Journal of Crop and Weed, 10(2):147-154(2014)

Time series modeling and forecasting of pulses production in India

VISHWAJITH K.P, B S DHEKALE, P K SAHU, P MISHRA AND MD NOMAN

Dept. of Agricultural Statistics
Bidhan Chandra Krishi Viswavidyalaya,
Mohanpur, Nadia, West Bengal – 741252

Received: 22-07-2014, Revised: 15-09-2014, Accepted: 25-09-2014

ABSTRACT

Pulses are known as poor man's meat as these are comparatively cheaper sources of protein in balancing human diet. In a populous developing country like India, production of pulses play pivotal role in nutritional security of the country. Production of pulses depends on many production factors like rainfall, fertilizer etc. and also on area under crops and its productivity. Analysis of production behaviour, modelling and forecasting of production taking all these factors in to consideration play vital role in human nutritional security. In this paper attempt has been made to analyse and forecast the production scenario of pulses in major pulses growing states of India. Autoregressive Integrated Moving Average (ARIMA) methodology has been used to model and forecast the behaviour of pulses production with and without inclusion of the above factors of production. In most of the cases, inclusion of the factors of production in the model outperformed the simple ARIMA modelling. To take care of conditional variances, use of Generalised Autoregressive Conditional Heteroscedastic (GARCH) models are found in literature; the method has also been used in this study. From the forecasted value it is clear that among the Indian states, Madhya Pradesh has to play a major role in augmenting pulses production in India with its estimated share of 3661 thousand tons out of all India production of 14360 thousand tons in 2015. Comparative analysis reveals that uniform superiority of neither the ARIMA model nor the GARCH could be established in modeling and forecasting the production behavior of pulses in India.

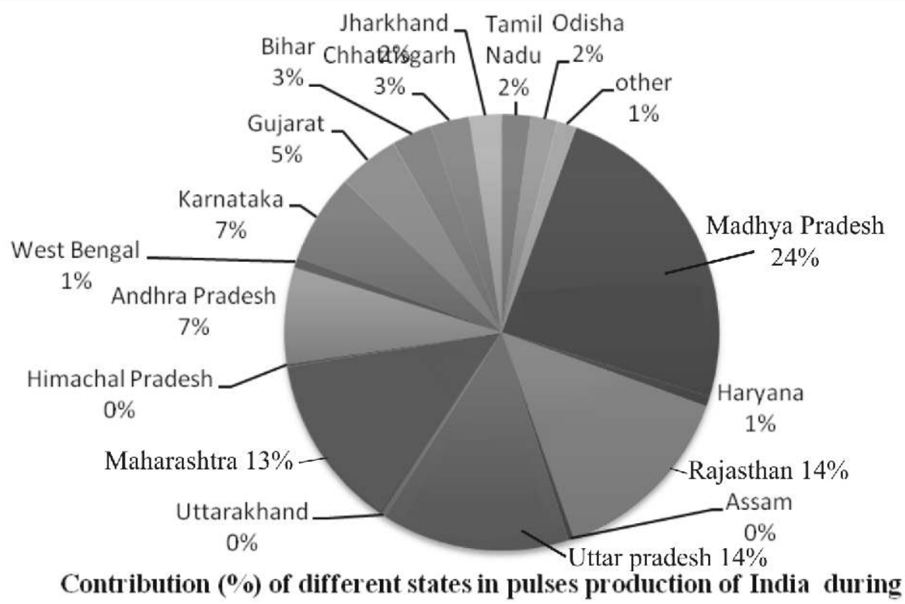
***Keywords*: ARIMA, forecasting, GARCH, production.**

