
Forecasting pigeon pea (Arhar) production in Karnataka-A Time series analysis

*Prof.Pushpa Savadatti

INTRODUCTION

Pulses are considered as very important for the Indian economy as they are contributing significantly to the sustainability of the sector being a legume crop and nutritionally also very important as a vital source of vegetable protein to the millions of the urban and rural Indians alike. India is one of the major producer of pulses in the world. India tops the list in case of production of Pigeonpea and Chickpea which are important pulse crops grown in the country. Karnataka is an important state contributing to the production of total pulses at all India level. It is also an important contributor to the Pigeonpea production in the country. But when we look at the production data at the state level it is observed that the production of Pigeonpea is oscillating over a period of time making the cultivation of the crop riskier. Though there are number of programmes under various names (Integrated Scheme of Oilseeds, Pulses, Oil Palm and Maize, ISOPOM; The National Food Security Mission, NFSM, etc.) have been implemented by the centre and state to enhance the production of Pigeonpea in particular and total pulses in general still not able to reach required level of production. The demand for Pigeonpea is increasing continuously as it is the main food item in the consumption basket of the mass in the state. Given the importance of the crop, forecasting the production of Pigeonpea would be of great importance to the policy makers, government and farmers to make appropriate decisions so that the Pigeonpea production will be enhanced in the near future at much higher level.

LITERATURE REVIEW

There is rich literature available on pulses on one hand and use of ARIMA process in forecasting the production of various agricultural commodities on the other. Some of the important works which were of value in building this paper have been discussed in this section. PriyaRampal (2017) in her paper had given the comprehensive picture of the pulse scenario in the country as a whole and also at the state level discussing supply side, demand aspect, prices, government policies over a period of time and made policy suggestions based on the analysis. Reddy, Bantilanand Geetha Mohan (2013) in their ICRISAT policy brief about pulses analysed the different aspects of pulse scenario in the country and came out with the suggestion that agricultural scientists had to focus on the region specific and crop specific research to invent the suitable cultivation practices, harvesting mechanisms, seeds development, etc., to enhance the pulse production in the country which is required at present. Singh, Shahi and Singh (2017) in their work on pulses described the situation of pulses in India in general and in Bihar in particular focussing on various aspects and made very important suggestions for increasing the pulse production through increased yield in Bihar state based on the analysis. Esther&Magadline (2017) forecasted pulse production in Kenya using univariate time series analysis ARIMA and results hinted at decreasing pulse production in Kenya during next decade hence, authors recommended that the government should take enough measures to increase the pulse production and thus ensure food security in the country. Another study by Prabakaran&Sivapragasam (2014) projected the area and production of rice in India with the help of Box-Jenkins ARIMA technique and best fitted model was ARIMA (1,1,1). The model adequacy was tested with the help of suitable tests. A study by Badmus&Ariyo (2011) forecasted area under maize and production of maize in Nigeria with the help of ARIMA technique based on the time series data for the period 1970 to 2005 and concluded that area under the maize crop could be increased with land reclamation and conservation measures if implemented effectively. There are number of studies used ARIMA technique to forecast the agricultural commodities (Arivarsi&Ganesan, 2015; Biswas, Dhaliwal, Singh & Sandhu, 2014; Savadatti, 2017). Based on the literature review the required theoretical framework is designed and presented below.

METHODOLOGY

Data Description

The present analysis aimed at forecasting the Pigeonpea production in Karnataka State. For this purpose, the secondary time series data dating from 1966-67 to 2015-16 were collected from the Centre for

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

Monitoring Indian Economy (CMIE) States of India data base. The frequency of the time series is annual and data refers to production of Pigeonpea in Karnataka State. The production is measured in thousand tons. Eviews 9 software is used for the present analysis.

THEORETICAL FARAMEWORK

The paper intends to project the production of Pigeonpeain Karnataka which is one of the important states contributing to the Arhar production at all India level. For this objective the time series technique known as Auto Regressive Integrated Moving Average (ARIMA) technique is adopted. In the present analysis univariate ARIMA procedure is used. This technique includes three steps mainly identification of the model, estimation of the parameters and diagnostic checking of the estimated model. The first step is to identify the appropriate ARIMA (p,d,q) model suitable for estimating the parameters. Here ‘p’ refers to the order of Autoregressive term. The Autoregressive processof order p is defined as (Gujarati &Sangeeta, 2007;p. 858)

$$(Y_t - \delta) = \gamma_1 (Y_{t-1} - \delta) + \gamma_2 (Y_{t-2} - \delta) + \dots + \gamma_p (Y_{t-p} - \delta) + U_t \text{ --- (1)}$$

and ‘q’ refers to the order of Moving average terms. Moving Average process of order q is written as (Gujarati &Sangeeta, 2007; p.859)

$$Y_t = \mu + \theta_0 \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \text{ --- (2)}$$

