International Journal of Current Advanced Research
ISSN: O: 2319-6475, ISSN: P: 2319-6505, Impact Factor: SJIF: 5.995
Available Online at www.journalijcar.org
Volume 7; Issue 2(F); February 2018; Page No. 10002-10006
DOI: http://dx.doi.org/10.24327/ijcar.2018.10006.1674

Research Article

ANALYSIS OF AREA UNDER GROUNDNUT CROP IN INDIA WITH ARIMA MODEL

Pushpa M. Savadatti*

Department of Economic Studies & Planning, Central University of Karnataka, Kalburgi, Karnataka

A R T I C L E   I N F O

Article History:
Received 5th November, 2017
Received in revised form 14th December, 2017
Accepted 16th January, 2018
Published online 28th February, 2018

Key words:
Groundnut, Autocorrelation, forecast, stationary

A B S T R A C T

The paper made an attempt to forecast the values for area under the ground nut crop based on the time series data collected for the period 1966-67 to 2015-16 using Auto Regressive Integrated Moving Average (ARIMA) process. The results revealed that the ARIMA (2,1,1) was the best fitted model and all the model adequacy tests like stationarity of the residuals, normality assumption, etc.,were satisfied by the fitted model. The forecasts for the period 2016-17 to 2025-26 indicated the declining trend in the area under the crop. This necessitates immediate attention from the policy makers and agricultural scientists to take appropriate measures to expand the area under the crop considering its importance in the Indian economy as an important source of edible oil for the majority of the people.

INTRODUCTION

Oilseeds occupy an important place in the Indian Economy. India is one of the largest producer of oilseeds in the world. Different oilseeds have been grown in the country namely groundnut, soybean, mustard, sunflower, sesame, rapeseed, safflower, etc. Groundnut is a very important cash crop grown in the different parts of the country. India accounts for about 18.60 percent of world area and 16.77 per cent of world production of groundnut next only to China during 2015-16 (Status Paper on Groundnut, 2017). Important states growing groundnut are Andhra Pradesh, Gujarat, Karnataka, Rajasthan and Tamil Nadu together accounted for major share of the groundnut area and production of the country. Groundnut is a highly valued food crop as it is a source of many nutrients. It is also known as Indian Almond and consumed by the common man due to its low price. Major part of the groundnut kernels is used for oil extraction also. Groundnut helps in maintaining soil fertility. Despite its importance groundnut also known as peanut is largely cultivated under rainfed conditions and area under groundnut crop is declining over a period of time. Groundnut area which accounted for 46 percent of the total oilseeds in 1966-67 declined to 18 percent in 2915-16 but that of groundnut area under irrigation increased from 5 percent to 25 percent during the same period (CMIE data base). In view of this the present work aimed at forecasting the area under the ground crop with the help of time series analysis called Auto Regressive Integrated Moving Average (ARIMA) process.

Corresponding author: **Pushpa M. Savadatti**
Department of Economic Studies & Planning, Central University of Karnataka, Kalburgi, Karnataka

The forecasting of the ground nut area will be helpful to policy makers in formulating appropriate policy measures in order to ensure increased area under the crop and enhanced production.