Pulses, the food legumes, have been grown since millennia and have been a vital ingredient of the human diet in India; as such has long been considered as the poor man's only source of protein. Pulses are one of the important segments of human diet in Indian sub continent along with cereals and oilseeds. The split grains of pulses, called *dal* are excellent source of high quality protein, essential amino and fatty acids, fibers, minerals and vitamins. These crops improve soil health by enriching nitrogen status, long-term fertility and sustainability of the cropping systems. It meets up to 80% of its nitrogen requirement from symbiotic nitrogen fixation from air and leaves behind substantial amount of residual nitrogen and organic matter for subsequent crops. The water requirement of pulses is about one-fifth of the requirement of cereals, thus effectively save available precious irrigation water (Reddy, 2009). India is the largest producer of pulses in the world, with 24.3% share in the global production (Anon., 2011). During the same year the major producers of pulses in the country are Madhya Pradesh (4.1 million tons), Uttar Pradesh (2.4 million tons), Maharashtra (2.2million tons), Rajasthan (2.4million tons), Andhra Pradesh (1.2million tons) followed by Karnataka (1.1million tons); these states together share about 69% of total pulse production while remaining 31% is contributed by Gujarat, Chhattisgarh, Bihar, Orissa and Jharkhand etc in year 2011. In-spite of the above, growths of pulses production in different states of India are poor; fig. 1 demonstrates the contribution of different states in all India pulses production during the year 2011. The skyrocketing prices of pulses since 2008 can be attributed to almost stagnant production leading to a decline in per capita availability of pulses. The pulses are generally grown in post monsoon period and are prone to losses due to drought stress if there is scarcity of rain. During 2009-10, India imported 3.5 million tons of pulses from the countries like Australia, Canada, and Myanmar as pulse production was lowered due to drought and which was unable to fulfill demand. Presently about 25 to 26 million hectares of land is under pulses cultivation in India producing about 17 million tons of pulses annually. Still, to meet the demand, about 2-3 million tons of pulses need to be imported every year. The yield (around 700 kg a hectare) is less than the global average and the per capita availability, one-fifth lower than what nutritionists recommend (IIPR, 2013).

As pulse production is affected severely by rainfall and other factors a need arise to forecast pulse production in advance for effective planning. Auto Regressive Integrated Moving Average (ARIMA) is the most general class of models for forecasting a time series. ARIMA model was introduced by Box and Jenkins in 1976 for forecasting variables. Sher and Ahmad (2008) used this method in Pakistan to forecast

Email: vishwajithkp@gmail.com

wheat production for 2015 using time series for 1979-2006. Yaseen, *et al.* (2006) studied and forecasted the cultivated area and production of sugarcane in Pakistan with the help of the Autoregressive Integrated Moving Average (ARIMA) model using the data for 1947-2002. Hossain *et al.* (2006) studied three types of forecasts, namely, historical, ex-post and ex-ante, using the world famous Box-Jenkins time series models for motor, mash and mung prices in Bangladesh. Though ARIMA models have got wide application in modelling time series data, this is being criticised for its assumption of linearity, and homoscedaticity. As such researchers were in search of better models. Generalised Autoregressive Conditional Heteroscedastic (GARCH) models was thought of and in literature one can find its use in time series modelling. Yaziz *et al.*(2011) studied ARIMA and GARCH expand models in forecasting crude oil prices and found that the GARCH model was better than ARIMA model. Paul *et al* (2009) studied India's volatile spice export data through the Box-Jenkins Autoregressive integrated moving average (ARIMA) approach and also through, GARCH nonlinear time-series model along with its estimation procedures. Mishra *et al.*(2013) also forecasted the production behaviour of onion in India using ARIMA modeling techniques. Lagrange multiplier test for testing presence of Autoregressive conditional heteroscedastic (ARCH) effects was also discussed. Comparative study of the fitted ARIMA and GARCH models was carried out from the viewpoint of dynamic one-step ahead forecast error variance along with Mean square prediction error (MSPE), Mean absolute prediction error (MAPE) and Relative mean absolute prediction error (RMAPE) to establish the superiority of GARCH model over ARIMA. All the above studies and other related studies have mostly considered modelling taking only the time series data of a particular phenomenon, but as mentioned earlier, one can not ignore the role of factors of production during modelling with time series data. The present study is an attempt to use the factors of production in the model. As such the study attempts to examine the production scenario, growth, trend and forecast the production of pulses in major growing states of India using ARIMA and GARCH model, taking in to consideration the factors like rainfall, fertilizer.



