Journal of Agriculture, Biology and Applied Statistics
Vol. 1, No. 2, 2022, pp. 101-121

URL : www.arfjournals.com
https://DOI: 10.47509/JABAS.2022.v01i02.05

Modeling and Forecasting of Urad in India

K. P. Vishwajith[1*], P. K. Sahu[1], Aditya Bhooshan Srivastava[2] and Rajani Gautam[3]
[1]*Department of Agricultural Statistics, Bidhan Chandra Krishi Viswavidyalaya, Mohanpur- 741 252, Nadia, West Bengal, India*
[2]*Department of Agricultural Economics, Acharya Narendra Deva University of Agriculture and Technology, Kumarganj, Ayodhya-224229 (Uttar Pradesh)*
[3]*Pandit S. N. Shukla University, Shahdol (M.P.) India*
**Corresponding email: vishwajithkp@gmail.com*

To cite this article

K. P. Vishwajith, P. K. Sahu, Aditya Bhooshan Srivastava & Rajani Gautam (2022). Modeling and Forecasting of Urad in India. *Journal of Agriculture, Biology and Applied Statistics*. Vol. 1, No. 2, pp. 101-121. *https://DOI: 10.47509/JABAS.2022.v01i02.05*

Abstract: In this study researcher has been made to apply the autoregressive integrated moving average (ARIMA) and Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model approach to investigate the trend in Urad area, production and productivity in Bihar, Madhya Pradesh, Uttar Pradesh, West Bengal, and India. Yearly data from 1970 to 2009 were used for forecasting up to 2020. In comparison, we get that in area ARIMA model outperformed GARCH model in all the states under study, whereas inclusion of auxiliary variables improve the model accuracy for production and productivity in maximum cases. Furthermore, according to the trend analysis analysis signifies that production of uradin many state has shown decreasing trend in recent period under study. Forecasted values are likely to help the policy maker in existing battle against food and nutritional security.

Keywords: Area, ARIMA, GARCH, trend, modelling, forecasting, production and productivity.

Introduction

India, with a large population of poor and malnourished people, has long encouraged a cereal-based diet based on subsidised staples like rice and wheat. Dietary habits, on the other hand, are changing today. Policymakers, researchers, and health advocates are exploring for strategies to combat malnutrition, not only hunger, in the country. Pulses (the dried, edible seeds of legumes) are becoming more popular as the focus shifts from calorie consumption to nutrients. 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.

Urad, like mung is primarily a warm season crop. It is also grown both in kharif and rabi seasons. It is a kharif season crop in Andhra Pradesh, Bihar, Gujarat, Karnataka, Madhya

Article History Received : 18 September 2022; Revised : 22 October 2022; Accepted : 30 October 2022; Published : 22 December 2022

Pradesh, Maharashtra, Orissa, Rajasthan, Tamil Nadu, Uttar Pradesh and West Bengal. But, some of these states like Andhra Pradesh, Tamil Nadu, Uttar Pradesh and West Bengal along with Assam cultivate urad in the rabiseason too. The crop is mainly grown for its beans, which are used as a whole or split. The total area under the crop has increased progressively to 3.24 million ha in 2007-08 (Jain, 2012). The leading states in area are Madhya Pradesh (18.21%), Uttar Pradesh (17.11%), Maharashtra (14.84%) and Andhra Pradesh (14.28%). Besides, it is also cultivated in Tamil Nadu (9.3%), Rajasthan (3.93%), Karnataka (3.91%) and Gujarat (3.14%). Uttar Pradesh with 21 percent share in all India production is the leading state. Uttar Pradesh, Maharashtra, Andhra Pradesh and Madhya Pradesh together grow around 67 percent. The yield level of urad was found extremely low at all India level (446 kgs/ha.). The highest yield of 853.65 kgs/ha was reported in Bihar. It is depressing to note that states of Karnataka, Madhya Pradesh and Tamil Nadu have exhibited very low yield of urad between three to four quintals per hectare.

In the arena of time series modelling and forecasting, Box-Jenkins ARIMA technique has come in a great way. Attempts have been made to forecast various crops using ARIMA techniques. Among these, the works of Sahu (2006) for forecasting of irrigated crops like potato, mustard and wheat. Mishra *et al.* (2013) for onion production in India. Rahman *et al.* (2013) Forecasting of lentil pulses production in Bangladesh and Vishwajith *et al.* (2016) for sugarcane production in major growing states of India are few to mentioned.

Though, ARIMA models have got wide application in modeling time series data, this is being criticized for its assumption of linearity, and homoscedaticity. As such researchers were in search of better models. Generalised Auto Regressive Conditional Heteroscedastic (GARCH) models was thought of and in literature one can find its use in time series modeling. Paul *et al.* (2009) studied India's volatile spice export data through the Box-Jenkins Auto Regressive integrated moving average (ARIMA) approach and also through, GARCH nonlinear time series model along with its estimation procedures. 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. Vishwajith *et al.* (2014) analyzed trend and forecasted production pulse production in India using ARIMA and GARCH models. All the above studies and other related studies have mostly considered modelling taking only the time series data of a particular phenomenon, but production of anycrops depends on many production factors like rainfall, temperature, relative humidity, fertilizer etc. There is not enough work on forecasting production by taking care of factors of production using ARIMAx model in Indian context. In this contest, [Vishwajith *et al.* (2016)], comparison has been made among the ARIMA,GARCH and ARIMAx models and selected the best model and forecasted the gram production upto 2020. The present study is a sincere attempt to use the factors of production in the model. As such the study attempts to examine the production scenario, growth and forecast the production of gram in major growing states of India using best model among ARIMA, GARCH and ARIMAx model.

Materials and Methods

The main approaches to the research problem with their methodologies are discussed here:

Source of Data

The data gathered is entirely secondary. The data on Uradproduction from 1970 to 2012 was collected from Directorate of Economics and Statistics.

Trend Models

The model can be described as a means of presenting a process/system. The statistical model generally traces the path of the process along with its statistical properties and implications. In the present topic, we are interested in studying the path and nature of the series under our preview through different models, which are briefly given in table 1.

Table 1: Different trend models

Model	*Form*
Linear Model	$Y_t=b_0+(b1t)$
Quadratic Model	$Y_t=b_0+(b_1t)+(b_2t^2)$
Compound Model	$Y_t=b_0(b^t)$ or $\ln(Y_t)=\ln(b_0)+\ln(b_1)$
Cubic Model	$Y_t=b_0+(b_1t)+(b_2t^2)+(b_2t^3)$
Exponential Model	$Y_t=b_0e^{(b1t)}$ or, $\ln(Y_t)=\ln(b_0)+(b_1t)$
Logarithmic Model	$Y_t=b_0+ b_1\ln(t)$
Growth Model	$\ln(Y_t)=b_0+b_1t\ Y_t$

Where, Y_t is the value of the series at time t and b_0, b_1, b_2, b_3 are the parameters.