LITERATURE REVIEW

In literature we find number studies done on area, production and yield of groundnut crop and other crops in India and other countries of the world using various econometric techniques and some of the relevant studies have been discussed in this section. The trends in area, production and productivity of groundnut crop in Telangana was analysed based on the time series data using exponential function and results revealed that area and production showed negative trend during the study period while that of productivity exhibited positive trend and accordingly relevant implications had been derived from the results (Shruthi *et al*., 2017). Trends in area, production and productivity of groundnut in India had been analysed using orthogonal polynomial technique based on the time series data and results showed that there were negative growth rates in area but positive growth rates observed in production and yield of groundnut during the study period (Mishra, 2017). Major challenges related to production and yield of groundnut were also discussed and based on the results suitable recommendations were made. A survey had been made on area, production and productivity of groundnut at the district (Anantapur), state (Andhra Pradesh) and national level (India) during 1996-2000 to 2001-2006 and based on the evidence collected some recommendations were made for the improvement of groundnut crop (Madhusudhana, 2017). Forecasting pulse production in Kenya had been done with the help of ARIMA model based on the time series data for the period 1961 to 2012 (Esther & Magdaline, 2017). The study based on the forecast values suggested for effective policy implementation for pulse production enhancement in the country. Simple exponential smoothing, Double exponential smoothing, Damped-Trend linear exponential smoothing and ARIMA models were applied to time series data to predict crop yield in Ghana and results indicated that the ARIMA model was better compared to other models in forecasting the yield (Choudhury & Jones, 2014). The results suggested that the findings of the study were useful for insurance underwriters. A study in Bangladesh forecasted the Garlic production based on ARIMA model and selected ARIMA (0,2,1) as the best fitted model for forecasting (Hossain & Abdulla, 2015). Groundnut oil prices were forecasted applying ARIMA method and Artificial Neural Networks (ANN) and a comparative analysis of these techniques had been done (Mishra & Singh, 2013). Forecasting of food grains area, production and yield had been done at all India level using ARIMA technique and results showed that ARIMA (2,1,2), ARIMA (4,1,0) and ARIMA (3,1,3) were best fitted to the data based on the various statistical criteria (Savadatti, 2017). The results suggested for timely measures on the part of the government for enhancement of food grains' supply to meet the rising demand in the country. Another study based on the ARIMA process forecasted the production of Groundnut in Turkey and the results indicated the increasing trend in the groundnut production (Celik *et al*., 2017). It is evident from the literature review that ARIMA technique is popular in time series forecasting. The review has given the required theoretical framework for the present analysis.

RESEARCH METHODOLGY

Theoretical Frame work

The forecasting of time series data using ARIMA technique involves three stages namely Model Identification, Model Estimation and Model adequacy tests and lastly Forecasting. The first step Model Identification requires that the time series under consideration are stationary if not, they have to be made stationary through appropriate procedure. The stationarity of the series can be checked with the help of Autocorrelations (AC) and Partial Autocorrelations (PAC) of the series. Also one can resort to statistical tests like Augmented Dicky Fuller (ADF) test for testing the stationarity. The ADF test requires estimation of the following regression (Gujarati & Sangeeta, 2007)

$$\Delta Y_t = \alpha_1 + \alpha_2 t + \emptyset Y_{t-1} + \sum_{i=1}^{n} \theta_i \Delta Y_{t-i} + \varepsilon_t \quad \text{---------(1)}$$

The number of lagged differenced terms to be included in the equation is decided based on the Akaike Information Criteria (AIC) or Shrewaz Information Criteria (SIC). Based on the test results if the series are stationary one may proceed to the next stage of model identification if not then, series are made to be stationary based on the appropriate technique i.e., differencing the series then series are said to be I(d) where 'I' stands for integration and 'd' for number of times the series to be differenced in order to make them stationary. Once the series are made stationary then proceed for identification of the suitable ARIMA model; ARIMA (p,d,q) where p indicates the number of autoregressive terms, q number of moving averages terms and d the number of times the series are differenced. The AR(p) and MA(q) processes may be written as follows (Gujarati &Sangeeta, 2017)

$$(Y_t - \delta) = \alpha_1(Y_{t-1} - \delta) + \alpha_2(Y_{t-2} - \delta) + \cdots + \alpha_p(Y_{t-p} - \delta) + U_t \text{ :AR(p) process ---(2)}$$

$$Y_t = \mu + \beta_0 U_t + \beta_1 U_{t-1} + \beta_2 U_{t-2} + \cdots + \beta_q U_{t-q} \text{ : MA (q) process ------------ (3)}$$