Contribution (%) of different states in pulses production of India during 2011

MATERIALS AND METHODS

Based on their relative contribution to Indian pulses basket, five major states along with whole India is considered for the present study. Data related to area, production and yield of pulses in major five states of production and India along with rainfall and fertilizer could be obtained for the period 1975 to 2009. To develop forecast models and subsequently use these models to forecast the series for the years to come, data for the whole period excepting last four years are used for model building, while data for last four years are used for model validation purpose.

Descriptive statistics are useful to describe patterns and general trends in a data set. It includes numerical and graphic procedure to summarize a set of data in a clear and understandable way. To examine the nature of each series these have been subjected to different descriptive measures. Statistical measure used to describe the above series are minimum, maximum, average, standard error, skewness, kurtosis, simple growth rate. Simple growth rates have been calculated using the formula,

SGR% $= \frac{(Y_t - Y_0)}{Y_0 n} \times 100$ where, Y_t and Y_0 are the values of the last year and the first year of the series; n is the number of years.

Autoregressive model (AR)

ARIMA models stands for Autoregressive Integrated Moving Average models. An ARIMA model is in-fact a combination of AR, MA models with integration.

Autoregressive model (AR)

The notation AR (*p*) refers to the autoregressive model of order *p*. The AR (*p*) model is written

$$X_t = c + \sum_{i=1}^{P} \alpha_i + X_{t-i} + \mu_t$$

where $\alpha_1, \alpha_2 ... \alpha_p$ are the parameters of the model, *c* is a constant and μ_t is white noise i.e. $\mu_t \sim WN(0, \sigma^2)$. Sometimes the constant term is omitted for simplicity.

Moving Average model (MA)

The notation MA (*q*) refers to the moving average model of order *q*:

$$X_t = \mu + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t$$

where the $\theta_1, ..., \theta_q$ are the parameters of the model, μ is the expectation of X_t (often assumed to equal 0), and the is the error term.

ARMA model

A time series $\{X_t\}$ is an ARMA (p, q) if $\{X_t\}$ is stationary and if for every t, $X_t - \phi_1 X_{t-1} - ...... - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + ... + \theta_q Z_{t-q}$ where, $\{Z_t\} \sim WN(0, \sigma^2)$ and the polynomials $(1 - \phi_1 Z - ...... - \phi_p Z^p)$ have no common factors.

ARIMA model

A time series $\{X_t\}$ is an ARIMA (p,d,q) if $Y_t = (1-B)^d X_t$ is a causal ARMA (p,q) process. This means $\{X_t\}$ satisfies $\phi^*(B) X_t \equiv \phi(B)(1-B)^d X_t = \theta(B) Z_t$, where, $\{Z_t\} \sim WN(0, \sigma^2)$ $\phi(Z)$ and $\theta(Z)$ are polynomials of degree p and q respectively and $\phi(Z) \neq 0$ for $|Z| \trianglelefteq 1$. The polynomial $\phi^*(Z)$ has a zero of order d at $z = 1$. The process $\{X_t\}$ is stationary if and only if d = 0 and in that case it reduces to ARMA (p,q) process.

Given a set of time series data, one can calculate the mean, variance, autocorrelation function (ACF), and partial autocorrelation function (PACF) of the time series. The calculation enables one to look at the estimated ACF and PACF which gives an idea about the correlation between observations, indicating the sub-group of models to be entertained. This process is done by looking at the cut-offs in the ACF and PACF. At the identification stage, one would try to match the estimated ACF and PACF with the theoretical ACF and PACF as a guide for tentative model selection, but the final decision is made once the model is estimated and diagnosed.

GARCH (p,q) Model

GARCH stands for Generalized Autoregressive Conditional Heteroscedasticity.

Generalized - It is developed by Bollerslev (1986) as a generalization of Engle's original ARCH volatility modelling technique.

Autoregressive - It describes a feedback mechanism that incorporates past observations into the present.

Conditional - It implies a dependence on the observations of the immediate past.

Heteroscedasticity - Loosely speaking, we can think of heteroscedasticity as time-varying variance.