ARIMA models stand for Auto regressive 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 as

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

where, $\alpha1$, $\alpha2$, ..., αp are the parameters of the model, c is a constant and μt is white noise i.e. $\mu t \sim WN(0, 2)$. 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:

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

where, the $\theta 1, \ldots, \theta q$ are the parameters of the model, μ is the expectation of Xt(often assumed to equal 0) and the t is the error term.

ARMA model : A time series {Xt} is an ARMA(p, q) if $\{X_t\}$ is stationary and if for every t,

$$X_t - \emptyset_1 X_{t-1} - \cdots - \emptyset_p X_{t-p} = z_t + \theta_1 Z_{t-1} + \cdots + \theta_q z_{t-q}$$

Where, {Zt} ~ WN(0, σ^2) and the polynomials $(1 - \emptyset_1 Z - \cdots - \emptyset_p Z^p)$ and $(1 - \emptyset_1 Z + \cdots + \emptyset_p Z^p)$ have no common factors.

ARIMA Model: A time series $\{X_t\}$ is an ARIMA (p, d, q) if $Y_t = (1\text{-}B)^d X_t$ is a casual ARIMA (p, q) process. This mean $\{X_t\}$ satisfies $\emptyset(B)\, X_t \equiv (B)\,(1-B)^d X_t = \theta(B) Z_t$

Where, $\{Z_t\} \sim$ WN $(0, \sigma^2)$ $\phi(z)$ and $\theta(z)$ are polynomials of degree p and q respectively and $\phi(z) = 0$ for $|Z| \leq 1$. The polynomial $\phi^*(Z)$ has a zero of order d at $z = 1$. The process {Xt} is stationary if and only if d = 0 and in that case it reduces to ARMA (p, q) process.

The stationarity requirement ensures that one can obtain useful estimates of the mean, variance and ACF from a sample. If a process has a mean that is changing in each time period, one could not obtain useful estimates since only one observation available per time period. This necessitates testing any observed series of data for stationarity. First the given data series are tested for stationarity through ADF and KPSS test. If the data are non-stationary, first order differencing was made to make data stationary. 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 Ft be the set of ε t up to time T, including εt for t d ≤ 0. As defined by Bollerslev (1986), "the process εt is a Generalized Autoregressive Conditional Heteroscedastic model of order p and q, denoted by GARCH (p, q), if å t given an information set F t has a mean of zero and conditional variance ht given as,

$$h_t = \alpha_1 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_q \varepsilon_{t-q}^2 + \beta_1 h_{t-1} + \cdots \beta_p h_{t-p}$$

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

Here, the conditional variance ht is the main component of a GARCH model and is expressed as a function of three terms namely: $\alpha_0, \alpha_0, \sum_{i=1}^{q} \alpha_i \varepsilon_{t-1}^2$ and $\sum_{j=1}^{p} \beta_j h_{t-j}$ are a constant, ARCH and GARCH term respectively.

We define ε_{t-1}^2,, 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, αi and βj, where i = 1, ..., q and j = 1, ..., p. To guarantee that the conditional variance $h_t > 0$, it needs to satisfy the following conditions: α 0 > 0, αi ≥ 0 and βj ≥ 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 ht is an Autoregressive Conditional Heteroscedastic process of order q or ARCH(q), if ht is given by

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

Where, q > 0 and $\alpha_0 > 0$ and $\alpha_i \geq 0$ for i = 1,…….q. Again, the condition $\alpha_0>0$ and $\alpha_i \geq 0$ 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 ht is given by $h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2$

The parameters α0 and α1 can be approximated by maximum likelihood estimation or MLE. The likelihood L of a sample of n observations x1, x2, ..., xn, is the joint probability function p(x1, x2, ..., xn) when x1, x2, ..., xn are discrete random variables. If x1, x2, ...,xn are continuous random variables, then the likelihood L of a sample of n observations, x1, x2, ..., xn is the joint density function $f(x_1, x_2, ..., xn)$. 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

$$\text{Equation } \frac{\partial L(\theta)}{\partial \theta} = 0$$

GARCH (1,1) : The most widely used GARCH (p, q) model for GARCH (1,1) takes the form of $h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$, where α_0 is constant term; $\alpha_1 \epsilon_{t-1}^2$ is ARCH term reflects the volatility from the previous period, measured as the lag of the squared residual from the mean equation and β_1 h_{t-1} is the 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. ARIMAx methodology ARIMAx model is a generalization of ARIMA model and is capable of incorporating an external input variable (X). Given a (k+1)- time-series process {(yt, xt)}, where yt and k components of xt are real valued random variables, ARIMAx model assumes the form

$$Y_t(1 - \sum_{s=1}^{P} \alpha_s L^s) = \mu + \sum_{i=1}^{q} \beta_s L^s x_t + (1 + \sum_{s=1}^{P} \gamma_s L^s) e_t$$

Where, L is the usual lag operator ($L^s y_t = y_{t-s}$ '$L^s x_t = x_{t-s}$, etc), $\mu \in R$, $\alpha_s \in R, \beta_s \in R^k$ and $\gamma_s \in R$ are parameters, et's errors and p, q and r are natural numbers specified in advance. The first step in building an ARIMAx model consists of identifying a suitable ARIMA model for the endogenous variable. The ARIMAx model concept requires testing for stationarity of exogenous variable before modelling. The factors which are found to affect the arhar productivity significantly in the step wise regression were used as a exogenous variable in ARIMAx models.

Among the competitive ARIMA, GARCH and ARIMAx models, the best fitted models are selected based on the maximum R^2, minimum value Akaike's Information Criterion (AIC), Bayesian Information Criterion (BIC), Mean Error (ME), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Percentage Error (MPE) and Mean Absolute Percentage Error (MAPE). In all three types of model, which has fulfilled most of the above criteria is selected. Best fitted models are again put under diagnostic checks through Ljung-Box- test, ACF and PACF graphs of the residuals. Only those models showing white noise are retained. Among these best fitted ARIMA, GARCH and ARIMAx models, one best model has been selected based on same model selection criteria mentioned above and forecast has been made upto 2020.

