e-ISSN: 2582-502X

Asiatic Society for Social Science Research. 2(1): June 2020, 131-136.

Research Article

www.asssr.in

(Peer Reviewed)

***Corresponding Author**

Mr. Subrat Kumar Mahapatra

Dept. Of Agril. Statistics, Odisha University of Agriculture & Technology (OUAT),

Bhubaneswar, Odisha

Email: smsubrat362@gmail.com

Received on 11.04.2019
Modified on 02.06.2020
Accepted on 26.06.2020

ARIMA Model for Forecasting of Black Gram Productivity in Odisha

Mr. Subrat Kumar Mahapatra[1] and Dr. Abhiram Dash[2]

[1] Dept. Of Agril. Statistics, Odisha University of Agriculture & Technology (OUAT), Bhubaneswar, Odisha.

[2]Assistant Professor, Dept. Of Agril. Statistics, Odisha University of Agriculture & Technology (OUAT), Bhubaneswar, Odisha.

ABSTRACT:

A study was conducted on forecasting the productivity of black gram in Odisha. Box-Jenkins Autoregressive integrated moving average (ARIMA) time-series methodology was considered for Black gram yield forecasting. Different ARIMA models are selected on the basis of Autocorrelation Function (ACF) and Partial autocorrelation Function (PACF) at various lags .The data from 1971-72 to 2006-07 are used for model building of different ARIMA models and from 2007-08 to 2015-16 is used for successful cross-validation of the selected model, which is based on the Mean absolute percentage error (MAPE). To check the stationarity, ARIMA Models are fitted to the original time series data as well as first difference data. Based on the significant coefficient of autoregressive and moving average components, the possible ARIMA Models are identified. Based on low value of Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE), the best fitted ARIMA models are selected. ARIMA (0,1,1) without constant found to be best fitted model for Black gram productivity having absolute percentage error ranging from 19.99% to 43.29% in cross-validation of the model. The best fitted ARIMA model has been used to forecast the productivity of black gram for the year 2016-17 to 2018-19. The model showed the forecast in productivity for the year 2018-19 to be about 221.45 kg per hectare with lower and upper limit 90.36 and 392.89 kg per hectare respectively.

KEY WORDS: ARIMA, Black gram, Productivity, ACF, PACF, RMSE, MAPE

Introduction

Pulses are an important commodity group that are grown in all the districts of Odisha. Pulses are the major source of protein & it provide high-quality protein complementing cereal proteins for the pre-dominantly substantial vegetarian population of the country. Currently Pulses are grown in about 18.7 lakh hector area with production of 9.4 lakh tonnes, having productivity 502 kg/ha. Pulse crops are considered as the important sources of Protein. Black gram (*Vigna mungo* L) is the important pulse crops in Odisha. The contribution of Black gram to the total pulse area in Odisha is 28.04%, during the year 2015-16. The share of Black gram towards total production of the State is 25%, during 2015-16.

Black gram is an important pulse crops in the context of Odisha, which is mostly grown from the ancient times. This is also one of the most important & highly priced pulses of India & other south Asian country.. The Mahanadi Delta, the Rusikulya plains and the Hirakud & Badimula regions are favorable for the cultivation of pulses. The Purpose plateau is dominated by pulses. The Rusikulya plain is the most important agricultural region of Orissa and is dominated by pulse.

Estimation of crop area, prediction/forecasting of Production & Productivity of crops are highly essentials procedures in supporting policy decisions. This is also highly useful for Land use allocation, food safety, security & environmental issues. Different types of Models & Various approaches have been used for forecasting of area, production & yield in agricultural crops. Several studies have been carried out on the univariate timeseries models known as ARIMA models. It is a very popular model for forecasting & Its popularity is due to its statistical properties as well as the known Box and Jenkins methodology. This study helps the policymakers to get an idea about the future requirements, enabling them to take appropriate measures like the selection of high yielding varieties, conducting trainings to farmers to improve cultural practices, adequate supply of inputs and use of latest technologies. The import and export of these crops can also be planned.

Materials and Methods

The secondary data on Productivity of Black gram for the period from 1971-72 to 2015-16 were collected from the Directorate of Agriculture & Food Production, Ministry of Agriculture & Farmers Empowerment, Govt of Odisha.