‘d’ refers to the number of differencing of the time series to make them stationary and it is also known as the order of integration I(d). During first step it is essential to make sure series under consideration are stationary which is characterised by the time invariant mean, variance and covariance of the data. For this Augmented Dicky Fuller (ADF) test is usedwhich requires estimating the following equation (Gujarati &Sangeetha, 2007; p. 836)

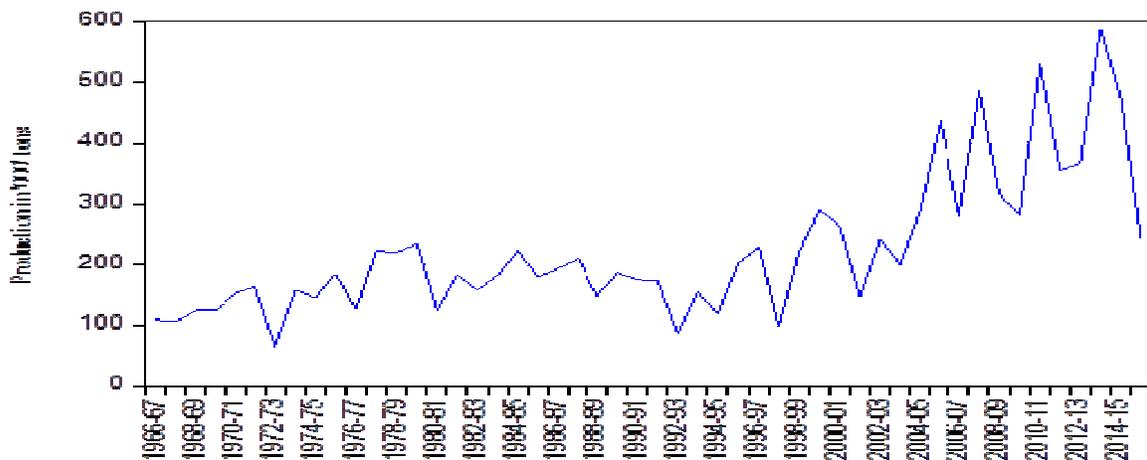
$$\Delta Y_t = \alpha_1 + \alpha_2 t + \beta Y_{t-1} + \sum_{i=1}^n \phi_i \Delta Y_{t-1} + \varepsilon_t \text{ --- (3)}$$

The ‘d’ will be identified with the help of ADF test and ‘p’ and ‘q’ will be identified with Autocorrelation (ACF) and Partial Autocorrelations (PACF) functions of the series. The next stage is to estimate the identified model with the help of Maximum Likelihood Method (MLM) and last step is to check for the adequacy of the estimated model. The model adequacy is based on checking for white noise test of the residuals. Forecast accuracy are done with the help of the Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and Theil’s inequality coefficient.

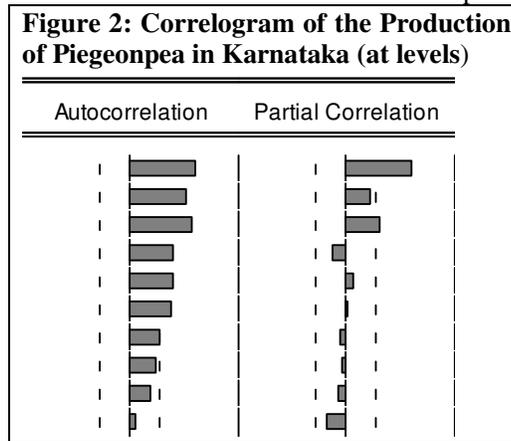
RESULTS AND DISUCSSION

In this sections results are presented. To understand the status of the production of Pigeonpea in Karnataka State graph of the production series for the study period are presented Figure 1. It may be seen from the figure that production of Pigeonpea showed an increasing trend during the study period but with lots of oscillations. There is sharp decline in production during 2015-16. This oscillation in output clearly indicates that most of the cultivation under the crop is rainfed. Hence, cultivation of Pigeonpea has become riskier. The production series are put to stationarity test using correlogram of the series which depicts the Autocorrelation and Partial Autocorrelation of the series.