AIC =2k-2 ln(L)   BIC=-2*ln(L) +k*ln(n)

$$ME=\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \hat{X}_i\right)$$

$$MPE = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i - \hat{X}_i}{X_i}\right) * 100$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|X_i - \hat{X}_i\right|$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{X_i - \hat{X}_i}{X_i}\right| * 100$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \hat{X}_i)^2}{n}}$$

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

where, $X, \bar{X}, \hat{X}$ are the values of the ith observation, mean and estimated value of the *i*th observation of the variable X and *k* is the number of parameters in the statistical model and *L* is the maximized value of the likelihood function for the estimated model

Result and Discussion

Trends in area, production and productivity of urad in major states of India

Knowing the overall performance, path of movement of the series was traced through different parametric trend models. To workout the trends in area, production and productivity of urad in major growing states different parametric models like linear, polynomial, logarithmic, compound, growth and exponential models were attempted. Among the significant competitive models, the best model was selected based on maximum value of R2 and minimum value of RMSE. The following section presents the result of this exercise.

From the table 2, it is clearly understood that except productivity series of Uttar Pradesh and Madhya Pradesh, all other data series are best fitted with non-linear trend models, in particular polynomial models in maximum cases. This polynomial nature of data series indicates the more than one point of fluctuations. Area under urad in India is significantly fitted with cubic model while production and productivity are fitted with quadratic trend model. Even though area under urad shows increased trend in recent years as indicated by positive b3 coefficient, production shows declining trend mainly due to dismal performance in productivity in recent years. In case of Uttar Pradesh area and productivity has increased exponentially and linearly respectively (Fig. 4.1.4.A. & 4.1.4.C), this combined effect can be clearly visualized in production performance of urad in Uttar Pradesh (Fig. 4.1.4.B.). Area and production of Madhya Pradesh have followed cubic trend while productivity has followed increasing linear trend. The coefficient of cubic time factor for both area and production; coefficient of linear trend for productivity in Madhya Pradesh are found to be positive, which clearly indicates the improvement in production performance of urad in recent years. On other hand production performance of urad in case of Maharashtra, Andhra Pradesh and Tamil Nadu are showing dismal performance in recent years as indicated by coefficients of respective trend models. This clearly a major concern towards food and nutritional security of the Indian people. One must think for resisting these tendencies so as to keep urad production at steady state.

From the fig. 1, one can see that area under urad in Uttar Pradesh has increased continuously throughout the study period which may be result for the highest annual growth rate (Table 4.1.4.A) among the states under study. In Madhya Pradesh urad cultivation has reached maximum in middle of eighties and then started declining thereafter. The negative growth rate of area under urad in case of Maharashtra and Madhya Pradesh can be visualized through fig. 1. We can see the more number of turning points between two consecutive years in production and productivity (Fig. 2&3) parameter of urad in most of the states under study which would have resulted in randomness nature of the data series. This variation nature also indicates the lack of stable varieties or hybrid as in case of other pulses.

Table 2: Trends in area, production and productivity of urad in major states of India.

Parameter	*Model*	*R2*	*RMSE*	*Constant*	*b1*	*b2*	*b3*
			Uttar Pradesh				
Area	Exponential	0.962	30.132	128.700	0.037		
Production	Quadratic	0.958	15.526	52.807	-2.082	0.199	
Productivity	Linear	0.605	54.118	250.378	5.807		
			Maharashtra				
Area	Cubic	0.366	45.803	487.601	-12.601	1.069	-0.020
Production	Cubic	0.777	29.228	143.727	-8.060	0.968	-0.018
Productivity	Quadratic	0.718	51.794	199.28	17.156	-0.256	
			Andhra Pradesh				
Area	Quadratic	0.795	63.225	57.594	27.351	-0.427	
Production	Quadratic	0.764	50.652	-24.184	23.464	-0.410	
Productivity	Quadratic	0.449	93.140	326.478	26.9	-0.543	

			Madhya Pradesh				
Area	Cubic	0.843	39.054	543.365	40.987	-2.589	0.039
Production	Cubic	0.346	13.433	132.386	7.903	-0.421	0.007
Productivity	Linear	0.822	20.96808	208.22	3.899		
			Tamil Nadu				
Area	Cubic	0.584	45.441	34.379	26.705	-1.001	0.012
Production	Quadratic	0.615	21.272	10.221	7.451	-0.136	
Productivity	Quadratic	0.459	46.328	253.868	10.998	-0.192	
			India				
Area	Cubic	0.819	173.686	1653.000	150.982	-4.841	0.048
Production	Quadratic	0.857	114.320	432.272	68.844	-1.166	
Productivity	Quadratic	0.833	23.940	269.998	11.689	-0.182	

Thus from the test of outliers, randomness test and trend analysis of area, production and productivity of urad in major states of India the following important features has emerged out:

1. Analysis of data for different series rejected the presence of outlier in all the parameters under study for all the states.
2. Area in case of Madhya Pradesh; production and productivity in maximum states have changed randomly during the study period.
3. Except productivity series of Uttar Pradesh and Madhya Pradesh, all other data series exhibits non-linear trend, polynomial in maximum cases.

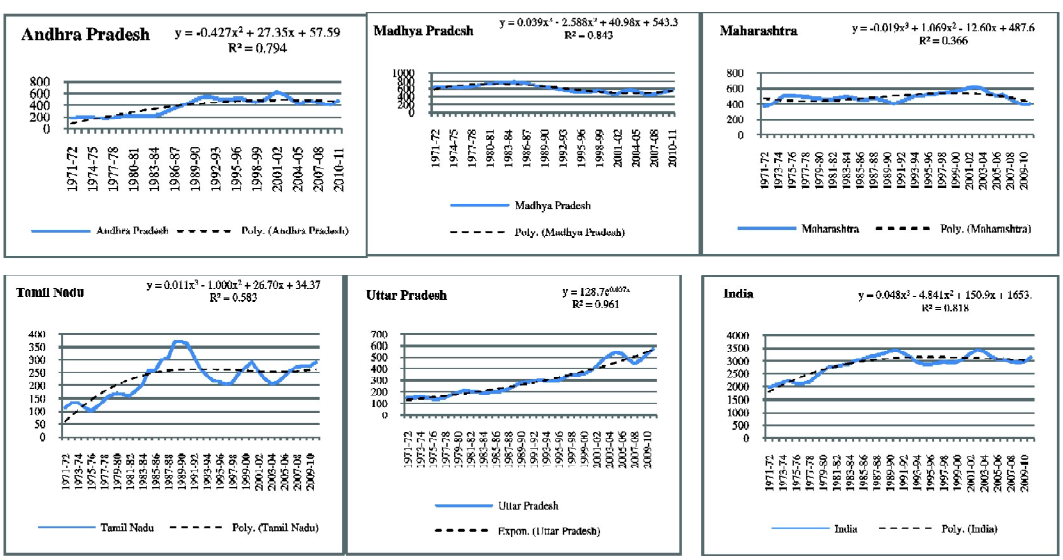

Figure 1: Observed and expected trends of area under urad in major states of India

4. Area under urad in Uttar Pradesh, Madhya Pradesh and Tamil Nadu has increasing trend duing recent period under study.
5. Production and productivity of urad in all the states including whole India excepting Uttar Pradesh and Madhya Pradesh have shown decreasing trend during recent period under study.

