model exponential smoothing were adopted. Holt's linear model found to be the best model because it has the ability to forecast with lowest RMSE and MAPE values compared the different models fitted for the data except for tomato production. In case of Tomato production, Damped trend model performed better among the selected models with less RMSE (215.30) and MAPE (13.01) values. Forecasted values were obtained using best fitted models.

Keywords: exponential smoothing, forecasting, time series analysis, agriculture

Introduction

The diverse agro-climatic condition in India enables it to grow a wide range of horticultural crops all year long. Due to higher profitable price factors, horticultural crop demand and export opportunities have been steadily rising. Vegetables, which have a high value as a product, enable farmers to earn a fair living even from tiny plots and at prices that are regularly higher than those for cereal crops. The country's horticulture production overtook food grain production in 2021–2022 with respective volumes of 315.72 million tonnes and 342.33 million tonnes. (Ministry of Agriculture and Farmers Welfare, 2022) [4].

Karnataka's economy has always been based primarily on agriculture. The state occupies a prominent place in the Horticulture map of the country and one of the major vegetable-growing State contributes about 5% to the total country's production. Horticulture generates 40% of the total income of the state. This accounts for 17% of the GDP of the state Horticulture has taken a front-line position in Karnataka agriculture and the sector is growing at a rapid pace. During the year 2020-21, an area of 5.21 lakh hectares was covered by vegetable crops and production was 108.85 lakh metric tonnes (Horticulture Department, Government of Karnataka) [2]. Karnataka is the third largest producer of onion and tomato in India and contributes about 8.93% and 10.23% respectively in production.

In agriculture, which is rife with uncertainty, reliable and timely forecasting offers crucial and valuable inputs for proper foresight and informed planning. Forecasting of various aspects of agriculture has become crucial in the current climate of change (Agrawal, 2005) [1]. Forecasting of area and production for decision makers has become even more crucial as a result of the fluctuation in agricultural commodity prices over the past ten years (Mase and Prokopy, 2014) [8].

To forecast crop yield and production, various forecasting techniques, such as the exponential smoothing method, have traditionally been used. Prediction is made by analysing underlying patterns in the past time-series data. Forecasting future events requires a knowledge of the past. Time series analysis explains why trends appear in historical data and how underlying patterns or processes may be able to account for them.

The exponential smoothing method is just one of many forecasting techniques. By taking into account the most recent data, the technique of exponential smoothing can identify significant changes in data. For short-term forecasts, exponential smoothing is a highly accurate forecasting technique. Simple moving averages and exponential smoothing both estimate future values using historical data or lags, but there is a significant distinction between the two: Simple moving averages take all previous observations into account equally, whereas exponential smoothing gives weights that decrease exponentially over time.

Masuda and Goldsmith (2008) [9] conducted a study on world Soybean production, yield and harvested area. Using damped trend exponential model, they finalized that world production of soybeans was predicted to be increased by 2.1 per cent annually to 359.7 million tonnes by 2030. Chaithra *et al*., (2019) [6] forecasted the area and production of cashew nut using ARIMA and Exponential smoothing, Holt's linear trend model performed better in their study with lowest RMSE and MAE.

Considering the fact that the growth of horticulture sector and the demand for consumption of vegetables an attempt has been made in this paper to forecast the area and production of onion and tomato in Karnataka.

Materials and Methods

Data

The secondary data on area and production of Onion and Tomato for Karnataka state were collected for period of 1997 to 2021 from India stat socio-economic statistics and horticultural statistics. Data from 1997 to 2017 were used for model building and from 2017 to 2021 were used for model validation purpose.

Exponential smoothing models

Exponential smoothing model is one of the successful parametric non-linear forecasting techniques to produce a smoothed time series. Exponential smoothing assigns exponentially decreasing weights as the observation get older.

Single exponential smoothing (SES) method

This method is also known as the method to estimate future value using a single weight/parameter. Recent observations are given higher weights, while distant observations are given lower weights.

Let S_{t+1} denote the estimator of the level at time t+1. Given St, once the observation at time t, S_t is previous observed value, becomes available, SES due to Brown (1963) [5] updates the level estimator via the recurrence equation.

$$S_{t+1} = \alpha Y_t + (1-\alpha) S_t$$

Where, α is a smoothing parameter taking values in the interval (0, 1).

S_{t+1} = the estimator at time t.

Y_t = is a weighted average of the latest observation.

