Journal of Pharmacognosy and Phytochemistry
Available online at www.phytojournal.com

E-ISSN: 2278-4136
P-ISSN: 2349-8234
www.phytojournal.com
JPP 2020; Sp9(2): 317-320
Received: 18-01-2020
Accepted: 20-02-2020

B Ramana Murthy
Assistant Professor, Department of Statistics and Computer Applications, Acharya N.G. Ranga Agricultural University, S.V. Agricultural College, Tirupati, Andhra Pradesh, India

S Govinda Rao
Assistant Professor, Department of Statistics and Computer Applications, Acharya N.G. Ranga Agricultural University, Agricultural College, Naira, Andhra Pradesh, India

SK Nafeez Umar
Assistant Professor, Department of Statistics and Computer Applications, Acharya N.G. Ranga Agricultural University, Agricultural College, Bapatla, Andhra Pradesh, India

Corresponding Author:
B Ramana Murthy
Assistant Professor, Department of Statistics and Computer Applications, Acharya N.G. Ranga Agricultural University, S.V. Agricultural College, Tirupati, Andhra Pradesh, India

Statistical model for forecasting production of ginger in India

B Ramana Murthy, S Govinda Rao and SK Nafeez Umar

Abstract
The present research study was carried out to fit different Linear, Non – Linear and ARIMA models on production of Ginger (*Zingiber officinale*) in India for the period 1970-71 to 2018-19. Statistically the best fitted model was chosen on the basis of goodness of fit criteria's like R^2, Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). Among all the models, ARIMA (4, 1, 2) was found to be the best fitted model for future forecasting. The results were shown, there is an increasing trend for certain long time of period and then fluctuating for short time of period and again it was increasing for a certain long time of period. It means that for a certain long time of period, the production of Ginger in India has been increasing. Based on this trend to forecast production of Ginger crop for next four years. The forecasted production of Ginger in India for the years 2019-20 to 2020-23 to be 2212.3, 2114.9, 2211.8, 2506.3 in thousand metric tonnes respectively.

Keywords: Ginger, R^2, RMSE, MAPE, ARIMA, forecast, and production

1. Introduction
Ginger (*Zingiber officinale*) is an important commercial crop grown for its aromatic rhizomes which is used both as a spice and a medicine. Ginger of commerce is the dried rhizome. It is marketed in different forms such as raw ginger, dry ginger, bleached dry ginger, ginger powder, ginger oil, ginger oleoresin, ginger candy, ginger beer, brined ginger, ginger wine, ginger squash, ginger flakes etc.

India is the largest producer and consumer of ginger contributing about 31% of total global production followed by China, Nepal, Indonesia, Nigeria, and Thailand. The leading states producing ginger in India are Orissa, Kerala, Karnataka, Arunachal Pradesh, West Bengal, Sikkim and Madhya Pradesh. Among which Kerala contribute 33 per cent to the total production in India (Gupta, 2008). India produces 6, 83,000 tons of ginger per annum that is almost 1/3rd of world's total production (FAO). 30 per cent of total production of ginger in India is transferred to dry ginger, 50 per cent is taken as fresh or green ginger and the rest part is used as seed materials. Kerala is the largest producer of dry ginger in India, which has taken a major share in export.

Sudha *et al*. (2013) [6] studied to measure the growth trends of area, production and productivity of maize between 1970-71 to 2008-09 and to estimate the future projections up to 2015 AD by using the growth functions like linear, logarithmic, inverse, quadratic, cubic, compound, power and exponential. Based on highest coefficient of determination (R^2) and its adjusted R^2 they concluded that among all the models cubic function was found to be best fitted model for future projections of maize area, production and productivity.

Paland and Mazumdar (2015) [4] were studied forecasting groundnut production of India using nonlinear growth models. They have pointed out country's groundnut production is likely to be increased with a slow and steady rate.

Ramana Murthy *et al*. (2018) [5] studied the appropriate ARIMA model for forecasting sunflower production in India and they found that ARIMA (4, 1, 4) model was the best fitted model to forecast Sunflower Production in India and also they were observed that there is a downward and upward trend on production of sunflower in India for forecasted years.