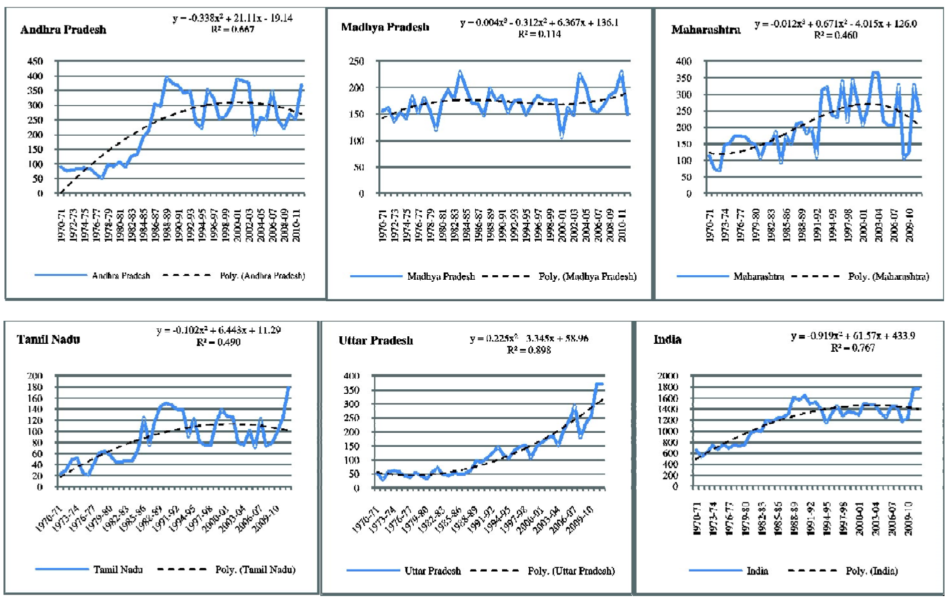

Figure 2: Observed and expected trends of production of urad in major states of India

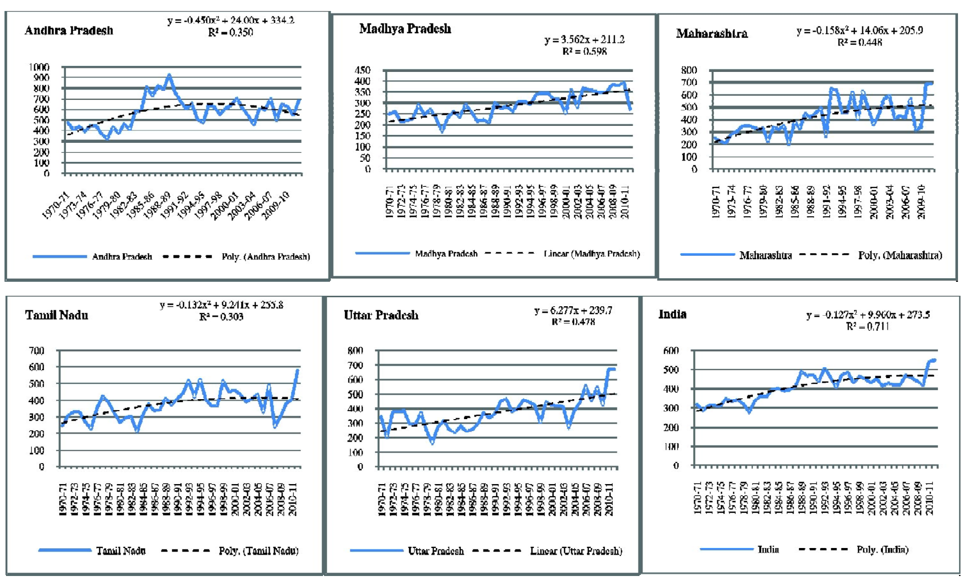

Figure 3: Observed and expected trends of productivity of urad in major states of India

Modeling and forecasting of area under urad

Results of stationarity test of area, production and productivity data series of urad in major states of India are presented in the table 4. From the table one can find that both KPSS and ADF test for the data series of area under urad reject the hypothesis that data are stationary. First order differencing was necessary for the series to make it stationary. After achieving stationarity, various ARIMA models are tried for each series and only best models among the competitive model for each series is selected based on minimum value of AIC, BIC, ME, RMSE, MAE, MPE, MAPE and maximum value of R2. On the other hand, in similar way, various GARCH models have been fitted and best GARCH model for each series is selected and presented in table 5. Developed models are also put under diagnostic checking through Ljung-Box test of residuals (Table 5).

From the table 5, it can be noted that, among the various competitive ARIMA model, ARIMA (0,1,2) for Andhra Pradesh, Madhya Pradesh and Tamil Nadu; ARIMA(2,1,0) for Uttar Pradesh; ARIMA(0,1,1) for Maharashtra and ARIMA(2,1,2) for whole India are found to be best fitted ARIMA models for modeling urad area. On other hand, urad area in all the states under study are bested fitted with GARCH(1) model whereas whole India is found not to have GARCH effect. The results of Ljung-Box test of residuals also reject the presence of significant auto correlation in the residuals of the best fitted ARIMA and GARCH (table 5).

Based on above mentioned error criteria and R2 value, best among the best fitted AIRMA and GARCH models are selected for forecasting area under urad in all the states under study. On comparing it is found that, ARIMA model was the best for modeling area under urad than GARCH all the states under study. It can also be noted that, all the model selection criteria except MAPE are suggesting ARIMA as best model than GARCH model. Hence best fitted ARIMA models are further put under diagnostic checking through ACF and PACF graphs of residuals (Fig. 4) and found that the residuals of selected models are free from significant correlations. These models are used for forecasting urad area up to 2020 (Fig. 5). From the figure 5, it is clearly visible that the selected models performed very well in all the states during the model building stage. The selected models are also validated for accuracy for last three years and observed that the actual and predicted values are almost in range for all the states including whole India (Table 6). The forecasted figures indicate that area under urad would increase marginally in Uttar Pradesh, Maharashtra, Andhra Pradesh and whole India; whereas in Madhya Pradesh and Tamil Nadu would decrease in 2020 as compared to 2012 figures. The same can be visualized through figure 5.