Holt's linear trend method

SES does not predict the time series with a local linear trend well. It can be shown that the SES forecast for a time series with a local linear trend tends to lag behind the series itself (Brown 1963) [5]. Holt (1957) [7] extended SES by adding one additional upgrading equation for the slope (trend) named Holts method to better forecast time series with a local linear trend.

$$S_t = \alpha Y_t + (1-\alpha)(S_{t+1} + \beta_{t+1})$$
$$b_t = \beta(S_t - S_{t+1}) + (1-\beta) b_{t-1}$$

Where,

S_t = Smoothed value at time period t.

S_{t+1} = Smoothed value at time period $(t-1)$.

α = Level Smoothing Constant ranges 0 to 1.

Y_t = Actual area or production or productivity at time period t.

b_t = Trend estimate of the time period t

b_{t+1} = Trend estimate of time period $(t-1)$

β = Trend Smoothing Constant

Brown's Linear trend.

Brown (1963) [5] also introduced a double exponential smoothing method for local linear trend forecasting, which employs a single smoothing parameter to smooth both the level and the trend. This method is suitable for series with a linear trend and no seasonality.

$$S_t = \alpha Y_t + (1-\alpha)(S_{t-1})$$
$$b_t = \alpha(S_t - S_{t-1}) + (1-\alpha) b_{t-1}$$

Where,

S_t = smoothened value at time period t

S_{t-1} = smoothened value at time period $(t-1)$

α = level smoothing constant

Y_t = actual value at time period

Y_t = trend estimate of time period t

b_{t-1} = trend estimate of the period $(t-1)$

Damped trend model

This method includes a damping parameter in addition to the smoothing parameters (level) and (trend), which also have values between 0 and 1 as in Holt's method. (0<ϕ<1)

$$\hat{y}_{t+hlt} = l_t + (\phi + \phi^2 + \cdots + \phi^h) b_t$$
$$S_t = \alpha X_t + (1-\alpha)(S_{t-1} + \phi b_{t-1})$$
$$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)\phi b_{t-1}$$

If ϕ=1, the method is the same as Holt's linear method. For values between 0 and 1, dampens the trend such that it finally approaches a constant in future. In fact, the forecasts converge to $l_t + \phi b_t/(1-\phi)$ as $h \to \infty$ for any value 0<ϕ<1. As a result, short-run predictions trended while long-run forecasts stay unchanged.

Where,

S_t = smoothened value at time period t

S_{t-1} = smoothened value at time period $(t-1)$

α = level smoothing constant

Y_t = actual value at time period

Y_t = trend estimate of time period t

b_{t-1} = trend estimate of the period $(t-1)$

Ljung Box test

This test statistic is employed to test for autocorrelation.

$$Q(m) = n(n+2) \sum_{j=1}^{m} \frac{r_j^2}{n-j}$$

Where,

r_j = the accumulated sample autocorrelations,

m = the time lag.

R^2

It is known as co efficient of determination. It indicates the proportion of variation present in response explained by the model

$$R^2 = 1 - \frac{SS_{res}}{SS_r}$$

Where, SS_{res} and SS_r are the residual and total sum of squares, respectively.

RMSE

Root Mean Square Error is the standard deviation of the residuals. It measures the spread of these residuals.

$$\text{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i 2}$$

Where, $e = (y_t - \hat{y}_t)$

MAPE

It is a measure of accuracy of a method for constructing fitted time series values in statistics, specifically in trend estimation. It usually expresses accuracy as a percentage.

$$M = \frac{1}{N}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \times 100$$

Where,

y_t = actual value

$\hat{y}_t$ = forecast value

Result and Discussions

Forecasting of Area and production of Onion

The data of area and production of Onion were tested for autocorrelation using Durbin Watson test since it was time series data. The test statistic values were found to be 0.38 and 0.36, respectively for area and production. As values did not lie between 1.5 to 2.5, exponential smoothing models were used to forecast the area and production.

Different exponential smoothing models were fitted for area and production of Onion. Based on the Ljung-Box Q statistic given in table 1, it could be said that the holt's model fitted well with data as the statistic was non-significant for both area and Production of Onion. Model fit statistics like R^2 (53.45%), RMSE (27.20), and MAPE (16.42) for Area and R^2 (64.32%), RMSE (656.30), and MAPE (19.65) for Production. These fit statistic values of Holt's model were also more satisfactory than the other exponential smoothing models, as given in the same table. The estimates of the parameters were given in table 2 for Onion data set. Holt's model was observed to have better forecasting accuracy so forecasting was done for next four years using Holt's model, as presented in Table 3. Outcomes emanated from the investigation were in line with the findings of Prabakaran *et al*. (2013) [10].