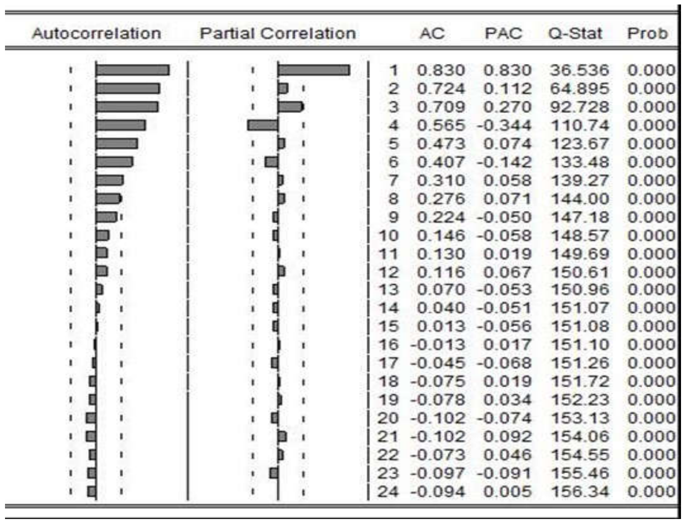
The identification of appropriate ARIMA model depends on the pattern identifiable in correlogram of ACs and PACs of stationary series.Once the appropriate model is identified next step is to estimate the model with the suitable technique especially non-linear estimation method is used to estimate the parameters of the identified model. After estimation next check whether the estimated model fits data adequately or not by applying various diagnostic checks like statistical significance of the estimated parameters, stationarity of the estimated residuals, etc. If the model is found to be satisfactory then go to next step otherwise start from step one – model identification until a suitable model is found as this is an iterative procedure. The next step is to use the fitted model for forecasting the series under consideration. Forecasting accuracy will be checked with relevant tests.

Data

The present analysis is based on the annual time series data at all India level was collected from the Centre for Monitoring Indian Economy (CMIE) data source for the period 1965-66 to 2015-16. The area under the groundnut crop is measured in '000' hectares. The data analysis was done with the help of the e-view 9 software and the results are presented below.

RESULTS AND DISCUSSIONS

The time series modelling requires the series under consideration have to be stationary. The data are converted to natural log form. The time series data collected on area under groundnut crop is tested for stationarity based on the correlogramof ACs and PACs of the series presented in Figure 1.



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.830	0.830	36.536	0.000
		2	0.724	0.112	64.895	0.000
		3	0.709	0.270	92.728	0.000
		4	0.565	-0.344	110.74	0.000
		5	0.473	0.074	123.67	0.000
		6	0.407	-0.142	133.48	0.000
		7	0.310	0.058	139.27	0.000
		8	0.276	0.071	144.00	0.000
		9	0.224	-0.050	147.18	0.000
		10	0.146	-0.058	148.57	0.000
		11	0.130	0.019	149.69	0.000
		12	0.116	0.067	150.61	0.000
		13	0.070	-0.053	150.96	0.000
		14	0.040	-0.051	151.07	0.000
		15	0.013	-0.056	151.08	0.000
		16	-0.013	0.017	151.10	0.000
		17	-0.045	-0.068	151.26	0.000
		18	-0.075	0.019	151.72	0.000
		19	-0.078	0.034	152.23	0.000
		20	-0.102	-0.074	153.13	0.000
		21	-0.102	0.092	154.06	0.000
		22	-0.073	0.046	154.55	0.000
		23	-0.097	-0.091	155.46	0.000
		24	-0.094	0.005	156.34	0.000

Figure 1 Correlogram of the series at levels

Source: Data Analysis

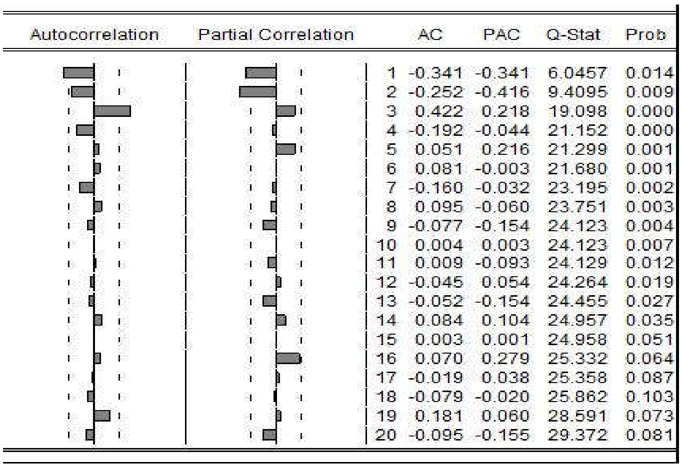
It can be seen from the above figure that autocorrelations die down slowly indicating that the series under consideration may not be stationary. So the series are transformed by differencing once and the correlogram of the transformed series is presented in Figure 2. The figure indicates that autocorrelation cuts off after first lag though the autocorrelation at third lag appears to be significant and remaining all autocorrelations are insignificant. Similarly, partial autocorrelations also cut off after two lags and remaining all are statistically insignificant signalling that the series are stationary at first difference. To get the clear picture Augmented Dicky Fuller (ADF) test has been used for testing the stationarity of the series and various specifications have been tried which are presented in Table 1.