Figure 1: Production of Arhar (Pigeonpea) in Karnataka



The plot of the correlogram of the series are displayed in Figure 2. It may be observed from the Figure that ACF dies down indicating the presence of the significant autocorrelations among the series. Hence, to confirm the results the ADF test is used and the results of the same are presented in Table 1.



The various specifications like equations without constant and trend, with constant only and with constant and trend have been tried and the results displayed that (Table 1) in all the three cases the ADF test statistics is less than the critical values at 1%, 5% and 10% level of significances respectively indicating that the series are not stationary this is in confirmation with the results of correlogram of the series presented in Figure 2. We need stationary series for further analysis. The techniques accepted for making series stationary is to difference the non-stationary series and accordingly we differenced the production series once and then checked for stationarity with the help of the correlogram plot and ADF test. The correlogram graph is presented in Figure 3 and ADF test results for differenced production series in Table 2. Figure 2 shows that ACF and PACF cut off after first and second lag respectively though ACF at 3rd lag is significant representing that the differenced production series are stationary.

Table 1: Augmented Dickey Fuller Stationary Test Results (at levels)			
Without constant and Trend			
		t-statistics	Probability
Augmented Dickey-Fuller test statistic		1.087206	0.9256
Test critical values:	1% level	-2.615093	
	5% level	-1.947975	
	10% level	-1.612408	
With constant and no Trend			
		t-statistics	Probability*
Augmented Dickey-Fuller test statistic		-0.446668	0.8922
Test critical values:	1% level	-3.577723	
	5% level	-2.925169	
	10% level	-2.600658	
With constant and Trend			
		t-statistics	Probability
Augmented Dickey-Fuller test statistic		-5.412575	0.0003
Test critical values:	1% level	-4.156734	
	5% level	-3.504330	
	10% level	-3.181826	

*MacKinnon (1996) one-sided p-values.

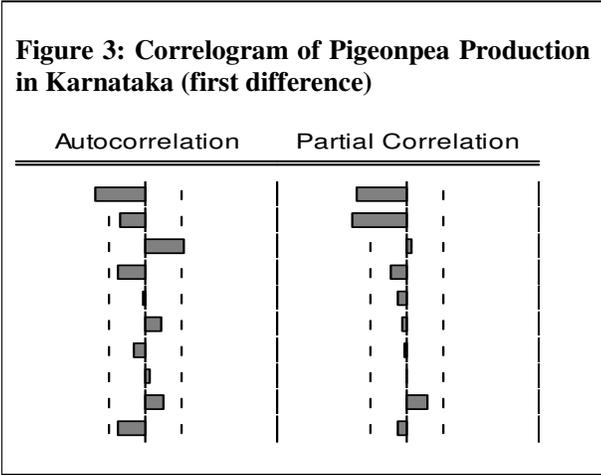


Table 2: Augmented Dickey Fuller Stationary Test Results (1st diff)

Wit constant		
	t-statistics	Probability*
Augmented Dickey-Fuller test statistic	-9.900213	0.0000
Test critical values:	1% level	-2.615093
	5% level	-1.947975
	10% level	-1.612408
*MacKinnon (1996) one-sided p-values.		

Table 2 presented the ADF test results for the first differenced series and calculated test statistics is greater than the critical value at 1% level of significance, amply signifying that the differenced series are stationary. The correlogram of the stationary series are used for deciding AR and MA terms i.e., ‘p’ and ‘q’ terms respectively as the value for ‘d’ is one in the ARIMA model. The various specifications like ARIMA (3,1,2), ARIMA (3,1,1), ARIMA (2,1,2), ARIMA (2,1,1), ARIM (1,1,1), ARIMA (2,1,0), ARIMA (0,1,2), and ARIMA (0,1,2) have been tried based on the significance of the AC and PAC in the correlogram of the differenced series and the model best fitted has been selected and the results are presented in Table 3.