GARCH is a mechanism that includes past variances in the explanation of future variances. More specifically, GARCH is a time series technique that allows users to model and forecast the conditional variance of the errors. It is used to take into account excess kurtosis and volatility clustering. To formally define GARCH, let $\varepsilon_1, \varepsilon_2, ..., \varepsilon_T$ be the time series observations denoting the errors and let F_t be the set of ε_t up to time T, including ε_t for t d" 0. As defined by Bollerslev (1986), "the process ε_t is a Generalized Autoregressive Conditional Heteroscedastic model of order p and q, denoted by GARCH(p, q), if ε_t given an information set F_t has a mean of zero and conditional variance h_t given as $h_t = \alpha_0 + \alpha_1 \varepsilon^2_{t-1} + ... + \alpha_q \varepsilon^2_{t-1} + \beta_1 h_{t-1} + ... + \beta_p h_{t-p}$

$$= \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}$$

Here the conditional variance h_t is the main component of a GARCH model and is expressed as a function of three terms namely:

- α_0 - a constant term
- $\sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2$ ARCH term
- $\sum_{j=1}^{p}\beta_j h_{t-j}$ GARCH term

We define ε^2_{t-i}, as the past i period's squared residual from the mean equation while the h_{t-j} is the past j period's forecast variance. The order of the GARCH term and ARCH term are denoted by p and q respectively. The unknown parameters which needs to be estimated are α_0, α_i and β_j, where i = 1, . . . , q and j = 1, . . . , p. To guarantee that the conditional variance $h_t > 0$, it needs to satisfy the following conditions: $\alpha_0 > 0$, α_i e" 0, and β_j e" 0.

ARCH (q)

The ARCH model is a special case of a GARCH specification in which, there is no GARCH terms in the conditional variance equation. Thus ARCH(q) = GARCH (0, q). The process ε_t is an Autoregressive Conditional Heteroscedastic process of order q or ARCH(q), if h_t is given by

$$h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + ... + \alpha_q\varepsilon_{t-q}^2 = \alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2$$

where q > 0 and $\alpha_0 > 0$, and α_i e ≥ 0 for i = 1, . . . , q. Again, the conditions $\alpha_0 > 0$ and αi ≥ 0 are needed to guarantee that the conditional variance $h_t > 0$. To carry out the process of parameter estimation, consider the simplest model which is the GARCH (0,1) model, where h_t is given by $h_t = \alpha_0 + \alpha_1\varepsilon^2_{t-1}$.

The GARCH model for exogenous variables is given as follows

$$= \alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2 + \sum_{j=1}^{p}\beta_j h_{t-j} + \gamma Z_t$$

Where γZ_t = is the inclusion of exogenous factor.

The parameters α_0 and α_1 can be approximated by maximum likelihood estimation or MLE. The likelihood L of a sample of n observations $x_1, x_2, . . . , x_n$, is the joint probability function $p(x_1, x_2, . . . , x_n)$ when $x_1, x_2, . . . , x_n$ are discrete random variables. If $x_1, x_2, . . . , x_n$ are continuous random variables, then the likelihood L of a sample of n observations, $x_1, x_2, . . . , x_n$, is the joint density function $f(x_1, x_2, . . . , x_n)$. Let L be the likelihood of a sample, where L is a function of the parameters $\theta_1, \theta_2, . . . , \theta_k$. Then the maximum likelihood estimators of $\theta_1, \theta_2, . . . , \theta_k$ are the values of $\theta_1, \theta_2, . . . , \theta_k$ that maximize L. Let θ be an element of Ω. If Ω is an open interval, and if L(θ) is differentiable and assumes a maximum on θ, then MLE will be a solution of the equation $\frac{\partial L(\theta)}{\partial\theta} = 0$.