Table 1: Model fit statistics and ljung-box q statistics for area and production of onion

Onion	Model	R2	RMSE	MAPE	Ljung-Box		
					statistics	DF	sig.
Area	Simple	0.30	32.36	22.46	16.82	17	0.46
	Holt's	0.53	27.20	16.42	13.79	16	0.61
	Brown	0.28	32.86	25.56	18.39	17	0.36
	Damped	0.53	27.95	19.99	13.83	15	0.53
Production	Simple	0.53	736.58	20.78	8.45	17	0.95
	Holt's	0.64	656.30	19.65	15.05	16	0.52
	Brown	0.51	748.06	35.05	7.88	17	0.96
	Damped	0.64	672.41	38.61	15.33	15	0.40

Table 2: Estimate of the fitted linear trend model parameter for area and Production of Onion

Onion	Models	Parameters	Estimates	SE	T statistic	p- value
Area	Holt's	Alpha (Level)	0.78**	0.287	3.13	<0.001
		Gamma (Trend)	0.64*	0.252	2.43	0.01
Production	Holt's	Alpha (Level)	0.60**	0.193	3.10	0.005
		Gamma (Trend)	0.41*	0.074	1.37	.046

**: Significant level at 1%, *: Significant level at 5%

Table 3: Forecasted values for Area and Production of Onion

Year	Forecasted Area (ha)	Forecasted Production (Tonnes)
2022	208.60	2999.97
2023	213.32	3115.20
2024	218.04	3230.43
2025	222.76	3345.65

In order to check the adequacy of the Holt's model, residual analysis was carried out, ACF and PACF plot of the residuals were obtained for both area and production, in which all the spikes were found within limits as shown in figure 1 and 2. Indicating that the model was adequate.