Ramana Murthy *et al*. (2018) [5] attempted to study forecasting groundnut area, production and productivity of India using ARIMA model. They have concluded that ARIMA (2, 1, 3), ARIMA (3, 0, 3) and ARIMA (2, 1, 3) models were best fitted to forecast area, production and productivity of groundnut in India for four leading years and also they have found that there was a decreasing trend on area and fluctuations on production and productivity from the period 2016-17 to 2019-2020.

The objective of the present study was to fit different linear, non-linear and appropriate Box-Jenkins Auto Regressive Integrated Moving Average (ARIMA) models for the time series of

Ginger production in India and to make four year forecasts with appropriate model.

2. Materials and Methods

The data of study for a period of 49 years (1970-71 to 2018-19) in India pertaining to Production ('000 MT) of Ginger were collected from the source of Ministry of Agriculture, Govt. of India in www.indiastat.com. In order to examine the nature of change and degree of relationship in production of Ginger in India by various linear, non-linear and ARIMA statistical models were worked out by using SPSS 22 version.

2.1 Linear and Non-Linear Growth Models: The linear and non-linear growth models for the crop characteristic i.e., Production of Ginger in India are estimated by fitting the following functions.

Parametric Trend models

Model	Functional form
Linear function	$y_t = a + bt$
Logarithmic function	$y_t = a + b\ln(t)$
Inverse function	$y_t = a + b/t$
Quadratic function	$y_t = a + bt + ct^2$
Cubic function	$y_t = a + bt + ct^2 + dt^3$
Compound function	$y_t = ab^t$
Power function	$y_t = at^b \quad (or) \quad \ln(y_t) = \ln(a) + b\ln(t)$
S- Curve function	$y_t = Exp(a + b/t) \quad (or) \quad \ln(y_t) = a + b/t$
Growth function	$y_t = Exp(a + bt) \quad (or) \quad \ln(y_t) = a + bt$
Exponential function	$y_t = ae^{bt} \quad (or) \quad \ln(y_t) = \ln(a) + bt$

2.2 Auto Regressive Integrated Moving Average (ARIMA)

The ARIMA methodology is also called as Box-Jenkins methodology (Box and Jenkins 1976) [2]. The Box-Jenkins procedure is concerned with fitting a mixed ARIMA model to a given set of data. The main objective in fitting ARIMA model is to identify the stochastic process of the time series and predict the future values accurately. This method shave also been useful in many types of situations which involve the building of models for discrete time series and dynamic systems. However the optimal forecast of future values of a time series are determined by the stochastic model for that series. A stochastic process is either stationary or non-stationary. The first thing to note is that most time series are non-stationary and the ARIMA models refer only to a stationary time series. Since the ARIMA models refer only to a stationary time series the first stage of Box-Jenkins model is for reducing non-stationary series to a stationary series by taking the differences.

The ARIMA (p, d, q) process is given by

$$y_t = \theta_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \dots\dots\dots + \phi_p y_{t-1} + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \theta_3\varepsilon_{t-3} - \dots\dots - \theta_q\varepsilon_{t-q} \quad (1)$$

Where y_t and ε_t are the actual value and random error at time period t, respectively. $\phi_i (i = 1,2,3,\dots\dots, p)$ and θ_j (j=1, 2, 3,…..,q) are model parameters. p and q are integers and often referred to as orders of the model. Random errors ε_t are assumed to be independently and identically distributed with a mean of zero and a constant variance of σ^2.

The main stages in setting up a Box-Jenkins forecasting model are as follows:

1. Identification
2. Estimating the parameters
3. Diagnostic checking
4. Forecasting

3. Results and Discussion

In the present study, the data for production of Ginger for the period of 49 years (1970-71 to 2018-19) were used for the study.