GARCH (1,1)

The most widely used GARCH(p, q) model for GARCH (1,1) takes the form of $h_t = \alpha_0 + \alpha_1\varepsilon^2_{t-1} + \beta_1 h_{t-1}$

α_0 - Constant term

$\alpha_1\varepsilon^2_{t-1}$ - ARCH term reflects the volatility from the previous period, measured as the lag of the squared residual from the mean equation

$\beta_1 h_{t-1}$ - GARCH term, it is the last periods forecast variance

The (1, 1) in GARCH (1, 1) refers to the presence of a first-order GARCH term (the first term in parentheses) and a first-order ARCH term (the second term in parentheses). We can interpret the period's variance as the weighted average of a long term average (the constant), the forecasted variance from last period (the GARCH term), and information about the volatility observed in the previous period.

Among the competitive models, best models are selected based on minimum value of Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Square Error (MSE) and Mean Absolute Percentage Error (MAPE), maximum value of Coefficient of Determination (R^2) and of course the significance of the coefficients of the models. Best fitted models are put under diagnostic checks through auto correlation function (ACF) and partial autocorrelation function (PACF) of the residuals.

$$MAE = \frac{\sum_{i=1}^{n}\left|X_i - \hat{X}_i\right|}{n}, R^2 = \frac{\sum_{i=1}^{n}\left(\hat{X}_i - \bar{X}\right)^2}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i - \hat{X}_i\right)^2}{n}}, MAPE = \frac{\sum_{i=1}^{n}\left|\frac{X_i - \hat{X}_i}{X_i}\right|}{n} \text{x100}$$

In the present study, first appropriate ARIMA and GARCH models are identified based on criteria mentioned above; the best models are used for forecasting purpose. In the second phase attempts have been made to introduce rainfall and fertilizer in the respective best models and their comparative models to examine their effects of such inclusion, particularly on model accuracy.

RESULTS AND DISCUSSION

In Andhra Pradesh, since 1975, the area under pulses has increased from 1341 thousand hectare to 2227 thousand hectare registering a growth of almost 10.90% (Table 1). For whole India, average area under pulses being 22778 thousand hectare. State wise figures show that in Madhya Pradesh has increased production from 1912 thousand tonnes to 3480 thousand tonnes registering a growth of almost 12.47 % during the period. The average yield under pulses in Uttar Pradesh being 837.43 kg ha^{-1}. Since 1975 the area under pulses in Maharashtra has increased from 1341 thousand hectare to 4347 thousand hectare registering a growth of almost 21.53 %. The Rajasthan

Table1: ***Per Se*** **performance of pulses production in India**

Area ('000ha)	Maharashtra	Uttar Pradesh	Andhra Pradesh	Rajasthan	Madhya Pradesh	India
Mean	3033.20	2838.121	1650.466	3387.755	4699.952	22778.788
Standard Error	138.174	31.846	46.419	87.913	53.786	113.025
Kurtosis	0.333	1.159	-0.441	0.156	-0.353	1.525
Skewness	-1.041	-1.216	0.869	0.045	-0.654	-1.105
Minimum	1341.000	2341.143	1341.000	2268.927	3953.967	20953.333
Maximum	4347.888	3072.728	2227.064	4530.377	5132.233	23740.000
SGR (%)	21.534	-10.189	10.907	-12.020	-1.599	-0.263
Production ('000t)						
Mean	1518.111	2376.704	786.447	1458.899	2756.300	12664.242
Standard Error	95.598	38.937	57.878	57.129	78.985	180.641
Kurtosis	-0.668	0.584	-0.047	0.341	-0.818	-0.049
Skewness	-0.433	-0.989	0.851	0.670	-0.214	-0.541
Minimum	480.087	1825.467	355.025	880.700	1912.660	10236.667
Maximum	2444.333	2709.467	1524.667	2307.967	3480.433	14510.000
SGR (%)	25.995	-13.429	25.023	-16.341	12.470	6.906
Yield(kg.ha^{-1})						
Mean	479.606	837.434	461.081	415.657	606.384	555.828
Standard Error	14.453	8.663	22.041	9.382	19.618	8.530
Kurtosis	-0.935	1.431	0.501	-0.466	-1.138	-0.540
Skewness	-0.062	-1.347	0.556	0.318	-0.357	-0.599
Minimum	349.333	704.667	260.000	325.333	404.667	447.000
Maximum	644.000	907.000	766.667	539.000	764.333	632.000
SGR (%)	13.655	-2.228	21.278	-1.965	13.197	7.206

yield for pulses had a standard error of 9.382 kg ha^{-1}. Madhya Pradesh recorded highest yield 764 kg ha^{-1} in 2007 and lowest at 404 kg ha^{-1} in 1975.