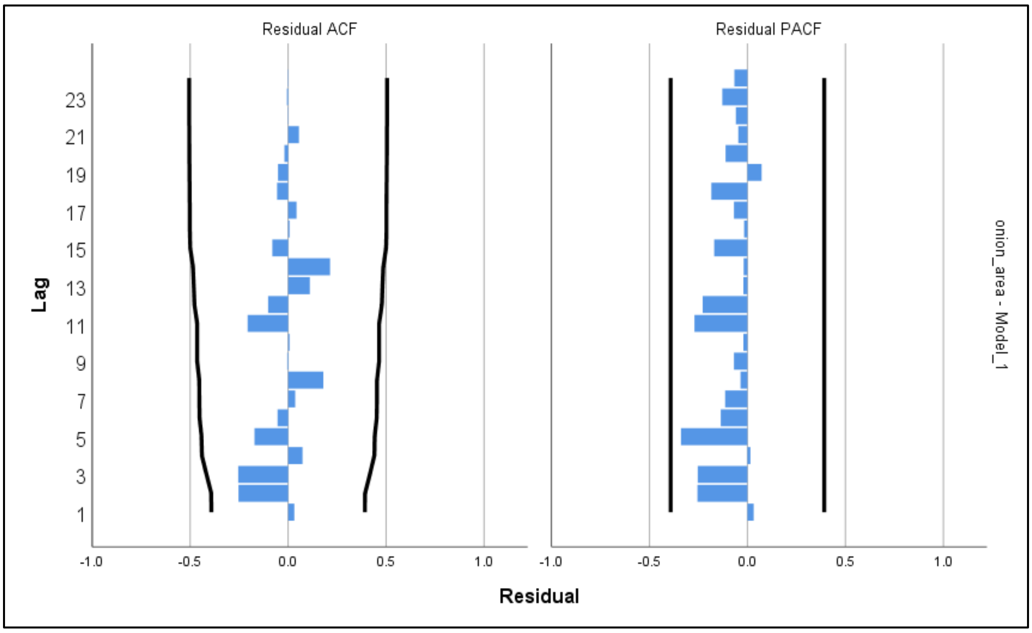

Fig 1: Residual autocorrelation and partial autocorrelations for Onion Area

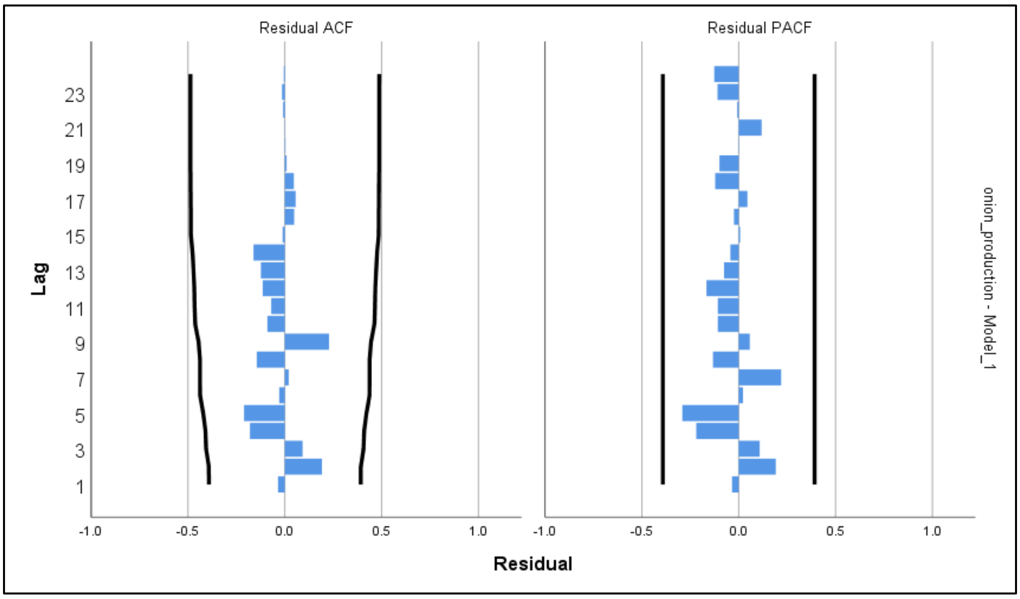

Fig 2: Residual autocorrelation and partial autocorrelations for Onion production

Forecasting of Area and production of Tomato

The data of area and production of Tomato were tested for autocorrelation using Durbin Watson test since it was time series data. The test statistic values were found to be 0.63 and 0.26, respectively for area and production. As values did not lie between 1.5 to 2.5, exponential smoothing models were used to forecast the area and production.

Different exponential smoothing models were fitted for area and production of Tomato. Based on the Ljung-Box Q statistic given in table 4, it could be said that the holt's model fitted for Area and Damped trend model fitted for Production data as the statistic was non-significant. Model fit statistics like R^2 (56.57%), RMSE (7.65), and MAPE (12.54) for Area and R^2 (84.20%), RMSE (215.30), and MAPE (13.01) for Production. These fit statistic values of Holt's model and Damped trend model were more satisfactory than the other exponential smoothing models, as given in the same table. The estimates of the parameters were given in table 5 for Tomato data set. For Area, Holt's model was observed to have better forecasting accuracy so forecasting was done for next four years using Holt's model. For Production, Damped trend model was observed to have better forecasting accuracy so forecasting was done for next four years using Damped Trend model as presented in Table 6.

Table 4: Model fit statistics and Ljung-Box Q statistics for Area and Production of Tomato

Tomato	Model	R2	RMSE	MAPE	Ljung-Box		
					statistics	DF	sig.
Area	Simple	0.54	7.66	12.67	6.50	17	0.99
	Holt's	0.56	7.65	12.54	6.75	16	0.98
	Brown	0.50	7.94	12.88	6.39	17	0.99
	Damped	0.56	7.87	12.86	6.72	15	0.96
Production	Simple	0.81	231.07	12.15	9.67	17	0.92
	Holt's	0.84	216.16	13.02	8.81	16	0.92
	Brown	0.80	236.98	14.90	8.93	17	0.94
	Damped	0.84	215.30	13.01	8.74	15	0.89

Table 5: Estimate of the fitted linear trend model parameter for area and Production of Tomato

Onion	Models	Parameters	Estimates	SE	T statistic	P value
Area	Holt's	Alpha (Level)	0.53	0.161	1.864	0.04
		Gamma (Trend)	0.15*	0.093	0.307	0.03
Production	Damped	Alpha (Level)	0.08*	0.119	0.689	0.04
		Gamma (Trend)	0.32*	0.210	0.002	0.05
		Phi (Trend damping factor)	0.98**	0.023	43.193	<0.001

**: Significant level at 1%, *: Significant level at 5%

Table 6: Forecasted values for Area and Production of Tomato

Year	Forecasted Area (ha)	Forecasted Production (Tonnes)
2022	67.18	2454.82
2023	68.52	2530.45
2024	69.85	2606.07
2025	71.18	2681.70

In order to check the adequacy of the Holt's model and Damped Trend Model, residual analysis was carried out, ACF and PACF plot of the residuals were obtained for both area and production, in which all the spikes were found within limits as shown in figure 3 and 4. Indicating that the model was adequate.