In case of Time Series analysis, the Box-Jenkins Methodology applies the ARMA or ARIMA Models to find out the best fit of a time series model to the past of a time series. To overcome the difficulty in describing the dynamic structure of the data by fitting Autoregressive (AR) & Moving Average (MA) models, the autoregressive moving average (ARMA) models was introduced. The Autoregressive Moving average model which include the order of differencing is known as Autoregressive Moving Average (ARIMA) models. The order of differencing is used to stationaries the data. The ARIMA model with parameter (p,d,q) is fitted by the univariate Box-Jenkins techniques (Box and Jenkins,1976). This model includes Autoregressive (AR) of order p, differencing to make stationary series of degree d & Moving average (MA) of order q.

If the time series data having constant mean & variance over time, then it is said to be stationary. First the original Productivity data of black gram plotted, then it is verified for stationarity. If the data is found to be non-stationary from the graph, then the first difference of the data is plotted and further checked for stationarity (Dash, et al. (2017) In this manner we proceed till the data become stationary. The maximum order of differencing (d) is usually

2.

By using the Box-Ljung test, the adequacy of the selected model is checked. By using Box-Ljung test of the residuls, A formal testing of the fitness of the model is also done (Ljung and Box (1978)) in following manner-

Null hypothesis is set as H_0: The errors are distributed randomly.

And the alternate hypothesis H_1: The errors are non-random.

The Box-Ljung test statistic,

$$Q = n(n+2)\sum_{k=1}^{m}\frac{r_k^2}{n-k}$$,

Here n is the number of observations &

R_k is the estimated autocorrelation of the series at lag k=1,2,….m

Where m is the number of lags being considered.

$r_k$is the estimated autocorrelation of the series at lag k = 1,2,…,m and *m* is the number of lags being considered. Here the null hypothesis (H_o) is rejected, it means the error are not independent, if

$Q \geq \chi^2_{1-\alpha,\, h}$

the null hypothesis(H_0) is accepted, it means the error are independent. if $Q < \chi^2_{1-\alpha,\, h}$

Where,$\chi^2_{1-\alpha,\, h}$ is the χ^2**(Chi-square distributions)** table value with 'h' degrees of freedom and level of significance α such that $P(\chi^2_h > \chi^2_{1-\alpha,\, h=}) = 1-\alpha$

In this case , the degree of freedom, h= (m-p-q)

Here, p & q are the numbers of AR & MA terms, respectively. The Box-Ljung test is done by the Statistical software IBM SPSS 20.0

To select the best fit model, the following model fit statistics are used-

1. RMSE-Root Mean Square Error- RMSE = $\frac{\sum_t e_t^2}{n-2}$

2. MAPE- Mean absolute percentage Error- MAPE = $\frac{\frac{\sum_t |Y_t - \hat{Y}_t|}{yt}}{n}$ ×100

The models among the selected ARIMA models have lowest value of Root mean square error & Mean absolute percentage error is considered to be the best-fit model from the respective data set.

Nearly 20% of the data at the end period is used for cross validation of the selected model, this data is not used in case of model building. In this research, the data from 1971-72 to 2006-07 are used for model building & data from 2007-08 to 2015-16 are used for the cross validation of the data. For this purpose, the percentage of forecasting error is to be calculatd.

For this, the percentage error is calculated.

% of Forecasting Error = $\left(\frac{Y-\hat{Y}}{Y}\right)$ ×100

Where, Y is the observed value of remaining eight years is the forecast values of remaining eight years Lower the value of forecasting error percentage, better is the prediction by the selected model.

The ARIMA model is used for forecasting, after its successful cross-validation. In ARIMA techniques mainly short-term forecasting is used. This is because if we go on predicting for longer periods than the error associated with the prediction will increase. So, ARIMA should be used for short term forecasting (Sarika *et al.* 2011)

Results and Discussion

Different ARIMA model is fitted to the data on productivity of Black gram for the purpose of forecasting. Data from 1971-72 to 2006-07 is used for model building. & For cross validation of the selected model, data from 2007-08 to 2015-16 are used. Forecasting is done for the years 2016-

17,2017-18 and 2018-19 by using the selected best fit ARIMA Model.