From table 3, one can see that simple ARIMA models are by and large better forecasting model compared to GARCH model in terms of MAPE and MAE value excepting area under Maharashtra and MP where simple GARCH has outperform simple ARIMA. So far about the model with rainfall, fertilizer or both, from table 2, it is clear that accuracy of the model would be increased by the inclusion of these factors of production for both ARIMA or GARCH model in majority of the states and also for India. But one can't be sure of superiority of either the model over its counterpart. Thus, from the present study one can very well think for inclusions of different factors of productions in the best fitted time series model to increase the accuracy and to trace the path of production behaviours in a better way compared to simple time series model.

Time series ARIMA or GARCH models can effectively be validated by comparing the observed values with those of the values obtained through these models. A close look into the table 3 reveals that neither the GARCH nor the ARIMA model for the series under consideration would be taken as

Table 2: Forecasting ARIMA models for area, production and yield of total pulses in major states of India

States	Parameter	Best fitted ARIMA models	R^2	RMSE	MAPE	MAE	MaxAPE	MaxAE
Madhya Pradesh	A	ARIMA (1,1,2)	0.747	292.725	4.761	202.505	41.133	1462.00
	P	ARIMA (0,1,5)	0.770	67.221	10.302	52.018	46.182	133.466
	Y	ARIMA (1,1,2)	0.778	369.456	11.678	259.925	70.435	1014.00
Maharashtra	A	ARIMA (1,1,3)	0.936	202.019	4.345	126.637	18.658	603.841
	P	ARIMA (1,1,5)	0.956	122.94	6.048	85.146	16.834	278.845
	Y	ARIMA (0,1,5)	0.905	27.505	4.354	20.344	11.287	48.083
UP	A	ARIMA (0,1,4)	0.909	58.533	1.393	38.255	6.974	168.425
	P	ARIMA (1,1,4)	0.811	106.638	2.999	69.955	15.612	328.636
	Y	ARIMA (0,1,2)	0.513	36.453	3.374	27.428	14.67	103.423
AP	A	ARIMA (1,1,2)	0.963	51.901	2.25	36.935	6.789	126.737
	P	ARIMA (1,1,2)	0.975	47.253	4.811	35.031	16.263	115.193
	Y	ARIMA (1,1,5)	0.955	24.055	3.332	14.435	14.673	64.951
Rajasthan	A	ARIMA (1,1,5)	0.603	319.849	6.821	216.489	21.587	614.666
	P	ARIMA (1,1,5)	0.652	215.048	11.058	153.641	35.104	372.979
	Y	ARIMA (0,1,3)	0.521	39.686	7.886	31.879	17.241	66.28
India	A	ARIMA (1,1,5)	0.747	381.355	1.148	261.774	3.173	736.68
	P	ARIMA (1,1,2)	0.616	635.971	4.146	517.284	9.169	1071.00
	Y	ARIMA (0,1,5)	0.902	16.115	2.195	11.821	9.329	43.844

Table 3: Comparison between the ARIMA and GARCH model of area, production and yield of pulses in India

