

Original Research Article

<https://doi.org/10.20546/ijcmas.2018.702.019>

## ARIMA Model for Forecasting of Area, Production and Productivity of Rice and Its Growth Status in Thanjavur District of Tamil Nadu, India

M. Hemavathi<sup>1\*</sup> and K. Prabakaran<sup>2</sup>

<sup>1</sup>Agricultural Statistics, Institute of Agricultural Science, Banaras Hindu University, Varanasi, Uttar Pradesh, India

<sup>2</sup>Agricultural Statistics, Agricultural College and Research Institute, Madurai, Tamil Nadu Agricultural University, Tamil Nadu, India

\*Corresponding author

### ABSTRACT

Crop acreage estimation and crop yield forecasting are two components, which are crucial for proper planning and policy making in the agriculture sector of the country. This research is a study model of forecasting area, production and productivity of rice in Thanjavur district as it is a “Rice bowl of Tamil Nadu”. Data for the period of 1990-91 to 2014-15 were analyzed by time series methods. Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF) were calculated for the data. Appropriate Box-Jenkins Auto Regressive Integrated Moving Average (ARIMA) model was fitted. Validity of the model was tested using standard statistical techniques. For forecasting area, production and productivity ARIMA (0, 1, 2), (0, 1, 1) and (0, 1, 1) model respectively were used to forecast five leading years. The results also shows area forecast for the year 2020 to be about 158.15 thousand hectare with upper and lower limit 200.85 and 122.80 thousand hectares respectively, production forecast to be about 637.05 thousand tonnes with upper and lower limit 1057.63 and 216.47 thousand tonnes respectively and productivity forecast to be about 3.79 thousand kg per ha with upper and lower limit 5.83 and 1.75 thousand kg per ha respectively. The growth pattern was examined by fitting an exponential function ( $Y = AB^t$ ). The result showed that the compound growth rate of rice was negative and non significant for area, production and yield for the study periods.

#### Keywords

Forecasting, Area, Production, Productivity, Rice, ARIMA, CGR

#### Article Info

##### Accepted:

04 January 2018

##### Available Online:

10 February 2018

### Introduction

Rice is among the three leading food crops of the world, with maize and wheat being the other two and it is one of the most important cereal crop of India occupying an area of 43.39 million hectare with an annual production of 104.32 million tonnes with an average productivity of 2404 kg/ha (2015-16).

India is the second largest producer and consumer of rice and accounts for 22.3% of global production. (Source: Annual Report 2015-16, Ministry of Agriculture and Farmers Welfare). Rice plays a vital role in the national food security and would continue to remain so because of its wider adaptability to grow under diverse ecosystems. Rice contributes 40.8% of total food grain and remains the

principal source of livelihood for more than 58% of the population.

### **Rice scenario in Tamil Nadu**

Tamil Nadu secured third position in rice productivity in India and from total gross cropped area (58.97 La ha) of the state, paddy alone cultivated in 17.75 L ha (31%) and it ranks first in area and production among other cereal crop cultivated in Tamil Nadu state (source: Policy Note 2015-16, Govt of TN), because it is the main staple food of the state with regard to human nutrition and caloric intake. Thanjavur district stands unique from time immemorial for its agricultural activities and is rightly acclaimed as the granary of south India lying in the deltaic region of Cauvery river which supports for agriculture. It is known as “Rice bowl of Tamil Nadu”, among 32 districts of Tamil Nadu, Thanjavur district rank second in area(1.77 L ha) and production (8.5 L tones) of rice after Thiruvarur district during 2014-15. The major crops cultivated in this district are paddy, pulses, gingelly, groundnut and sugarcane. Paddy is the principle crop grown in three season viz. Kuruvai, Samba and Thaladi.

In recent years the agricultural problem has become much more severe and intense. Thanjavur district has differed by deficit of irrigation. Thanjavur district agriculture has continued to be a gamble in the hands of monsoon, failure of rainfall and excessive rains and consequent floods also affected this area. Considering the importance and need of rice for daily life it's necessary to know the availability of rice grain both in present and future year which will aid in predicting demand as well as to meet demand for growing population.

Based on requirement of rice grain, we formulated the objective of the study is to develop appropriate ARIMA models for the

time series of rice crop area, production and yield in Tamil Nadu and to make five year forecast with appropriate prediction interval.