Table 1 Stationary Test Results (ADF)

		t-Statistic	Prob.*
At levels without Constant and Linear Trend			
Augmented Dickey-Fuller test statistic		-1.664527	0.0903
Test critical values:	1% level	-2.616203	
	5% level	-1.948140	
	10% level	-1.612320	
At levels with Constant			
Augmented Dickey-Fuller test statistic		1.121620	0.9971
Test critical values:	1% level	-3.577723	
	5% level	-2.925169	
	10% level	-2.600658	
At levels with constant and Linear trend			
Augmented Dickey-Fuller test statistic		-1.213691	0.8958
Test critical values:	1% level	-4.170583	
	5% level	-3.510740	
	10% level	-3.185512	
First Difference with Constant			
Augmented Dickey-Fuller test statistic		-3.388744	0.0011
Test critical values:	1% level	-2.616203	
	5% level	-1.948140	
	10% level	-1.612320	

*MacKinnon (1996) one-sided p-values.
Source: Data Analysis

The specifications without constant and linear trend, with constant and with constant and linear trend showed that null hypothesis of unit root is not rejected as the probability is > 0.05 indicating that the series are non-stationary at levels. Then the series are differenced once and tested for stationarity and the probability is < 0.05 hence, concluded that the differenced series are stationary (Table 1). The next stage is model identification based on the ACs and PACs of the stationary series. The correlogram of ACs and PACs are presented in Figure 2. Based on the observation various specifications of the ARMA have been tried and the results of the selected ARIMA model are presented in Table 2. The confidence intervals for the estimated coefficients of the ARIMA models are presented in Table 3.



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.341	-0.341	6.0457	0.014
		2	-0.252	-0.416	9.4095	0.009
		3	0.422	0.218	19.098	0.000
		4	-0.192	-0.044	21.152	0.000
		5	0.051	0.216	21.299	0.001
		6	0.081	-0.003	21.680	0.001
		7	-0.160	-0.032	23.195	0.002
		8	0.095	-0.060	23.751	0.003
		9	-0.077	-0.154	24.123	0.004
		10	0.004	0.003	24.123	0.007
		11	0.009	-0.093	24.129	0.012
		12	-0.045	0.054	24.264	0.019
		13	-0.052	-0.154	24.455	0.027
		14	0.084	0.104	24.957	0.035
		15	0.003	0.001	24.958	0.051
		16	0.070	0.279	25.332	0.064
		17	-0.019	0.038	25.358	0.087
		18	-0.079	-0.020	25.862	0.103
		19	0.181	0.060	28.591	0.073
		20	-0.095	-0.155	29.372	0.081

Figure 2 Correlogram of the series (First difference)

Source: Data Analysis

The adequacy of the estimated model is tested based on the various tests and the same is discussed here. First criteria for model adequacy is the significance of the coefficients estimated and the results presented in Table 2 showed that the coefficients of AR(1), AR(2) and MA(1) are significant at 1% level. R^2 is also significant at 1% level though the R^2 value is low.

Table 2 ARIMA (2,1,1) Results

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.012238**	0.005873	-2.083753	0.0430
AR(1)	-0.939154***	0.142788	-6.577283	0.0000
AR(2)	-0.625376***	0.119506	-5.233021	0.0000
MA(1)	0.597257***	0.221969	2.690729	0.0100
R-squared	0.366784***			
Adjusted R-squared	0.309219			
S.E. of regression	0.062781		Akaike info criterion	-2.581166
Sum squared resid	0.173423		Schwarz criterion	-2.388123
Log likelihood	68.23857		Hannan-Quinn criter.	-2.507926
F-statistic	6.371638		Durbin-Watson stat	1.970466
Prob(F-statistic)	0.000391			