	MAPE								MAE							
	ARIMA				GARCH				ARIMA				GARCH			
	Simple	R+F	R	F	Simple	R+F	R	F	Simple	R+F	R	F	Simple	R+F	R	F
Maharashtra																
A	**4.35**				4.45				126.64				**122.11**			
P	6.05	**2.81**	2.89	3.53	8.00	3.61	3.11	3.26	85.15	**65.03**	66.65	82.22	106.42	71.22	72.47	75.45
Y	4.35	2.52	2.67	**2.51**	5.46	6.54	6.62	6.76	20.34	20.39	21.57	**20.27**	25.16	30.80	32.15	81.66
Uttar Pradesh																
A	**1.39**				1.47				**38.26**				41.19			
P	3.00	2.81	2.89	2.89	3.54	**2.56**	3.47	2.67	69.96	**65.03**	66.65	67.45	82.17	73.32	79.33	81.45
Y	3.37	**2.64**	3.38	3.31	2.87	2.73	2.73	2.87	27.43	**21.30**	27.47	20.37	23.35	22.30	22.24	23.32
Andhra Pradesh																
A	**2.25**				3.17				**36.94**				57.14			
P	4.81	**3.50**	4.76	3.78	5.38	3.75	3.89	4.01	35.03	25.67	36.48	27.93	40.82	**23.45**	36.12	29.22
Y	3.33	**2.78**	4.26	3.28	4.15	3.11	3.45	2.96	14.44	12.16	19.66	14.98	18.47	**11.45**	22.43	16.29
Rajasthan																
A	**6.82**				8.78				**216.49**				275.00			
P	**11.06**	11.07	11.54	11.49	14.00	13.38	13.38	14.35	**153.64**	157.07	163.58	163.06	204.00	184.63	185.70	197.78
Y	7.89	7.63	**7.49**	7.62	8.65	7.91	8.48	8.55	31.88	30.91	**30.31**	30.88	35.23	32.67	34.66	34.97
Madhya Pradesh																
A	4.76				**2.47**				202.51				**113.66**			
P	5.04	**5.01**	5.06	5.03	6.91	7.26	7.14	7.27	136.83	**135.51**	137.47	135.66	182.93	197.35	196.50	197.00
Y	2.89	2.86	**2.82**	2.85	3.49	2.66	9.97	9.67	16.57	16.09	16.45	15.68	19.76	**14.53**	15.51	14.33
India																
A	**1.15**				1.39				**261.77**				316.32			
P	4.15	2.55	2.96	2.79	3.43	**2.42**	2.77	2.62	517.28	310.36	366.07	346.12	421.99	**303.12**	324.05	345.66
Y	**2.20**	2.75	3.12	2.99	2.62	2.57	2.98	2.54	11.82	14.87	12.44	14.21	14.00	**11.26**	11.97	13.72

Note: *A= Area, P= Production, Y= Yield, R = Rainfall, F = Fertilizer, Simple = without the factors, (R+F) = both Rainfall and Fertilizer included;Bold faces letters are for minimum values*

Table 4: Model validations and forecasting of area, production and yield of pulses in India

ARIMA							GARCH			
State	Model (p,d,q)	2006 Observed	2006 Predicated	2007 Observed	2007 Predicated	2015 Predicated	Models (p,q)	2006 Predicated	2007 Predicated	2015 Predicated
Maharashtra										
A	(1,1,3)	3729	3890	3886	3502	4413	(1,1)	4415	4484	5100
P	(1,1,5)	2328	2187	2350	2145	2883	(1,1)	2453	2479	2881
Y	(0,1,5)	644	634	628	627	674	(1,1)	578	590	651
Uttar Pradesh										
A	(0,1,4)	2448	2384	2341	2387	2096	(1,1)	2432	2387	2429
P	(1,1,4)	1850	1826	1825	1932	1993	(1,1)	1981	1932	2093
Y	(0,1,2)	756	731	785	785	799	(1,1)	815	785	862
Andhra Pradesh										
A	(1,1,2)	2016	2285	2016	2084	2326	(1,1)	2248	2272	2461
P	(1,1,2)	1497	1641	1525	1634	1949	(1,1)	1493	1515	1708
Y	(1,1,5)	751	682	767	697	795	(1,1)	671	674	639
Rajasthan										
A	(1,1,5)	4530	3184	3117	2996	3051	(1,1)	3681	3581	3240
P	(1,1,5)	1620	1523	1364	1274	1227	(1,1)	1316	1316	1252
Y	(0,1,3)	375	409	453	394	384	(1,1)	376	381	378
Madhya Pradesh										
A	(1,1,2)	4108	4251	4026	4167	5070	(1,1)	4140	4138	4060
P	(0,1,5)	3203	3328	2454	2578	3661	(1,1)	3137	3245	3599
Y	(1,1,2)	780	712	609	631	789	(1,1)	750	774	879
India										
A	(1,1,5)	23070	22146	22970	22554	21530	(1,1)	22494	22309	21778
P	(1,1,2)	14117	13730	14510	13804	14360	(1,1)	13707	13818	14482
Y	(0,1,5)	612	621	632	629	667	(1,1)	607	616	658