### **Materials and Methods**

This study is based on secondary data of rice crop for forecasting area, production and yield of rice in thanjavur district. The area, production and yield data for rice crop data collected for the period from 1990-91 to 2014-15 from various sources like Directorate of Economics and Statistics, Season and Crop report and Tamilnadu state website etc.

### **Forecasting of area, production and productivity using ARIMA Model**

The annual data on rice crop cultivated area, production and yield of Thanjavur district for the period from 1990-91 to 2014-15 were used for forecasting the future values using ARIMA models. The ARIMA methodology is also called as Box-Jenkins methodology. The Box-Jenkins procedure is concerned with fitting a mixed Auto Regressive Integrated Moving Average (ARIMA) model to a given set of data. The main objective in fitting this ARIMA model is to identify the stochastic process of the time series and predict the future values accurately. These methods have also been useful in many types of situation which involve the building of models for discrete time series and dynamic systems. But, this method was not good for lead times or for seasonal series with a large random component (Granger and Newbold, 1970).

Originally ARIMA models have been studied extensively by George Box and Gwilym Jenkins during 1968 and their names have frequently been used synonymously with general ARIMA process applied to time series analysis, forecasting and control. 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 model 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 reducing non-stationary series to a stationary series by taking first order differences.

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

Identification  
Estimating the parameters  
Diagnostic checking and  
Forecasting

### **Estimation of growth rates**

An exponential function of the form

$$Y = AB^t$$
$$\text{Log } Y = \log A + t \log B$$

Where,

Y = area / production / yield  
A = constant  
B = regression coefficient, and  
t = time in years

Compound growth rate (CGR) = (Antilog of B - 1) x 100 was used to calculate the compound growth rates in area, production and yield of rice.

### **Results and Discussion**

In this study, we used the data for rice crop cultivated area, production and yield for the period 1990-91 to 2014-15. As we have earlier stated that development of ARIMA model for any variable involves four steps: Identification, Estimation, Verification and Forecasting. Each of these four steps is now

explained for rice crop cultivated area, production and yield as follows.

### **Model identification**

For forecasting rice crop area, production and yield, ARIMA model estimated only after transforming the variable under forecasting into a stationary series. The stationary series is the one whose values vary over time only around a constant mean and constant variance. There are several ways to ascertain this. The most common method is to check stationarity through examining the graph or time plot of the data is non-stationary. In this case graph of  $X_t$  was stationary in mean and so no need of differencing the data.

The next step is to identify the values of p and q. For this, the autocorrelation and partial autocorrelation coefficients of various orders of  $X_t$  are computed (Table 1). The Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF) (Fig 1, 2 and 3) show that the order of p and q can at most be 0.5. We entertained eight tentative ARIMA models and chose that model which has minimum AIC (Akaike Information Criterion) and SBC (Schwartz Bayesian Criterion). The models and corresponding AIC and SBC values are given in Table 2. So the most suitable model is ARIMA (0, 1, 2) for rice area, ARIMA (0, 1, 1) for rice production and ARIMA (0, 1, 1) for rice productivity has the lowest AIC and SBC values.

### **Model estimation and verification**

Rice cultivated area, production and productivity model parameters were estimated using SPSS package. Results of estimation are reported in Table 3, Table 4 and Table 5. The model verification is concerned with checking the residual of the model to see if they contain any systematic pattern which still can be removed to improve on the chosen ARIMA.