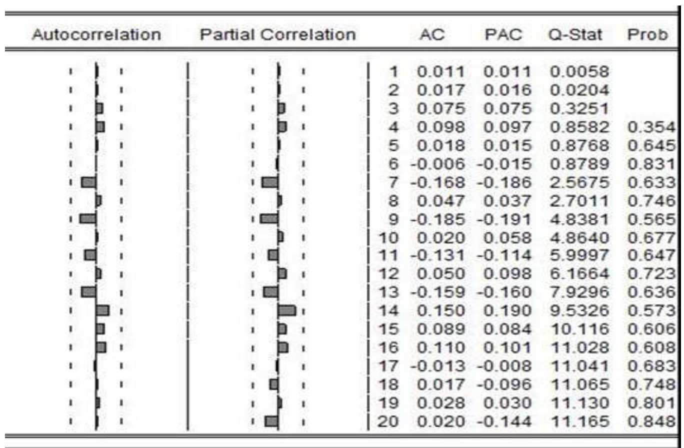
***: indicates significance at 1% level
Source: Data Analysis

Table 3: 95% Confidence Interval for Estimated Coefficients

Variable	Coefficient	Low	High
C	-0.012238	-0.024074	-0.000402
AR(1)	-0.939154	-1.226924	-0.651385
AR(2)	-0.625376	-0.866225	-0.384528
MA(1)	0.597257	0.149909	1.044605
SIGMASQ	0.003539	0.001951	0.005127

Source: Data Analysis

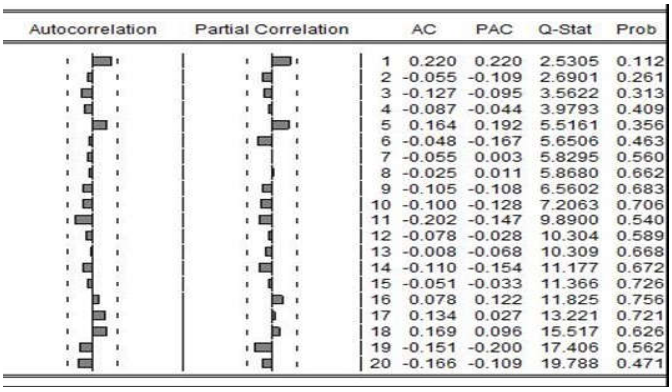
Secondly, there is need to check the residuals of the estimated ARIMA model for stationarity. The correlogram of the ACs and PACs of the estimated residuals is shown in Figure 3. None of the ACs and PACs of the residuals are significant hence, concluded that there is no pattern left in the results and residuals are stationary. Correlogram of the squared residuals are also presented in Figure 4 showing that there is no problem of heteroscedasticity among the residuals. Thirdly, the presence of multicollinearity among the explanatory variables of the model is tested with the help of Variance Inflating Factor (VIF) and the results are presented in Table 4. The VIF for all the variables is around 3<10 hence, there is no problem of multicollinearity among explanatory variables (Gujarati & Sangeetha, 2007).



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.011	0.011	0.0058	
		2	0.017	0.016	0.0204	
		3	0.075	0.075	0.3251	
		4	0.098	0.097	0.8582	0.354
		5	0.018	0.015	0.8768	0.645
		6	-0.006	-0.015	0.8789	0.831
		7	-0.168	-0.186	2.5675	0.633
		8	0.047	0.037	2.7011	0.746
		9	-0.185	-0.191	4.8381	0.565
		10	0.020	0.058	4.8640	0.677
		11	-0.131	-0.114	5.9997	0.647
		12	0.050	0.098	6.1664	0.723
		13	-0.159	-0.160	7.9296	0.636
		14	0.150	0.190	9.5326	0.573
		15	0.089	0.084	10.116	0.606
		16	0.110	0.101	11.028	0.608
		17	-0.013	-0.008	11.041	0.683
		18	0.017	-0.096	11.065	0.748
		19	0.028	0.030	11.130	0.801
		20	0.020	-0.144	11.165	0.848