The original plot of productivity data of black gram as shown in Fig. 1(a) reveals that the data is non-stationary. Thus, the first difference on productivity data of black gram is plotted in Fig 1(b), This plot shown data the first difference data is found to be stationary.

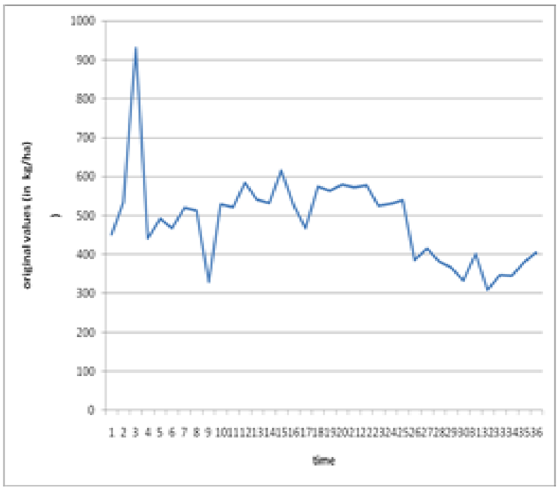

Figure 1(a) Original Values of productivity

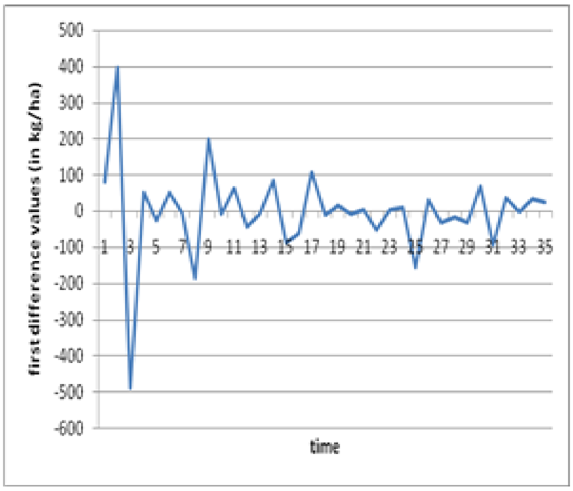

Figure 1(b) First difference values of black gram vs time of productivity of black grams time

The Autocorrelated Function(ACF) & Partial Autocorrelated Function(PACF) plot of first difference values of black gram productivity is displayed in Figure 2, which suggests that the *q* & *p* value that would be suitable for black gram productivity is *q=1* & *p=0,* Thus ARIMA(0,1,1) model is found to be the best fitted model.