**Table.1** Auto correlations and partial auto correlations for rice area, production and productivity

Lags	Rice Area				Rice Production				Rice Productivity			
	ACF	SE	PACF	SE	ACF	SE	PACF	SE	ACF	SE	PACF	SE
1	.433	.189	.433	.200	.300	.189	.300	.200	.247	.189	.247	.200
2	.103	.185	-.104	.200	.206	.185	.128	.200	.125	.185	.068	.200
3	.021	.181	.021	.200	.032	.181	-.067	.200	.039	.181	-.007	.200
4	-.059	.176	-.082	.200	-.064	.176	-.093	.200	-.007	.176	-.024	.200
5	-.002	.172	.074	.200	.005	.172	.061	.200	-.122	.172	-.125	.200
6	.003	.168	-.031	.200	.118	.168	.150	.200	.041	.168	.108	.200
7	.142	.163	.191	.200	.151	.163	.083	.200	.102	.163	.103	.200
8	-.098	.159	-.324	.200	-.098	.159	-.253	.200	-.050	.159	-.118	.200
9	-.137	.154	.087	.200	-.175	.154	-.170	.200	-.125	.154	-.130	.200
10	-.161	.149	-.206	.200	-.243	.149	-.080	.200	-.204	.149	-.172	.200
11	-.279	.144	-.106	.200	-.318	.144	-.167	.200	-.215	.144	-.099	.200
12	-.166	.139	-.048	.200	-.130	.139	-.002	.200	-.060	.139	.095	.200
13	-.105	.133	.006	.200	-.091	.133	-.055	.200	-.056	.133	-.059	.200
14	-.058	.128	-.121	.200	-.083	.128	-.069	.200	-.128	.128	-.176	.200
15	-.071	.122	.066	.200	-.104	.122	-.027	.200	-.064	.122	-.037	.200
16	.015	.115	.010	.200	-.144	.115	-.055	.200	-.161	.115	-.141	.200

**Table.2** AIC and SBC values for tentative ARIMA models

	ARIMA (p, d, q)	AIC	SBC
<b>1. Rice area</b>	<b>011</b>	542.720	545.076
	110	544.061	546.417
	111	540.994	544.528
	<b>012</b>	<b>540.288</b>	<b>543.822</b>
	210	544.559	548.093
	112	542.559	547.014
	211	542.338	547.050
	212	543.114	549.004
<b>1. Rice production</b>	<b>011</b>	<b>647.043</b>	<b>649.399</b>
	110	648.530	648.530
	111	649.082	652.616
	012	649.101	652.635
	210	650.203	653.737
	112	651.291	656.003
	211	650.250	654.962
	212	652.475	658.365
<b>1. Rice productivity</b>	<b>011</b>	<b>391.205</b>	<b>393.561</b>
	110	393.107	395.464
	111	393.294	396.828
	012	393.295	396.829
	210	393.969	397.504
	112	395.413	400.125
	211	395.434	400.146
	212	397.105	402.996

**Table.3** Estimates of the fitted ARIMA model for rice area

Model Fit statistics								Box- Ljung Q	
St. R-sq.	R-sq.	RMSE	MAPE	MAE	MaxAPE	MaxAE	NormalizedBIC	Statistics	Sig.
0.266	0.171	1.739E4	8.508	1.366E4	30.952	4.308E4	19.925	<b>10.733</b>	<b>0.826</b>

**Table.4** Estimates of the fitted ARIMA model for rice production

Model Fit statistics								Box- Ljung Q	
St. R-sq.	R-sq.	RMSE	MAPE	MAE	MaxAPE	MaxAE	NormalizedBIC	Statistics	Sig.
0.279	0.007	1.652E5	24.337	1.295E5	90.533	3.306E5	24.295	<b>15.957</b>	<b>0.527</b>

**Table.5** Estimates of the fitted ARIMA model for rice productivity

Model Fit statistics								Box- Ljung Q	
St. R-sq.	R-sq.	RMSE	MAPE	MAE	MaxAPE	MaxAE	Normalized BIC	Statistics	Sig.
0.303	-0.043	800.08	18.581	584.888	53.626	1.880E3	13.634	<b>13.658</b>	<b>0.691</b>

St. R-sq.-Stationary R-Square, R-Sq.- R-Square, RMSE-Root Mean Square Error, MAPE-Mean Absolute Percentage Error, MAE- Mean Absolute Error, Max APE- Maximum Absolute Percentage Error, Max AE- Maximum Absolute Error.

**Table.6** Forecasted values of rice cultivated area, production and productivity with 95% Confidence Level (CL)

Year	Forecasted values (thousand hectares)	LCL	UCL	Forecasted values (thousand tonnes)	LCL	UCL	Forecasted values (thousand kg/ha)	LCL	UCL
2015-16	164.98	131.56	204.58	644.76	304.05	985.46	3.82	2.17	5.47
2016-17	160.28	124.45	203.55	642.83	280.50	1005.16	3.81	2.06	5.57
2017-18	159.56	123.89	202.65	640.91	258.17	1023.63	3.80	1.95	5.66
2018-19	158.86	123.35	201.75	638.98	236.87	1041.08	3.80	1.85	5.75
2019-20	158.15	122.80	200.851	637.05	216.47	1057.63	3.79	1.75	5.83