Figure 3 Correlogram of the Residuals

Source: Data Analysis



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.220	0.220	2.5305	0.112
		2	-0.055	-0.109	2.6901	0.261
		3	-0.127	-0.095	3.5622	0.313
		4	-0.087	-0.044	3.9793	0.409
		5	0.164	0.192	5.5161	0.356
		6	-0.048	-0.167	5.6506	0.463
		7	-0.055	0.003	5.8295	0.560
		8	-0.025	0.011	5.8680	0.662
		9	-0.105	-0.108	6.5602	0.683
		10	-0.100	-0.128	7.2063	0.706
		11	-0.202	-0.147	9.8900	0.540
		12	-0.078	-0.028	10.304	0.589
		13	-0.008	-0.068	10.309	0.668
		14	-0.110	-0.154	11.177	0.672
		15	-0.051	-0.033	11.366	0.726
		16	0.078	0.122	11.825	0.756
		17	0.134	0.027	13.221	0.721
		18	0.169	0.096	15.517	0.626
		19	-0.151	-0.200	17.406	0.562
		20	-0.166	-0.109	19.788	0.471

Figure 4 Correlogram of the Squared Residuals

Source: Data Analysis

Table 4 Multicollinearity Test for the Estimated Model

Variable	VIF*
C	NA
AR(1)	3.019902
AR(2)	1.365157
MA(1)	3.406640

*VIF: Variance Inflating Factor

Source: Data Analysis

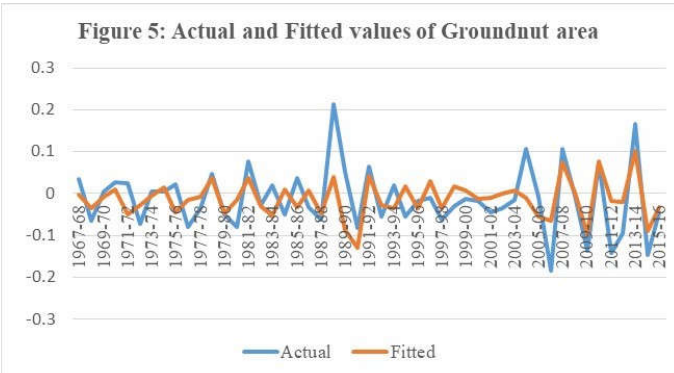
Fourthly, residuals are also tested for the normality assumptions and the results are presented in Table 5. The Jarque-Bera Probability is greater than 0.05 so the normality assumption is satisfied.

Table 5 Normality Test Results (Residuals)

Tests	Test Statistics
Skewness	0.582896
Kurosis	3.656345
Jarque-Bera /	3.654299
Probability	0.160871

Source: Data Analysis

The actual and fitted values of the area under groundnut is presented in Figure 5 below. It may be observed from the figure that the estimated model could capture the pattern present in the time series data.



Source: Data Analysis

From the above analysis it may be concluded that the estimated ARIMA model is adequate and fits the data well hence, the model is used for forecasting the values for the area for the period 2016-17 to 2026-27 and the forecasted values are presented in Table 6 below.

Table 6 Forecasted Values of Groundnut Area

Year	Area ('000' hectares)
2016-17	8.293079
2017-18	8.280841
2018-19	8.268604
2019-20	8.256366
2021-22	8.244128
2022-23	8.231891
2023-24	8.219653
2024-25	8.207416
2025-26	8.195178
2026-27	8.182941

Source: Data Analysis

The forecast accuracy of the model is tested and the same are shown in Table 7.

Table 7 Forecast Accuracy Measures

Mean Absolute Error	0.113416
Mean Absolute Percentage Error	1.300417
Theil's Inequality Coefficient	0.008348

Source: Data Analysis

The results show that forecast accuracy is fairly good as the Mean Absolute Error is low and Theil's coefficient is near to zero. The graph of the actuals and forecasted values are presented in Figure 6.