Modeling and forecasting of urad production

From stationarity test for the production series of urad, it is observed that, all the data series are non-stationary in nature (Table 4). The non-stationary data series are made

Table 4: Test of stationarity of area, production and productivity of urad in India

State	*ADF Value*	*P Value*	*Conclusion*	*KPSS*	*P Value*	*Conclusion*
			Area			
Uttar Pradesh	-1.655	0.742	Non Stationary	1.787	0.010	Non Stationary
Maharashtra	-1.123	0.907	Non Stationary	0.545	0.031	Non Stationary
Andhra Pradesh	-0.141	0.990	Non Stationary	1.540	0.010	Non Stationary
Madhya Pradesh	-1.568	0.743	Non Stationary	1.272	0.010	Non Stationary
Tamil Nadu	-1.915	0.607	Non Stationary	0.845	0.010	Non Stationary
India	-0.852	0.948	Non Stationary	1.200	0.010	Non Stationary
			Production			
Uttar Pradesh	-0.480	0.978	Non Stationary	1.873	0.010	Non Stationary
Maharashtra	-0.204	0.990	Non Stationary	1.389	0.010	Non Stationary
Andhra Pradesh	-0.728	0.959	Non Stationary	1.329	0.010	Non Stationary
Madhya Pradesh	-1.932	0.600	Non Stationary	0.241	0.032	Non Stationary
Tamil Nadu	-1.593	0.733	Non Stationary	0.982	0.010	Non Stationary
India	-0.983	0.928	Non Stationary	1.444	0.010	Non Stationary
			Yield			
Uttar Pradesh	-1.628	0.719	Non Stationary	1.427	0.010	Non Stationary
Maharashtra	-0.732	0.959	Non Stationary	1.456	0.010	Non Stationary
Andhra Pradesh	-1.768	0.664	Non Stationary	0.594	0.023	Non Stationary
Madhya Pradesh	-2.352	0.436	Non Stationary	1.766	0.010	Non Stationary
Tamil Nadu	-0.847	0.949	Non Stationary	0.911	0.010	Non Stationary
India	-1.368	0.821	Non Stationary	1.499	0.010	Non Stationary

Table 5: Best selected ARIMA and GARCH models for area under urad in India.

States	*Models*	*Model Selection Criteria*								*Ljung-Box test for residuals*	
		AIC	*BIC*	*ME*	*RMSE*	*MAE*	*MPE*	*MAPE*	R^2	χ^2	*P value*
Uttar Pradesh	ARIMA(2,1,0)*	298.380	304.820	-0.065	13.181	10.840	-0.556	3.860	0.989	5.633	0.845
	GARCH(1)	331.743	339.931	-1.453	19.565	14.700	5.010	-1.212	0.974	5.633	0.247
Maharashtra	ARIMA(0,1,1)*	341.940	346.850	-0.109	22.276	17.882	-0.068	3.722	0.859	3.500	0.967
	GARCH(1)	361.102	369.290	1.032	25.904	20.541	4.230	0.110	0.779	0.043	0.836
Andhra Pradesh	ARIMA(0,1,2)*	339.970	346.410	0.090	22.074	14.902	0.218	3.720	0.978	3.154	0.977
	GARCH(1)	381.478	389.796	9.002	35.262	27.211	6.783	1.997	0.939	3.635	0.132
Madhya Pradesh	ARIMA(0,1,2)*	316.620	323.060	-0.052	15.953	11.830	0.027	2.027	0.976	8.722	0.559
	GARCH(1)	364.694	372.882	3.575	25.832	21.984	3.725	0.397	0.933	3.602	0.176
Tamil Nadu	ARIMA(0,1,2)*	312.560	319.000	-0.278	15.112	12.089	0.078	5.492	0.959	2.441	0.992
	GARCH(1)	417.927	426.245	1.562	45.750	33.781	15.625	-2.706	0.665	1.383	0.343
India	ARIMA(2,1,2)*	418.670	428.340	-0.129	62.352	48.360	0.064	1.708	0.980	5.256	0.874
	NO GARCH										

Note: * indicates the best model and used further for forecasting purpose.

Table 6: Validation and forecasting of area under urad (000' hectare) in India on the basis of selected best model

States	*Model*	*2010*		*2011*		*2012*		*2016*	*2018*	*2020*
		Observed	*Predicted*	*Observed*	*Predicted*	*Observed*	*Predicted*	*Predicted*	*Predicted*	*Predicted*
Uttar Pradesh	ARIMA(2,1,0)	607.00	532.41	556.00	582.09	555.00	598.58	590.84	633.37	649.34
Maharashtra	ARIMA(0,1,1)	367.19	420.71	482.00	381.83	364.00	382.17	383.51	384.19	384.86
Andhra Pradesh	ARIMA(0,1,2)	429.00	402.46	464.00	434.75	540.00	530.05	483.28	662.20	550.80
Madhya Pradesh	ARIMA(0,1,2)	505.50	498.86	591.70	507.86	557.40	504.74	492.26	486.02	479.77
Tamil Nadu	ARIMA(0,1,2)	259.72	268.89	304.44	261.83	308.26	266.29	284.14	293.06	301.99
India	ARIMA(2,1,2)	2958.08	2841.63	3247.68	2885.65	3215.88	2975.64	3059.46	3122.74	3268.13

stationary by first order differencing. After achieving stationarity, various ARIMA model are tried for each and every series and only best model for each series is selected based on minimum value AIC, BIC, ME, RMSE, MAE, MPE, MAPE and maximum value of R2 and presented in table 8. From the table it can be noted that, ARIMA(3,1,2) for Maharashtra and Madhya Pradesh; ARIMA(3,1,4), ARIMA(2,1,3), ARIMA(0,1,2) and ARIMA(1,1,0) for Uttar Pradesh, Andhra Pradesh, Tamil Nadu and whole India respectively are found to be best fitted ARIMA models for modeling uard production in different states under study. By following same criteria mentioned above, best GARCH models are selected for various