Table 3: Results of ARIMA Model:ARIMA (2,1,0)

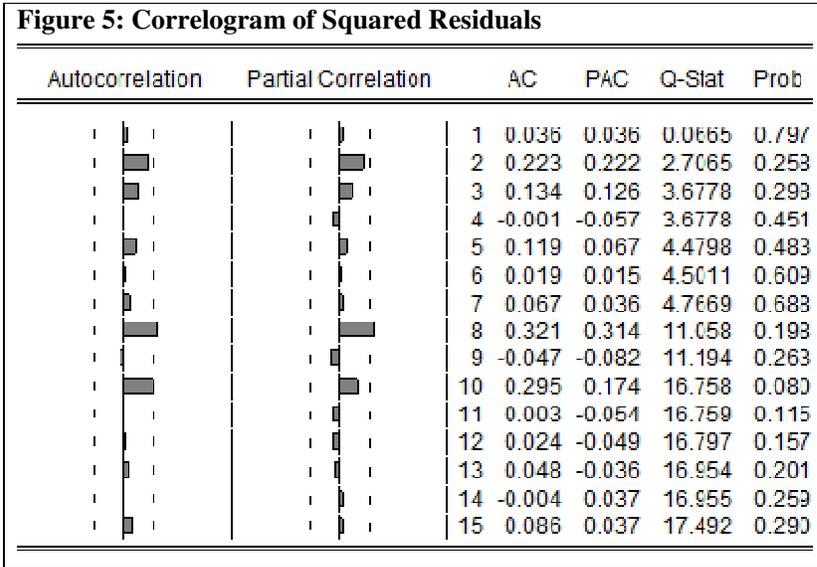
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.260510	4.371399	1.432153	0.1590
AR(1)	-0.794942*	0.155152	-5.123628	0.0000
AR(2)	-0.637206*	0.130731	-4.874166	0.0000
R-squared	0.462372			
Adjusted R-squared	0.426530			
S.E. of regression	75.38734	Akaike info criterion		11.58801
Sum squared resid	255746.3	Schwarz criterion		11.74245
Log likelihood	-279.9063	Hannan-Quinn criter.		11.64661
F-statistic	12.90033	Durbin-Watson stat		1.973623
Prob(F-statistic)	0.000003			

*indicate significance at 1% level

The estimated model has been checked for its adequacy as it is the second step of the ARIMA process. The estimated coefficients of the AR (1) and AR (2) terms are significant at 1 % level as their probabilities are less than 0.01 percent. The model fairly explains the changes in the dependent variable as R² is 0.4265 and F-value is quite high signalling that the R² is significant at 1 percent level. The D-W statistics is around 2 (1.9736) means there is no problem of auto-correlation in the residuals of the estimated model. Further the residuals are tested for the presence of any pattern and the correlograms of residuals and squared residuals are presented in Figures 4 and 5 respectively.

Figure 4: Correlogram of residuals

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.062	-0.062	0.2024	
		2	-0.080	-0.084	0.5431	
		3	-0.145	-0.158	1.6928	0.193
		4	-0.066	-0.099	1.9348	0.380
		5	-0.001	-0.043	1.9348	0.586
		6	0.092	0.052	2.4312	0.657
		7	0.159	0.151	3.9321	0.559
		8	-0.052	-0.021	4.0996	0.663
		9	0.129	0.184	5.1317	0.644
		10	-0.103	-0.027	5.8179	0.668
		11	-0.061	-0.031	6.0619	0.734
		12	0.037	0.055	6.1525	0.802
		13	-0.168	-0.223	8.1015	0.704
		14	0.139	0.089	9.4721	0.662
		15	-0.148	-0.222	11.080	0.604

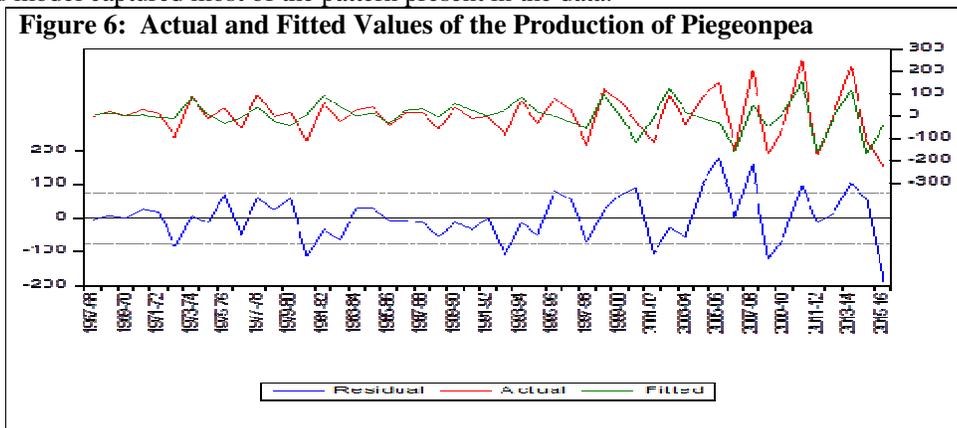


It may be observed from the figures 4 and 5 that none of the autocorrelations and partial autocorrelations in residuals and squared residuals are significant revealing that no pattern is left in the residuals as the Q-statistics probability is greater than 0.05 specifying that all the AC and PAC are statistically insignificant. Hence, it may be inferred that the estimated ARIMA model has been able to capture the pattern present in the time series data fairly well. Next residuals are examined for the normality assumption and results are presented in Table 4.