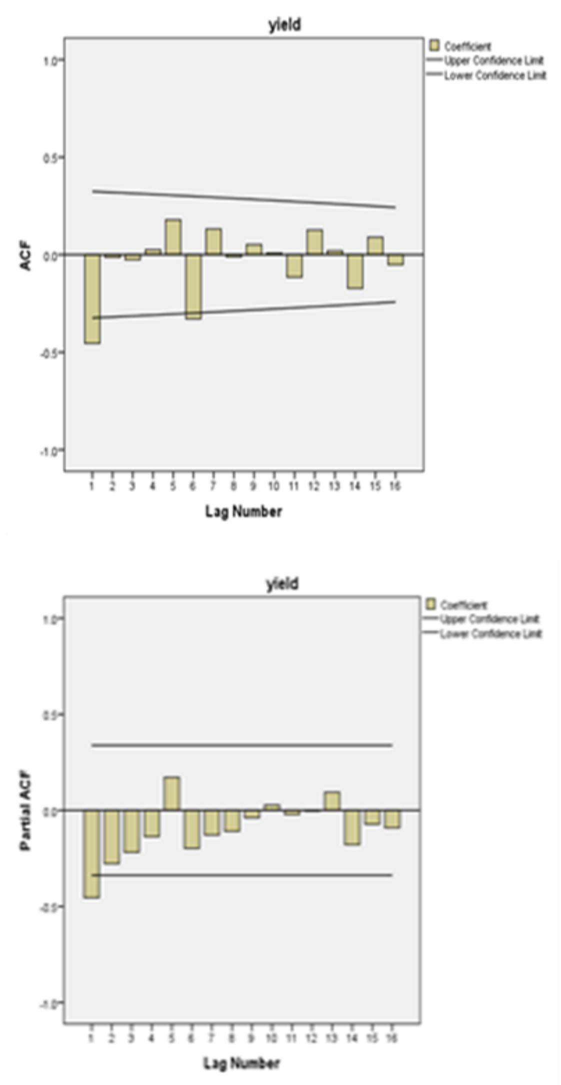

Figure 2: first difference values of ACF & PACF plot of black gram productivity

Table 1 : Autoregressive (AR) & Moving Average(MA) coefficient of the fitted ARIMA model considered for black gram productivity forecasting in Odisha

	Fitted ARIMA model	Constant (μ)	Coefficient of AR components		Coefficient of MA components	
			α_1	α_2	θ_1	θ_2
Productivity	011	0.173 (0.150)	-	-	0.972** (0.290)	-
	011-Without constant	-	-	-	0.810** (0.115)	-

**significant at 1% level of significance

The study of Table 2 shows that all the fitted model satisfy the assumptions of normality of error as they all have non-significant S-W Statistic and also all the models are found to be adequate due to non-significant of Ljung-Box Q Statistic. ARIMA(0,1,1) without constant model having lowest value of Root Mean Square Error & Mean Absolute Percentage Error, so this model is considered as best fitted Model.

Table 2: Fitted ARIMA model for black gram productivity- Model fit statistics & Residual Diagnostics

	Fitted ARIMA model	Model fit statistics		Residual	
		Root Mean Square Error	Mean Absolute Percentage Error	Ljung-Box Q Statistics	Shaoiro-Wilk(S-W) Statistics
Productivity	011	108.909	14.536	12.293	0.951
	011 (without constant)	**107.108**	**14.399**	**10.585**	**0.957**

The cross validation of the selected best fitted ARIMA model i.e. ARIMA(0,1,1,) without constant for productivity of black gram presented on the table 3 shows that the absolute percentage error are quite low, thus the selected model is Successfully cross validated.

Table 3: Cross Validation of the selected best fitted ARIMA (0,1,1) without constant model for productivity of black gram Odisha

Year	Actual value (in kg/ha) (Y)	Forecasted value (in kg/ha) ($\hat{Y}$)	Error ($Y-\hat{Y}$)	Absolute % Error $\left(\frac{\lvert Y-\hat{Y}\rvert}{Y}\right) X\,100$
2007-08	425	340.04	084.96	19.99059
2008-09	455	329.55	125.45	27.57143
2009-10	417	318.95	098.05	23.51319
2010-11	430	308.27	121.73	28.3093
2011-12	418	297.51	120.49	28.82536
2012-13	457	286.70	170.30	37.26477
2013-14	455	275.84	179.16	39.37582
2014-15	456	264.95	191.05	41.89693
2015-16	448	254.06	193.94	43.29018

The forecasted values for the black gram productivity are obtained from the best fitted ARIMA model, i.e. ARIMA(0,1,1) are presented in the following Table (Table 4), which indicates that the black gram productivity The forecasted values for the productivity of black gram are obtained from respective best fit ARIMA model are presented in table 4 shows productivity gradually decrease in their forecasted values i.e. from 2016-17 to 2018-19.

Table 4: Forecasted values (with 95% confidence limits) for productivity of black gram

	Year	Forecasted Value	L.C.L-Lower confidence Limit (95%)	U.C.L-Upper Confident Limit (95%)
Productivity (in kg/ha)	2016-17	243.16	107.62	416.97
	2017-18	232.28	98.87	405.02
	2018-19	221.45	90.36	392.89

The following Figure (Figure 3) represents the observed values & fitted values for the black gram productivity along with U.C.L & L.C.L as obtained from their best fitted ARIMA model