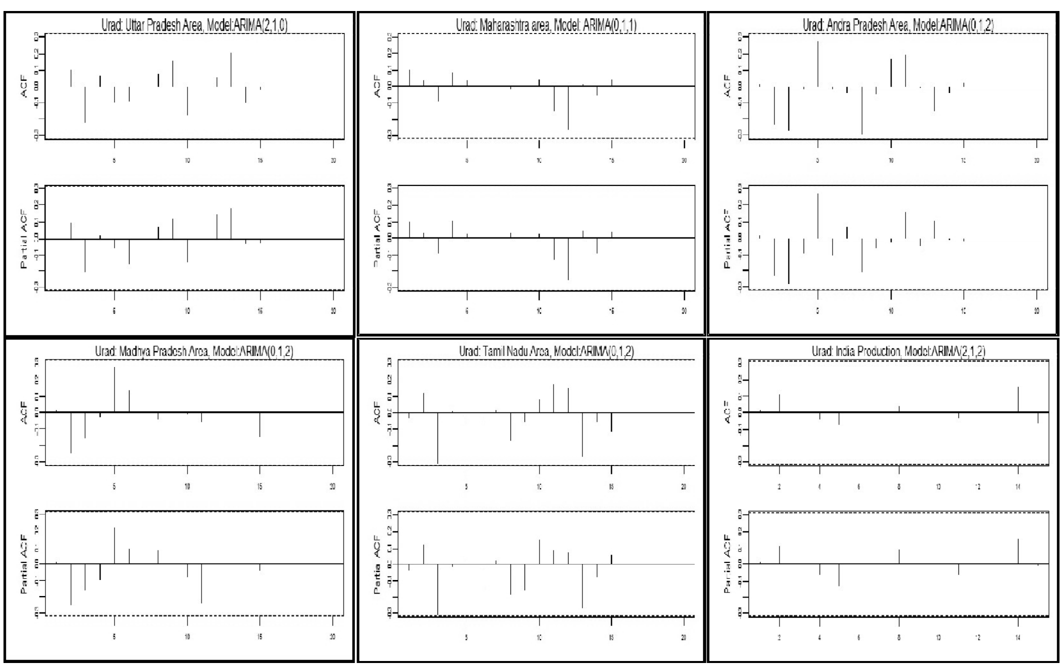

Figure 4: ACF and PACF graphs of residuals for the best fitted models of area under urad in India

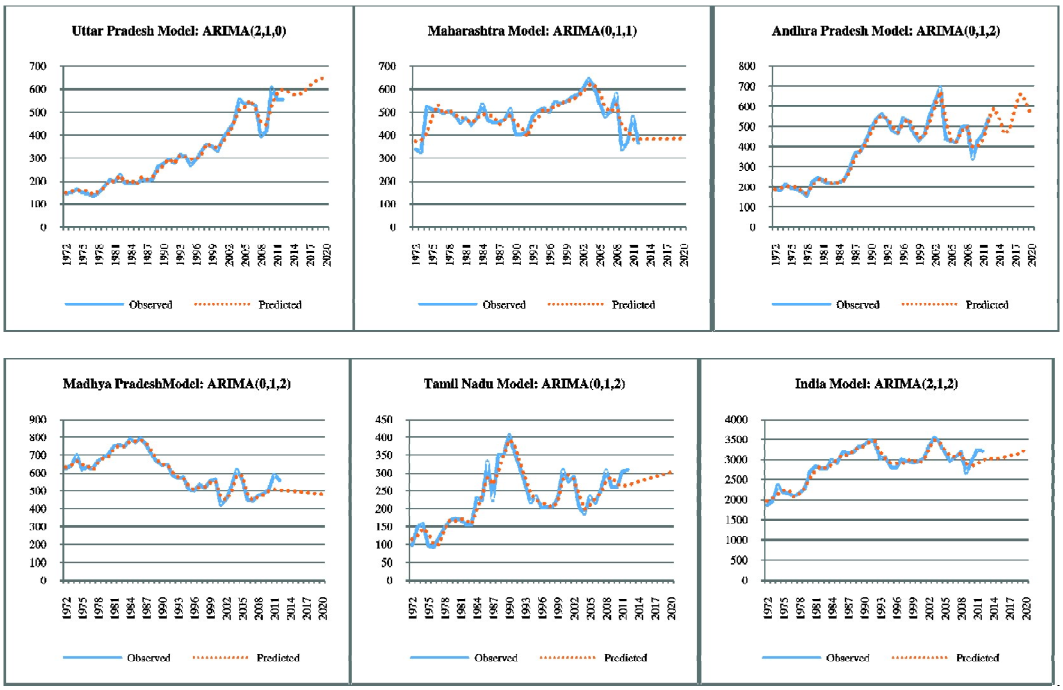

Figure 5: Observed and forecasted area ('000 ha) under urad cultivation using best selected model in India

states under study and found that GARCH(1) as a best fitted GARCH model for modeling urad production in all the major states under study including whole India. In ARIMAx, first all the independent variables which are found to contribute significantly to the productivity of urad crop are modeled and forecasted up to 2020 using ARIMA technique. Then these forecasted values are used as independent variables in the ARIMAx model. As in case of ARIMA and GARCH, here also best ARIMAx model has been selected based on minimum value of error criteria and maximum value of R2 and off course significance of all the coefficients. From the table 8, the best ARIMAx model for modeling the urad production are ARIMAx(3,1,1) for Maharashtra and Madhya Pradesh; ARIMAx(3,1,4), ARIMAx(2,1,3), ARIMAx(0,1,2) and ARIMAx(1,1,0) for Uttar Pradesh, Andhra Pradesh, Tamil Nadu and whole India respectively. The results of Ljung-Box test of residuals also reject the presence of significant auto correlation in the residuals of the best fitted ARIMA, GARCH and ARIMAx model (Table 7).

Best among the best selected ARIMA, GARCH and ARIMAx models is selected based on minimum value of AIC, BIC, ME, RMSE, MAE, MPE, MAPE and maximum value of R2. The best among best selected models for various states under study are presented in table 7. The ARIMAx models are found to best models for modeling urad production in maximum states under study, whereas GARCH model is found to be less efficient model in all the states. These selected are put under diagnostic checking through ACF and PACF graphs of residuals (8) and found that the residuals of selected models are free from significant correlations and used for forecasting uradproduction up to 2020 (Fig. 7). It is clearly visible in the figure that actual and predicted values are almost close to each other in model building stage in all the states expect for Madhya Pradesh and Maharashtra, where the ARIMAx model has failed to catch the lot of variation between two consecutive years throughout the study period. The selected models are also validated for accuracy by using last three years data and observed that the actual and predicted values are in range for all states except Andhra Pradesh (Table 8). The model has failed to catch the sudden change during the 2012 in Andhra Pradesh. From the forecasted figures, it can be seen that urad production would increase marginally in 2020 as compared to 2012 in Maharashtra, Madhya Pradesh and whole India whereas Uttar Pradesh, Andhra Pradesh and Tamil Nadu has the tendency to decrease its production capacity in future. The same can also be visualized through figures.