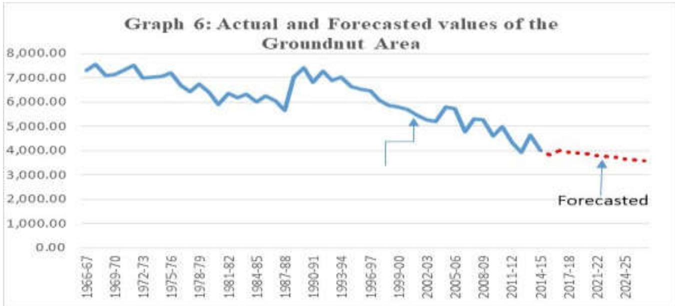


Figure 6 shows that area under the groundnut crop is declining during the study period and it further displays that the area tend to decline during the forecasted period as reflected by the forecasted values. This calls for urgent attention of the all concerned stakeholders for appropriated action in the matter.

CONCLUSION

The present study based on the time series data fitted the ARIMA model for forecasting the future values for the groundnut area in India. The best fitted model is ARIMA (2,1,1). Model fulfils most of the model adequacy tests hence, considered for forecasting the values for the period 2016-17 to 2026-27. The forecasts showed the declining trend in the area under the ground nut crop during the forecast period. It is observed that the area under groundnut is showing the declining trend during the study period 1966-67 to 2015-16 also and the same trend is likely to continue during the next decade also as evidenced by the forecasted values. This calls for an urgent attention of all the farmers, policy makers and agricultural scientists, agricultural extension workers considering the importance of the groundnut crop which is known as the Indian Almond. Groundnut is used as seeds, food and value added products like peanut butter, milk, burfi, biscuits, etc., (40% of the groundnut kernel) and 60% of the kernel is used for oil extraction (Status report on Groundnut, 2017). Ground nut has the highest share among oilseeds exports. Considering all this it is essential to take necessary measures on priority basis to enhance the area under the groundnut crop and thus production.

References

Celik, S., Karadas, K., & Eyduran, C.(2017). Forecasting the Production of Groundnut in Turkey using ARIMA Model. *The Journal of Animal and Plant Science*, 27(3), 920-928. DOI: www.thejaps.org.pk/docs/v-27-03/29

Choudary, A., & Jones, J. (2014). Crop Yield Prediction using Time Series Models. *Journal of Economic and Economic Education Research,* 15(3), 53-67. DOI: www.alliedacademies.org/articles/crop-yield-prediction-using-time-series-models.pdf

Esther, N.M., & Magdaline, N.W. (2017). ARIMA Modeling to Forecast Pulses Production in Kenya. *Asian Journal of Economics, Business and Accounting*, 2(3), 01-08. DOI: www.sciencedomain.org

Government of India Report (2017). Status Paper on Groundnut. Ministry of Agriculture and Farmers Welfare, GOI, New Delhi. DOI:http://nmoop.gov.in/Publication/StatusPaper_GNut_2017.pdf

Gujarati, D.N. and Sangeetha (2007). Basic Econometrics. McGraw Hill Publication, New Delhi.

Hossain, M.Md., & Abdulla, F. (2015). Forecasting the Garlic Production in Bangladesh by ARIMA Model. *Asian Journal of Crop Science*, 7(2), 147-153. DOI: 10.3923/ajcs.2015.147.153

Madhusudhanan, B. (2013). A Survey on Area, Production and Productivity of Groundnut Crop in India. *Journal of Economics and Finance*, 1(3), 01-07. DOI: www.iosrjournals.org

Mishra, C.M. (2017). Trends in Area, Production and Productivity of Groundnut in India: Issues and Challenges. *Journal of Research in Agriculture and Animal Science*, 4(7), 01-06. DOI:www.questjournals.org

Savadatti, P.M.(2017). Forecasting of Area, Production, and Productivity of Food Grains in India: Application of ARIMA Model. *Arthshastra Indian Journal of Economics & Research*. 6(6), 7-22.

Shruti, G., Rao, B.D., LatikaDevi,Y., & Masih, J. (2017). Analysis of Area, Production and Productivity of Groundnut Crop in Telengana. *Agricultural Scientist Digest*, 37(2), 151-153. DOI: 10.18805/asd.v37i2.7993

How to cite this article:

Pushpa M. Savadatti (2018) 'Analysis of Area Under Groundnut Crop in India with Arima Model', *International Journal of Current Advanced Research,* 07(2), pp. 10002-10006. DOI: http://dx.doi.org/10.24327/ijcar.2018.10006.1674

****