Note: A= Area in ('000ha), P = Production in ('000t) and Y= Yield in (kg ha^{-1})

uniformly better over the respective rivals. This fact has also been established through the comparison of MAPE and MAE for these two types of model for different series. Using the best fitted models predictions has been made for the year 2015.

Forecast values are presented in table 3, as obtained from the best fitted models; it can be seen that for whole India, pulses area will decrease marginally from 22.9 million hectare in 2007 to 21.53million hectare in 2015. With a major contribution, Madhya Pradesh produced 2.454 million tons in 2007 against predicated 2.578 million tons. The state is forecasted to produce 3.661 million tons during 2015. The study expects to have 14.360 million tons of pulses with an average productivity of 667 kg ha^{-1} during 2015. It is clear from the table that there will be consistent increase in production coupled with marginal decrease in area for all India. The increase in production would be attributed to increase in areas of major states in association with increase in productivity. Thus, the study expects Madhya Pradesh and other major states to play role in augmenting pulses production of India while the contribution of marginal states will further reduce.

From the present study one can conclude that there has been increase in production of pulses in India during the last three decades or more. Both ARIMA and GARCH models can be used for modelling pulses production in India; inclusion of the factors like fertilizer and rainfall increases the accuracy of the model. Superiority of either ARIMA or GARCH could not be establishing emphatically in modelling data of pulses. Pulses production for whole India is also expected to increase during the years to come and contribution from major states will continue to increase while that from the minor states will reduce. In-spite of growth in all fronts the major concern is that, the productivity of major contributing state as well as that of whole India would still be about 1/3rd of the productivity of USA and Canada with 1.8t ha^{-1}. India needs to augment productivity in pulses for nutritional security of its huge population.

ACKNOWLEDGEMENT

The authors acknowledge and thak Department of Science and Technology for awarding Inspire fellowship for pursuing Doctoral Programme.

REFERENCES

Anonymous, 2011. *Agricultural Statistics at a Glance*. Directorate of Economics and Statistics, Department of Agriculture and Cooperation, Ministry of Agriculture, Govt. of India.

Box, G.E.P. and Jenkins G.M, 1976. *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco.

Bollerslev, T. 1986. Generalized autoregressive conditional heteroscedasticity, *J. Econ.*, **31**: 307-27.

Hossain, Md Z., Quazi, A. S., Ali Md Z. 2006. ARIMA model and forecasting with three types of pulse prices in Bangladesh: a case study, *Int. J. Soc.Econ.*, **33** : 344 –53.

Mishra, P., Vishwajith, K.P., B. S. Dhekale and P.K. Sahu. 2013. Instability and forecasting using ARIMA model in area, production and productivity of onion in India. *J. Crop Weed*, **9**:96-01

Paul, R. K., Prajneshu, and Ghosh, H. 2009. GARCH nonlinear time series analysis for modelling and forecasting of India's volatile spices export data, *J. Indian Soc. Agric. Stat.,* **63**: 123-31.

Reddy, A.A. 2009. Pulses production technology: status and way forward, *Econ. political Weekly,* **44**: 73-80.

Sher, F. and Ahmad E. 2008. Forecasting wheat production in Pakistan, *The Lahore J. Eco.,* **13**: 57-85.

Yaziz, S.R., Ahmad, M.H. Nian, L.C. and Muhammad, N. 2011. A comparative study on Box-Jenkins and GARCH models in forecasting crude oil prices, *J. Applied Sci.,* **11**: 1129-35.

Yaseen, M., Zakir, M., Shahzad, I.U.D, Khan, M.I. Javed, M.A. 2005. Modelling and forecasting the sugarcane yield of Pakistan, *Int. J. Ag. Biol,* **7**: 180-83.