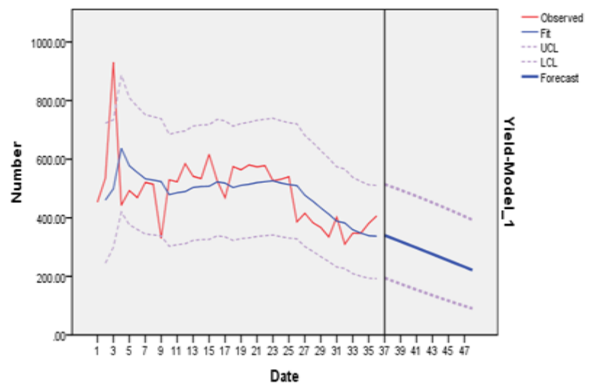

Figure 3: Observed & fitted values of Black

gram productivity along with U.C.L & L.C.L by using the Best fit ARIMA(0,1,1) without constant model

Conclusion

ARIMA(0,1,1) & ARIMA(0,1,1) without constant are the selected model for Black gram productivity of Odisha. In case of ARIMA(0,1,1), the constant is non-significant. So, ARIMA(0,1,1) without constant is also fitted. ARIMA(0,1,1) without constant model is the best fitted model due to the significant coefficient of MA(1) & lowest value of Root Mean Square Error & Mean Absolute Percentage Error. Cross validation of the best fitted model is done, after the successful cross validation results of the best fitted ARIMA (0,1,1) without constant model, lowest absolute percentage error i.e. 19.99% is found in 2007-08 & highest absolute percentage error i.e. 43.29% during the year 2015-16. After Successful forecasting by Using best fitted ARIMA (1,0,0) without constant model & From the above study it is found that the black gram productivity is expected to be decrease for the future year i.e. 2016-17,2017-18,2018-19.

References

Amin, Amanullah M, and A Akbar. 2014. "Time series modeling for forecasting wheat production of Pakistan." *The journal of Animal and Plant Sciences* 1444-1451.

Box, GEP, and GM Jenkins. 1976. "Time Series Analysis: Forecasting and Control." *Holden Day.*

Dash, A, D S Dhakre, and D Bhattacharjee. 2017. "Forecasting of food grain production in Odisha by fitting ARIMA model." *Journal of pharmacognosy and phytochemistry* 1126-1132.

Dasyam, R, B Bhattacharyya, and P Mishra. 2016. "Statistical modeling to area, production and yield of potato in west Bengal." *International Journal of Agriculture Sciences* 2782-2787.

Folarin, S B, and O S Iyiola. 2013. "Causality and invariability of autoregressive integrated moving average (arima) model." *Asian journal of current engineering and maths* 260 - 266.

Ljung, G M, and G E P Box. 1978. "On a Measure of a Lack of Fit in Time Series Models." *Biometrika* 297–303.

Odisha Agriculture Statistics. n.d. *Directorate of Agriculture and Food Production.* Bhubaneswar: Government of Odisha.

Padhan, P C. 2012. "Application of ARIMA Model for Forecasting Agricultural Productivity in India." *Journal of Agriculture & Social Sciences* 50-56.

Prabakaran, K, and C Sivapragasam. 2014. "Forecasting areas and production of rice in India using ARIMA model." *International Journal of Farm Sciences* 99-106.

Prakash, K, and B Muniyandi. 2014. "Application of ARIMA model for forecasting production of jasmine flower in Madurai district of Tamil Nadu, India." *American International Journal of Research in Humanities, Arts and Social Sciences.*

Rahman, N A F, M A Aziz, and M Mukhlesur. 2013. "Modeling on Grass Pea and Mung Bean Pulse Production in Bangladesh Using ARIMA Model." *IOSR Journal of Agriculture and Veterinary Science (IOSR-JAVS)* 20-31.

Raman, R K, T V Sathianandan, A P Sharma, and B P Mohanty. 2015. "Modeling and Modeling and Forecasting Marine Fish Production in Odisha Using Seasonal ARIMA Model." *The National Academy of Sciences* 393–397.

Sarika, Iquebal M A , and C Chattopadhyay. 2011. "Modeling and forecasting of pigeon pea (Cajanuscajan) production using autoregressive integrated moving average methodology." *Indian Journal of Agricultural Sciences* 520–523.