3.1 Model Identification.

Among several models Linear, Non-linear and ARIMA (p, d, q) studies the goodness of fitted models were examined by highest R^2 value, lowest RMSE (Residual Mean Square Error) and MAPE (Mean Absolute percentage Error) values. Based on these criterions ARIMA (4, 1, 2) is found to be the best fitted model for study period. The Coefficient of determination (R^2), Mean Absolute Percentage Error (MAPE) and Residual Mean Square Error (RMSE) are given by

$$R^2 = 1 - \left[\sum_{t=1}^{n} (y_t - \hat{y}_t)^2 / \sum_{t=1}^{n} (y_t - \bar{y})^2 \right] \quad (2)$$

$$MAPE = \frac{100}{n} \sum_{t=1}^{n} \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (3)$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} \left(y_t - \hat{y}_t \right)^2}{n}} \quad (4)$$

Where y_t is the actual observation for time period't' and $\hat{y}_t$ is the predicted value for the same period and $\bar{y}$ is the overall sample mean of observations. The models and the corresponding values are shown in table (1).

Table 1: Linear, Non-linear and Time series models of Ginger production in India

	Parameter				Criteria		
Model	**a**	**b**	**c**	**d**	**R^2**	**RMSE**	**MAPE**
Linear	-110.266	19.250**			0.628**	209.4847	70.8098
Logarithmic	-346.218*	243.097**			0.387**	269.0192	97.6735
Inverse	441.579**	-772.141*			0.125*	321.2679	114.9941
Quadratic	244.978**	-22.543**	0.836**		0.817**	146.8137	63.8343
Cubic	-97.643	55.759**	-3.040**	0.052**	0.929**	91.8137	23.7576
Compound	63.578**	1.058**			0.804**	151.3255	33.5627
Power	18.206**	0.900**			0.796**	240.4786	36.0788
S-Curve	5.913**	-3.900**			0.480**	323.5215	51.1341
Growth	4.152**	0.056**			0.804**	151.3254	33.5627
Exponential	63.578**	0.056**			0.804**	151.3254	33.5627
Time Series Models							
ARIMA (1, 0, 1)					0.886	121.240	16.277
ARIMA (3, 0, 3)					0.922	104.608	16.242
ARIMA (2, 1, 4)					0.935	95.829	13.918
ARIMA (3, 1, 4)					0.934	97.538	13.905
ARIMA (4, 1, 2)					0.936	95.037	13.857
ARIMA (2, 2, 3)					0.932	96.721	15.052
ARIMA (3, 2, 4)					0.934	98.118	14.447
ARIMA (3, 3, 3)					0.922	105.375	16.610

1. The value of the criterion for a model with bold numbers shows that the model is better than the other models with respect to that criterion.
2. ** ,* indicates significant at 1% and 5% level of probability respectively.

3.2 Model Estimation and Verification

The parameters of the model were estimated by using SPSS 22 package. The ARIMA (4, 1, 2) model is best fitted models for production of Ginger. The model verification (or) diagnosed by the Ljung-Box Q statistic. The Ljung-Box Q statistic is to check the overall adequacy of the model. The test statistic Q is given by

$$Q_n = nr(nr+2)\sum_{l=1}^{n}\frac{r_l^2(e)}{nr-l} \qquad (5)$$

Where $r_l(e)$ is the residual autocorrelation at lag l, nr is the number of residual, n is the number of time lags included in the test for model to be adequate, p-value associated with Q statistics should be large $(p-value>\alpha)$. The results of estimation are reported in Table 2.

Table 2: Estimates of the fitted ARIMA (4, 1, 2) model for Ginger Production

Model fit Statistics			Ljung-Box Q (18)	
R-Square	RMSE	MAPE	Statistic	p-value
0.936	95.037	13.857	2.918	0.996

3.3 Forecasting with ARIMA Model

After the identification of the model and its adequacy check, it is used to forecast the production of Ginger in the next four periods. Hence we used the identified ARIMA model to forecast production of Ginger for the years 2019-20 to 2022-23. The forecasting results are presented Table. 3. And also the diagrams of actual and forecasted values are presented in Fig.1.