Table 4: Normality Tests of Residuals

Test	Value
Skewness	-0.012064
Kurtosis	3.353144
Jarque-Bera (JB)	0.255807
JB Probability	0.879938

Skewness, Kurtosis and Jarque-Bera (JB) test results point to the fact that the residuals are normally distributed as the JB probability is greater than 0.05 hence, not able to reject the null of hypothesis of normality. So, the normality assumption of the residuals is fulfilled. The actual and fitted values of the production alongwith the residuals are presented in Figure 6. It may be seen from the figure that the estimated model captured most of the pattern present in the data.



Based on the above tests it may be concluded that the model fitted data well hence, used for forecasting the production of Pigeonpea for next ten years and the forecasted values and actual values of the production are presented in Table 5 and Figure 7 respectively.

Table 5: Forecast Values for Production of Pigeonpea in Karnataka	
Year	Forecasts ('000' tons)
2016-17	420.4684
2017-18	426.7288
2018-19	432.9893
2019-20	439.2499
2020-21	445.5104
2021-22	451.7708
2022-23	458.0314
2023-24	464.2919
2024-25	470.5524
2025-26	476.8129

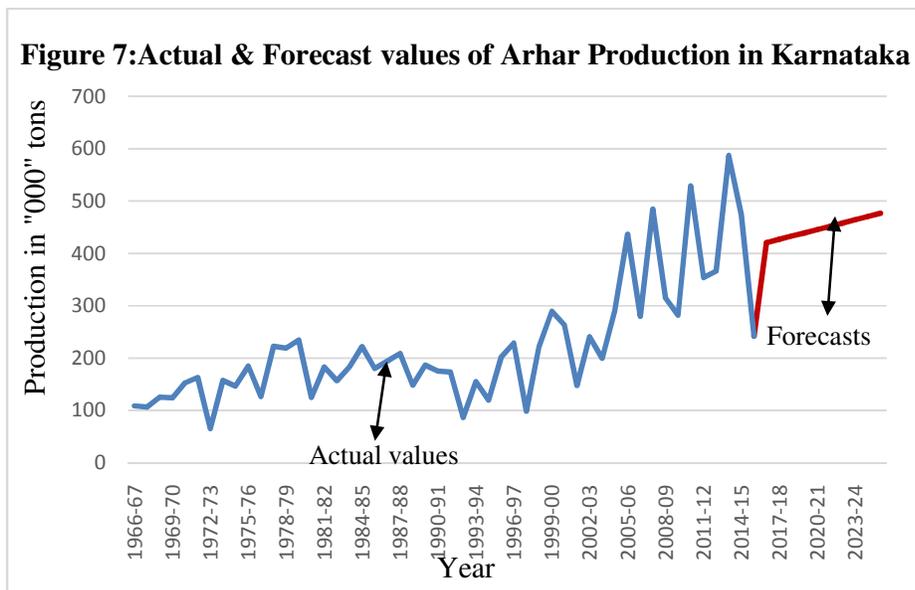


Table 5 depicts the forecast values for the production of Pigeonpea during 2016-17 to 2025-26 indicating increasing trend in the values. The production of Pigeonpea increases from 420 thousand tons in 2016-17 to around 476 thousand tons during 2025-26. The figure 7 portrays the actuals and forecast values which shows an increasing trend but at a lower level. It may be seen from the Figure that there is lot of oscillation in the production of Pigeonpea over a study period revealing the uncertainty involved in it. The forecast accuracy measures of the model have been shown in Table 6.

Table 5: Forecast Accuracy Measures	
RMSE	91.37752
MAE	74.28754

MAPE	41.63630
Theil's Inequality Coefficient	0.169051

The Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and Theil's inequality coefficient values are presented and it may be seen that the Theil's inequality coefficient is near to zero indicating that the predictive performance of the estimated ARIMA model is reasonably good (Pindyck&Rubinfeld, 1998).