Table 7: Best fitted ARIMA, GARCH and ARIMAx models for production of urad in India

States	*Models*	*Model Selection Criteria*								*Ljung-Box test for residuals*	
		AIC	*BIC*	*ME*	*RMSE*	*MAE*	*MPE*	*MAPE*	*R2*	χ^2	*P value*
Uttar Pradesh	ARIMA(3,1,4)*	288.620	303.360	0.072	7.725	5.784	-1.981	6.364	0.989	1.876	0.997
	GARCH(1)	296.925	305.113	1.372	11.508	8.222	8.990	-1.822	0.967	0.047	0.828
	ARIMAx(3,1,3)	353.360	368.100	-0.640	20.197	16.208	-8.305	19.138	0.911	2.757	0.987
Maharashtra	ARIMA(3,1,2)	350.600	362.060	0.274	20.821	17.050	0.056	8.465	0.897	0.010	0.920
	GARCH(1)	365.588	373.776	2.610	26.709	21.489	10.613	0.553	0.821	0.444	0.505
	ARIMAx(3,1,2)*	333.390	347.890	0.253	18.559	15.132	0.031	7.530	0.920	2.322	0.993
Andhra Pradesh	ARIMA(2,1,3)	339.730	351.190	0.844	17.817	12.890	0.169	5.929	0.972	10.070	0.434
	GARCH(1)	367.131	375.319	6.711	28.406	22.415	9.864	0.812	0.924	1.112	0.292
	ARIMAx(1,1,2)*	330.440	340.110	0.158	13.485	0.178	6.285	0.564	0.972	11.162	0.345
Madhya Pradesh	ARIMA(3,1,2)	241.840	253.120	0.137	5.225	4.104	0.039	2.370	0.792	2.473	0.991
	GARCH(1)	303.019	311.207	1.345	12.064	9.784	5.733	0.279	0.422	0.000	0.986
	ARIMAx(3,1,2)*	240.960	253.850	0.150	5.005	3.871	0.040	2.236	0.804	5.075	0.886
Tamil Nadu	ARIMA(0,1,2)	278.250	284.800	-0.165	8.777	7.248	0.009	8.945	0.939	8.149	0.614
	GARCH(1)	266.284	274.338	3.006	9.030	6.737	7.202	2.526	0.920	0.237	0.626
	ARIMAx(0,1,4)*	232.010	243.290	0.076	4.499	3.583	0.293	4.246	0.981	10.252	0.419
India	ARIMA(1,1,0)*	386.280	391.110	-0.269	46.684	38.567	0.168	3.257	0.978	3.620	0.963
	GARCH(1)	422.119	430.307	5.199	57.838	45.691	3.803	0.801	0.963	0.780	0.377
	ARIMAx(1,1,0)	407.360	415.420	-0.248	51.080	40.642	0.088	3.334	0.971	0.358	0.550

Note: * indicates the best model and used further for forecasting purpose.

Table 8: Validation and forecasting of urad production (000'tonnes) in India on the basis of selected best model

States	*Model*	*2010*		*2011*		*2012*		*2016*	*2018*	*2020*
		Observed	*Predicted*	*Observed*	*Predicted*	*Observed*	*Predicted*	*Predicted*	*Predicted*	*Predicted*
Uttar Pradesh	ARIMA(3,1,4)	257.00	259.28	372.00	302.62	371.00	306.38	322.07	331.24	356.00
Maharashtra	ARIMAx(3,1,2)	123.42	135.99	329.00	236.09	249.00	245.70	245.57	234.79	252.45
Andhra Pradesh	ARIMAx(1,1,2)	269.00	260.21	253.00	267.37	368.00	271.22	288.55	299.21	307.08
Madhya Pradesh	ARIMAx(3,1,2)	192.10	186.97	230.90	191.86	150.60	191.82	182.83	195.64	191.99
Tamil Nadu	ARIMAx(0,1,4)	98.71	105.69	123.81	121.61	178.81	134.86	155.57	160.47	165.44
India	ARIMA(1,1,0)	1235.79	1237.68	1759.62	1469.90	1766.05	1633.05	1698.18	1771.88	1840.83

Modeling and forecasting of urad productivity

From the stationarity tests for the series of urad productivity, both ADF and KPSS test rejects the hypotheses of stationarity (Table 4) for all productivity series of urad under study. First order differencing was necessary to make all the series stationary. After achieving stationary, we proceeds in similar way as in case of production and selected best ARIMA, GARCH and ARIMAx models for all the states under study and results of the same is