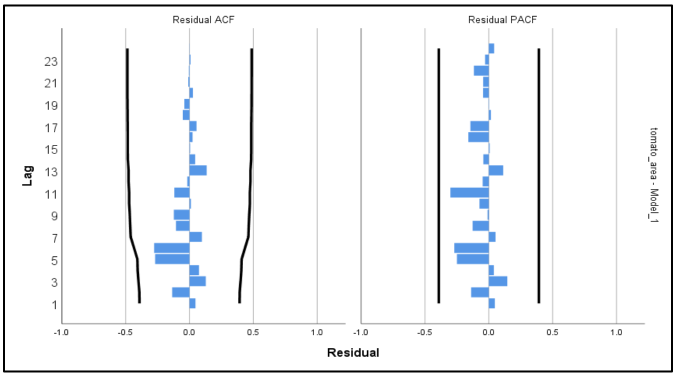

Fig 3: Residual autocorrelation and partial autocorrelations for Tomato Area

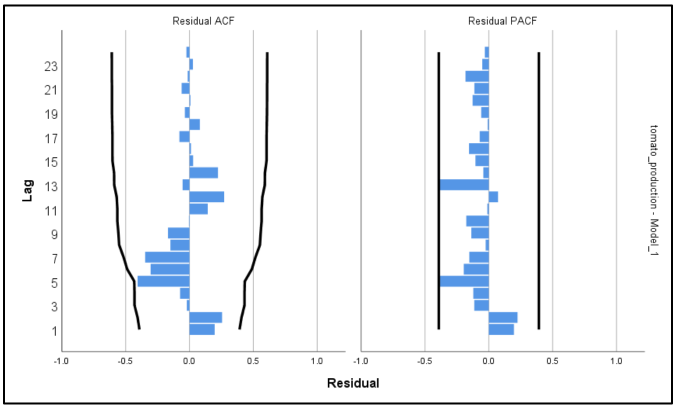

Fig 4: Residual autocorrelation and partial autocorrelations for Tomato production

Conclusion

Due to presence of Autocorrelation in the data, time series forecasting model Exponential smoothing models were adopted. Different exponential smoothing models were applied and model were fitted for Area and Production of Onion and Tomato. Holt's linear models were found to have better forecasting accuracy with lowest RMSE and MAPE values among the different models fitted for forecasting area and production of Onion. In case of Tomato production, Damped trend model, with RMSE value (215.30), was observed to have better forecasting ability among the fitted models. Area and Production of Onion and Tomato was forecasted for next four years. Forecasted values from this study are expected to help planners in recommending policies regarding production of Onion and tomato to strengthen the economy.

References

1. Agarawal R. Forecasting techniques in crops. Indian Agricultural Statistics Research Institute, New Delhi. 2004.
2. Anonymous. Department of Horticulture, Government of Karnataka. [http://www.horticulture.kar.nic.in/], 2021a. [visited on February, 04,2023].
3. Anonymous. Economic Survey of Karnataka. Planning, Programme Monitoring and Statistics Department. [https://planning.karnataka.gov.in/info-4/Reports/Economic+Survey+2021-22/en], 2021. [Visited on February, 01,2023].
4. Anonymous. Ministry of Agriculture & Farmers Welfare. [https://pib.gov.in/PressReleasePage.aspx?PRID=1841480], 2022. [Visited on February, 02,2023].
5. Brown RG. Smoothing forecasting and prediction of discrete time series, USA, Prentice- Hal; c1963.
6. Chaithra M, Pandit P, Bakshi B. Forecasting of area and production of cashew nut in Dakshina Kannada using ARIMA and exponential smoothing models. Journal of Reliability and Statistical Studies. 2019;30:61-76.
7. Holt CC. Forecasting seasonal and trends by exponentially weighted moving averages, ONR Research Memorandum, Carnigie Institute; c1957.
8. Mase AS, Prokopy LS. Unrealized potential: A review of perceptions and use of weather and climate information in agricultural decision making. Weather, Climate, and Society. 2014;6(1):47-61.
9. Masuda T, Goldsmith PD. World soybean production: Area harvested yield, and long- term projections, International Food and Agribusiness Management Review. 2008;12(4):143-161.
10. Prabakaran K, Sivapragasam C, Jeevapriya C, Narmatha A. Forecasting cultivated area and production of wheat in India using ARIMA model, Golden Research Thoughts. 2013;3(3):1-7.