CONCLUSIONS

In the present paper an effort has been made to forecast the Pigeonpea production in Karnataka state which is an important state in contributing to the pulse production of the country. The popular time series analysis known as ARIMA analysis has been adopted to forecast the values for the production series. The production series are integrated of order one i.e., $I(1)$. Based on the correlogram of the stationary time series data various specifications of ARIMA have been tried with the help of suitable econometric techniques and the best fitted model ARIMA (2,1,0) was selected for forecasting. The estimated ARIMA model satisfied most of the model adequacy tests. The forecasts for production of Pigeonpea suggest that the production will increase during next ten years though at a lower level. The lot of fluctuations are observed in the production series during the study period revealing the risk involved in the production of Pigeonpea in Karnataka State where Pigeonpea is mostly cultivated under dryland farming. Large part of the cultivation under the crop depends on rainfall which is erratic hence, resulting into unexpected oscillations in the production and thus making the crop cultivation riskier and hence, discouraging the farmers to allocate more land to Pigeonpea. This calls more comprehensive efforts on the part of the state government to implement the pulse programmes effectively and efficiently to cover the risk involved in it and to reduce the risk also. The measures are required very urgently considering the importance of the Pulse crop from nutritional and sustainable agriculture point of view.

REFERENCES

1. Arivarasi, R. &MadhaviGanesan (2015). Time Series Analysis of Vegetable Production & Forecasting using ARIMA Model. Asian Journal of Science & Technology, 6(10), 1844-1848. DOI: <http://www.journallist.com>.
2. Biswas, B., Dhaliwal, L.K., Singh, S.P. & Sandhu, S.K. (2014). Forecasting wheat production using ARIMA model in Punjab. International Journal of Agricultural Sciences, 10(1), 158-161.
3. Esther, M.N. &Magdaline, N.W. (2017). ARIMA Modeling to Forecast Pulses Production in Keyna. Asian Journal of Economics, Business and Accounting, 2(3), 1-8.
4. DOI: www.sciencedomain.org
5. Gujarati, D.N. and Sangeetha (2007). Basic Econometrics. Tata McGraw –Hill Publishing Company limited, New Delhi.
6. Oludare, A. (2011). Forecasting Cultivation Area and Production of Maize in Nigeria using ARIMA model. Asian Journal of Agricultural Sciences. 3(3), 171-176.
7. Prabakaran, K. &Sivapragasam, C. (2014). Forecasting areas and production of rice in India using ARIMA model. International Journal of Farm Sciences. 4(1), 99-106.
8. Pindyck, R.S. &Rubinfeld, D.L. (1998). Econometric Models and Economic Forecasts. McGraw Hill International Edition.
9. Rampal, P. (2017). Situational Analysis of Pulse Production and Consumption in India. Lansa Working paper series, Vol (20). DOI:<http://59.160.153.185/library/sites/default/files/LANSA%20Working%20paper%2020.pdf>
10. Reddy, A.A., Bantilan, MCS, Geetha Mohan (2013). ICRISAT Policy Brief 2013, Research Program on Green Legumes. DOI: http://oar.icrisat.org/6812/1/26_Policy_BriefIndia%20_2013.pdf
11. Savadatti, P.M. (2017). Trends and Forecasting Analysis of Area, Production and Productivity of Total Pulses in India. International Journal of Economics and Development, 5(12).
12. Singh, P., BrajeshShahi& Singh, K.M. (2017). Trends of Pulse Production, Consumption & Imports in India: Current Scenario & Strategies. MPRAC Paper No: 81589. DOI: <http://apra.ub.uni-muenchen.de/81589/>.

13. Sarwade W.K. ,: "Conceptual Development Of Supply Chain Management In Indian Retail Organised Market", International Journal of Management and Applied Science (IJMAS), Volume-2,Issue-2,pp 182-186 ,2016

14. Sarwade, Dr. W.K., Industrialization, Vision 2020 and Economic Development of Aurangabad Region of Maharashtra State (December 7, 2015). 12th International Conference on Business Management (ICBM) 2015. Available at SSRN: <https://ssrn.com/abstract=2706351> or <http://dx.doi.org/10.2139/ssrn.2706351>

15. Professor Sarwade W. K., Miss. Tandale Bhagyashri Jagannath "A Study of Economic Reforms and Performance of Private Insurance Sector in Marathwada Region" Sumedha Journal of Management Year : 2017, Volume : 6, Issue : 1

####