LCL – Lower Confidence Level UCL – Upper Confidence Level

**Table.7** Estimate of compound growth rate of area, production and productivity of blackgram in villupuram district during 1990-91 to 2014-15

Crop	Area	Production	Productivity
Rice	-0.487 <sup>NS</sup> (0.113)	-0.965 <sup>NS</sup> (0.230)	-0.480 <sup>NS</sup> (0.460)

# Parenthesis indicates the P value, NS – Non significant.

Fig.1 Rice area ACF & PAC

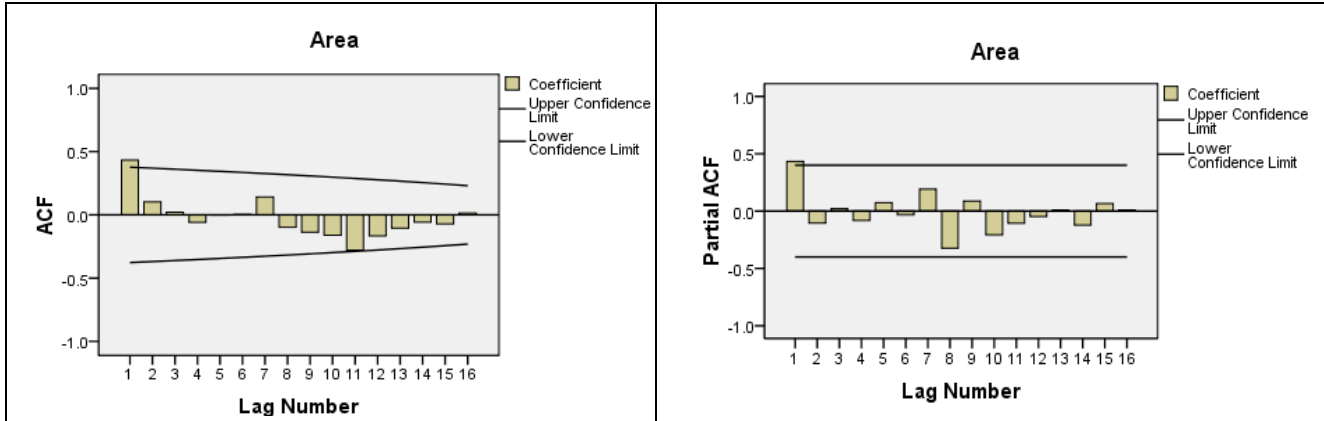


Fig.2 Rice production ACF & PACF

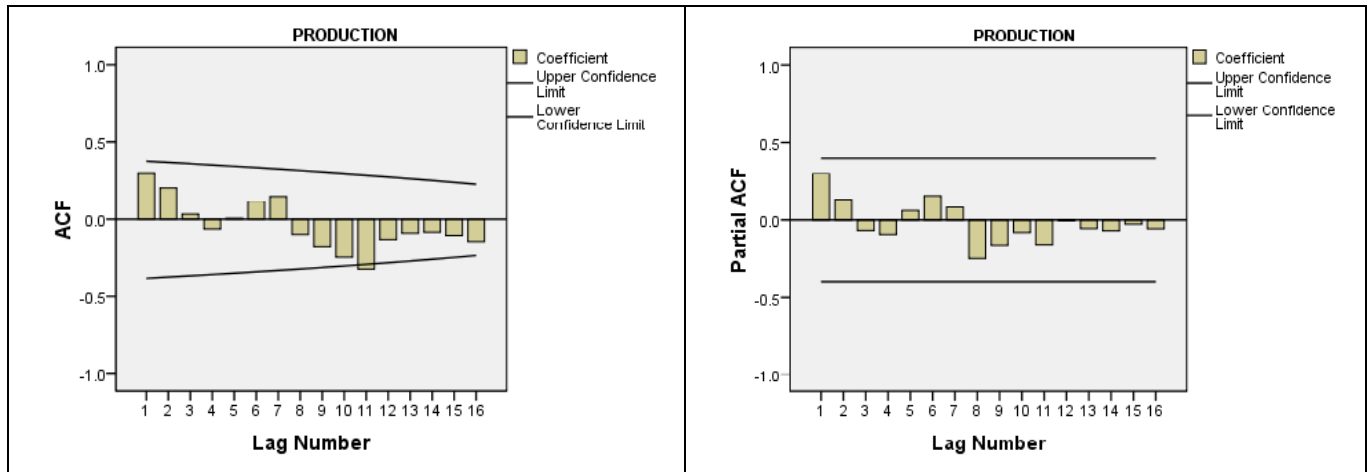
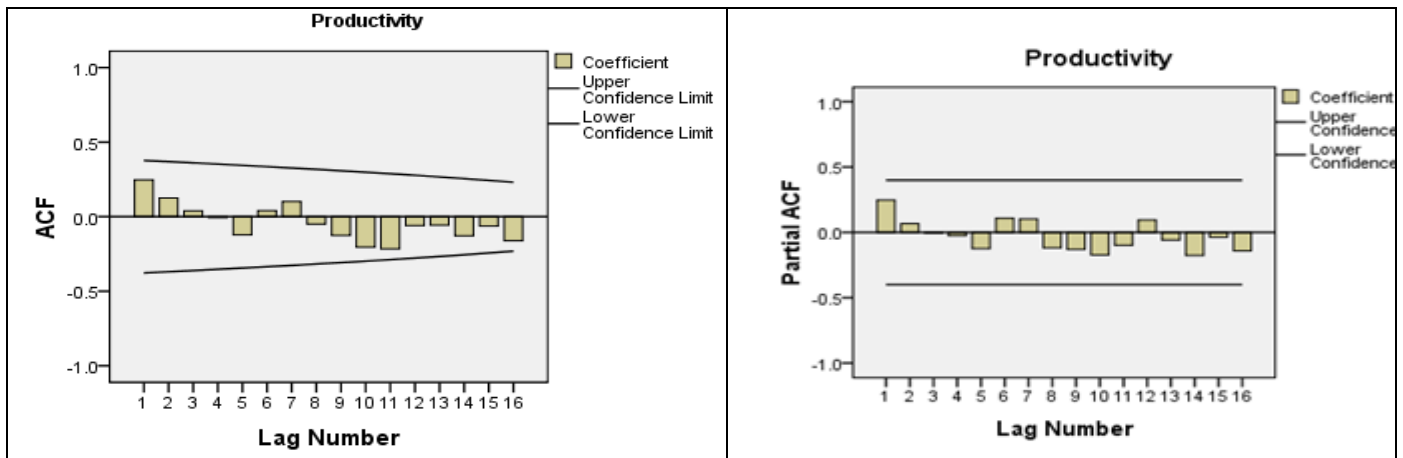
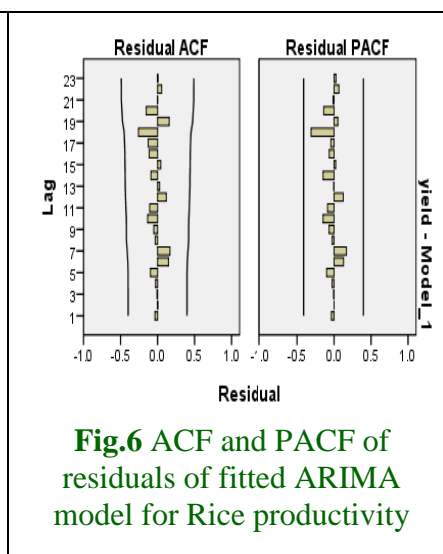
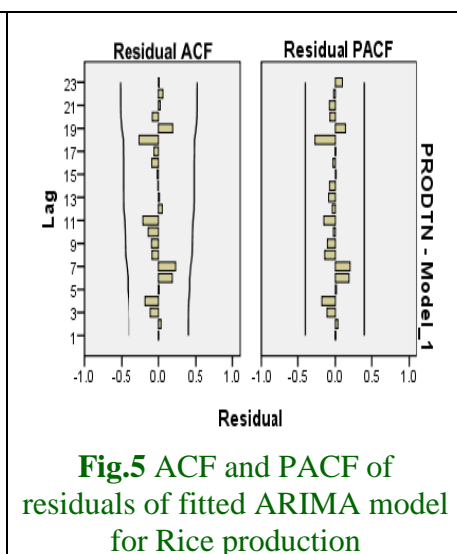
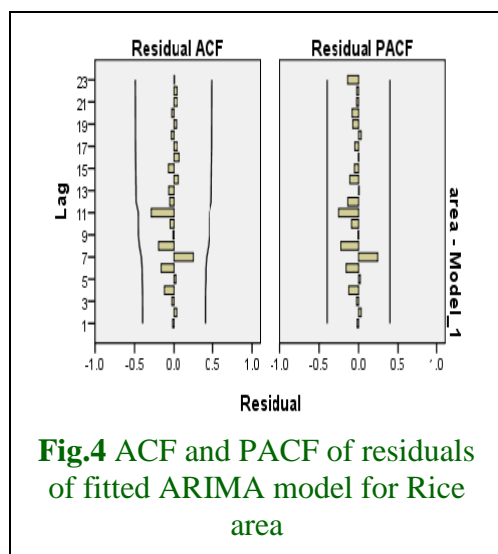