Table 3: Forecasted values of Ginger Production with 95% Confidence Level (CL)

Year	Production		
	Forecasted values	**LCL**	**UCL**
2019-20	2212.3	1870.9	2574.9
2020-21	2114.9	1644.0	2629.3
2021-22	2211.8	1666.1	2814.0
2022-23	2506.3	1805.9	3290.2

LCL: Lower Confidence Level
UCL: Upper Confidence Level

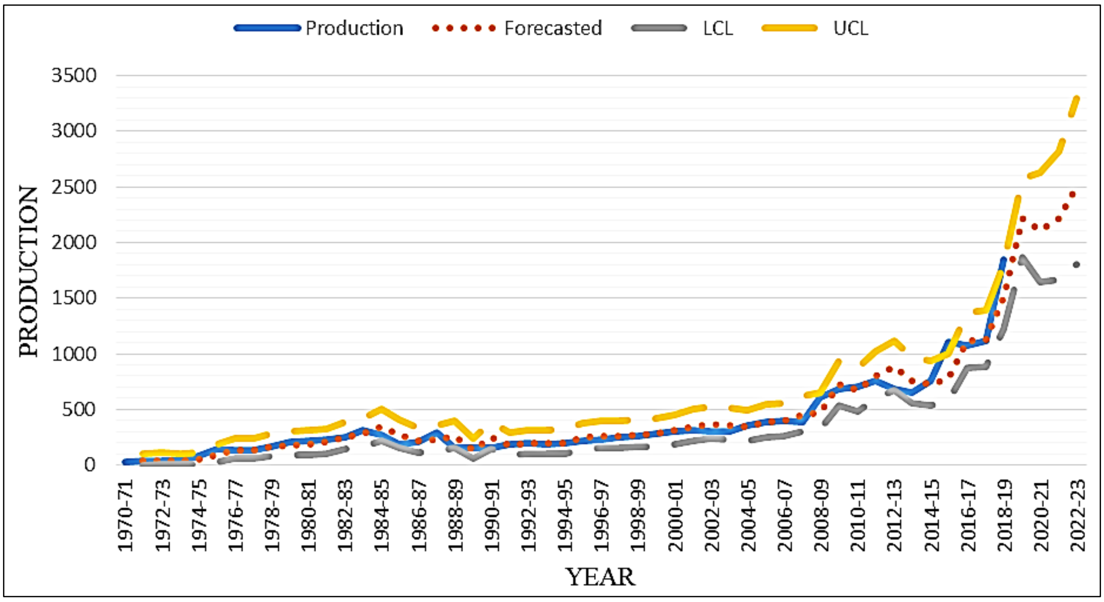

Fig 1: Forecasted Ginger Production (1970-71 to 2022-23)

4. Conclusion

In the present study, the developed ARIMA (4, 1, 2) is the best model for forecasting of Ginger production based on R^2 , RMSE and MAPE criterions. The study revealed that in coming next years there is a fluctuating trend on production up to 2020-21 and then it may increase up to certain long time of period. Ginger is an important commercial crop which is used both as a spice and a medicine and also it is a huge usage in our daily life. Ginger has a very long history of use in various forms of traditional/alternative medicine. It has been used to help digestion, reduce nausea and help fight the flu and common cold, to name a few. Ginger can be used fresh, dried, powdered, or as an oil or juice, and is sometimes added to processed foods and cosmetics. It is a very common ingredient in recipes. So that there is a need to improve production of Ginger in India.

5. References

1. Bijaya B Bag. Ginger Processing in India (*Zingiber officinale*): A Review, International Journal of Current Microbiology and Applied Sciences. 2018; 7(4). ISSN: 2319-7706.
2. Box GEP, Jenkin GM. Time series of analysis, Forecasting and Control, Sam Franscico, Holden Day, California, USA, 1976.
3. Gupta RK. Ginger-A wonderful Spice: An overview. International Journal of Plant Research. 2008; 21(1):1-10.
4. Paland S, Mazumdar D. Forecasting Groundnut Production of India using Nonlinear Growth Models. Journal Crop and Weed, 11 (Special Issue), 2015, 67-70.
5. Ramana Murthy B, Nafeez Umar SK, Hari Babu O. ARIMA Models for forecasting Sunflower Production in India. International Journal of Pure & Applied Bioscience. 2018; 6(6):1121-1126.
6. Sudha CHK, Rao VS, Suresh CH. Growth trends of maize crop in Guntur district of Andhra Pradesh. International Journal of Agricultural Statistical Sciences. 2013; 9(1):215-220.
7. www.indiastat.com.