Fig.3 Rice productivity ACF & PACF





This is done through examining the auto correlations and partial auto correlations of the residuals of various orders. The ACF and PACF of the residual (Fig. 4, 5 and 6) and Box L-Jung statistic non-significant result also indicate “good fit” of the model.

### Forecasting with ARIMA model

Mainly an ARIMA model is used to produce the best weighted average forecasts for a single time series (Rahulamin and Razzaque, 2000). The accuracy of forecasts for both Ex-ante and Ex-post were tested using the following tests (Markidakis and Hibbon, 1979) such as Mean square error (MSE) and Mean Absolute percentage error (MAPE). ARIMA models are developed basically to forecast the corresponding variable. To judges the forecasting ability of the fitted ARIMA model important measure of the sample period forecasts accuracy was computed. The Mean Absolute Percentage Error (MAPE) for rice cultivated area, production and productivity turns out to be 8.503, 24.337 and 18.581 respectively. This measure indicates that the forecasting inaccuracy is low. The table 5 shows that as the year goes on, there is continuous decreasing trend in rice area due to which production and yield both resulted in decreasing trend.

### Compound growth rate

Growth performance of rice obtained from the exponential function fitted to area, production and yield are presented in Table 6. It could be seen from table that the compound growth rate of rice was negative and non-significant for area (-0.487), production (-0.965) and productivity (-0.480). The results indicate that the changing climatic condition during the study period is the main factor for declining the area, production and productivity of rice in Thanjavur district.

In this study the developed model for rice cultivated area, production and productivity was found to be ARIMA (0, 1, 2), ARIMA (1, 1, 0) and ARIMA (0, 1, 1) respectively.

From the forecast available by using the developed model, it can be seen that forecasted rice cultivated area, production and productivity decreases the next four years. The validity of the forecasted value can be checked when the data for the lead periods become available. The compound growth rate of rice had shown negative and non-significant for area, production and productivity for the study period indicate that suitable measures to be taken to enhance the rice production in Thanjavur district.

## References

- Agricultural Statistics at a glance 2014 & Annual Report 2015-2016, Government of India, Department of Agricultural, Cooperation and Farmers welfare, Ministry of Agricultural and Farmers welfare.
- Agricultural Statistic 2015. Government of India, Department of Agricultural, Cooperation and Farmers welfare, Ministry of Agricultural and Farmers welfare.
- Biswa, R. and B. Bhattacharyya. 2013."ARIMA modelling to forecast area and production of rice in West Bengal". *Journal of Crop and Weed*, 9(2):26-31.
- Box, G.E.P. and G.M. Jenkin, 1976. Time Series of Analysis, Forecasting and Control, Sam Franscico, Holden Day, California. USA.
- Granger CWJ and Newbold P 1970. Spectral methods. *Econometrica* 30: 424-438.
- Makridakis, S. and M. Hibbon. 1979. Accuracy of forecasting: An empirical investigation, *J. Roy. Stat. Soc. A*. 41(2): 97-145.
- Policy note 2015-16. Government of Tamil Nadu, Ministry of Housing, Urban Development and Agriculture.
- Rahulamin, M.D., and M.A. Razzaque. 2000. Autoregressive Integrated Moving Average Modeling for Monthly Potato Prices in Bangladesh. *Journal of Financial Management and Analysis*, 13(1): 74-80.

### How to cite this article:

Hemavathi, M. and Prabakaran, K. 2018. ARIMA Model for Forecasting of Area, Production and Productivity of Rice and Its Growth Status in Thanjavur District of Tamil Nadu, India. *Int.J.Curr.Microbiol.App.Sci*. 7(02): 149-156. doi: <https://doi.org/10.20546/ijcmas.2018.702.019>