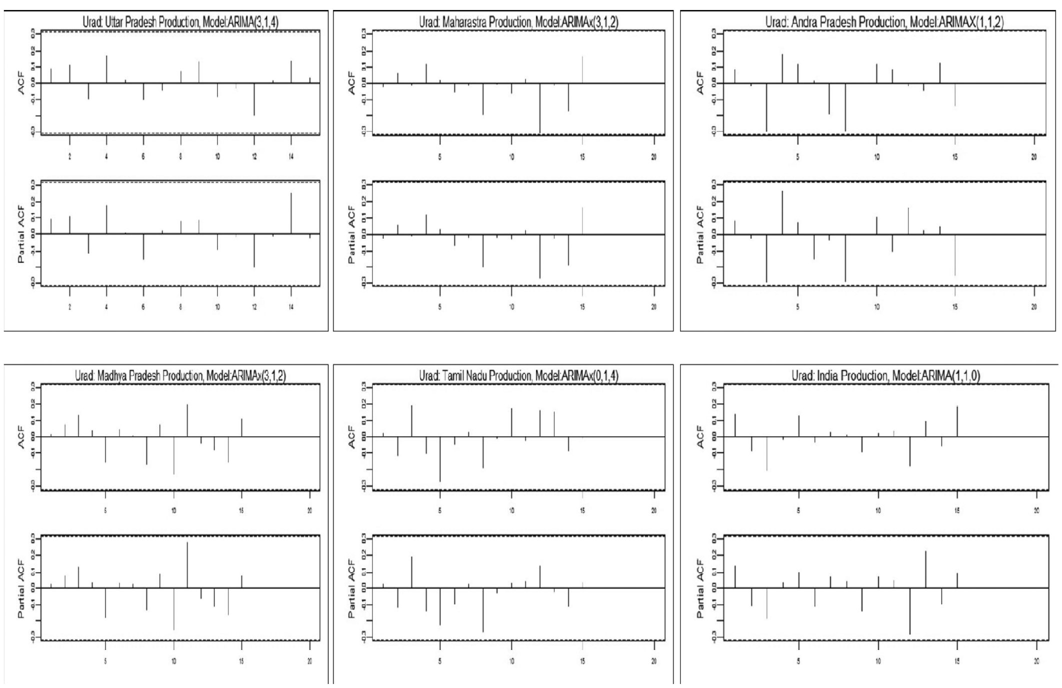

Figure 6: ACF and PACF graphs of residuals for the best fitted models of urad production in India

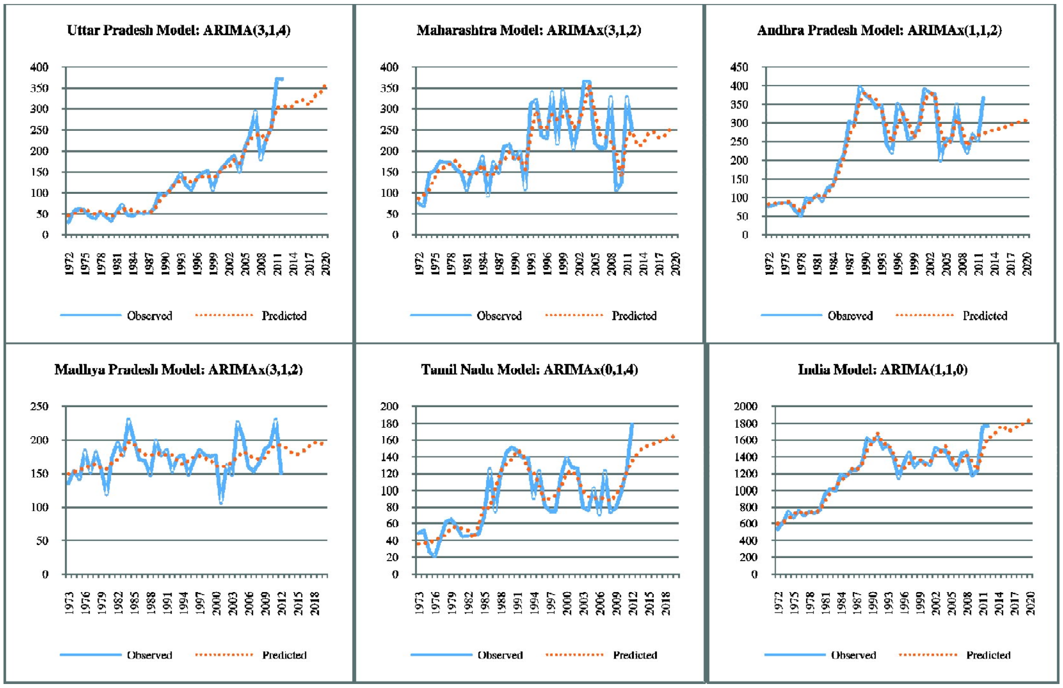

Figure 7: Observed and forecasted urad production ('000 tonnes) using best selected model in India

presented in the table 4.4.4.C1. From the table, the best fitted ARIMA model for modeling urad productivity are found to be ARIMA(1,1,2) for Andhra Pradesh, Madhya Pradesh and Tamil Nadu; ARIMA(3,1,2) for Uttar Pradesh and Maharashtra and ARIMA(0,1,3) for whole India. Among the various GARCH models, GARCH(1) found to be best GARCH model for Andhra Pradesh, Madhya Pradesh and whole India, whereas urad productivity of Uttar Pradesh, Maharashtra and whole India does not have GARCH effect. In similar fashion, best ARIMAx model among various competitive models are also been selected for urad productivity of different states under study. From the table 9, it can be noted that among the various ARIMAx models, ARIMAx(1,1,2) for Tamil Nadu and whole India; ARIMAx(3,1,2), ARIMAx(4,1,2), ARIMAx(2,1,2), ARIMAx(4,1,1) for Uttar Pradesh, Maharashtra, Andhra Pradesh and Madhya Pradesh respectively are found to be best ARIMAx model for modeling productivity of urad in respective states. The residuals of all the best selected models of ARIMA, GARCH and ARIMAx are put under Ljung-Box test (Table 9) and results revealed that there is no presence of significant auto correlation for residuals in the all the cases.

Best among the best selected ARIMA, GARCH and ARIMAx models for every state is selected based on minimum value of error criteria mentioned earlier and maximum value of R2. For modeling urad productivity in all states, ARIMAx model is found to best as compared to ARIMA and GARCH except in case whole India, where ARIMA(0,1,3) remains best model. These selected are put under second stage diagnostic checking through ACF and PACF graphs of residuals (Fig 8) and found that the residuals of selected models are independent. These model used for forecasting urad productivity up to 2020 (Fig. 9& Table 10). From the figure 9, one can see that actual and predicted values in all the states under study are close to each other during the model building stages except for Maharashtra where predicted value deviates from the observed due to presence of lot of variations between years.

The selected best of the best models are also validated by using recent three years data (Table 4.4.4.C2) and found that predicted values are close to actual values, only in case of Madhya Pradesh, where as models in remaining states including whole India fails to catch the sudden changes in urad productivity level during the validation period. The forecasted figures indicate that, urad productivity would in Madhya Pradesh would increase to 315.30 kg per hectare in 2020 as compared to 270.18 kg per hectare in 2012, whereas all other states including whole India would decrease productive capacity in future as compared to 2012. Hence, India needs to augment productivity in urad for nutritional security of its huge population.

Table 4.4.4.C1: Best fitted ARIMA, GARCH and ARIMAx models for productivity of urad in India.

States	*Models*	*Model Selection Criteria*								*Ljung-Box test for residuals*	
		AIC	*BIC*	*ME*	*RMSE*	*MAE*	*MPE*	*MAPE*	*R2*	χ^2	*P value*
Uttar Pradesh	ARIMA(3,1,2)	370.790	382.440	0.465	23.933	17.628	-0.384	4.775	0.924	4.826	0.903
	No GARCH										
	ARIMAx(3,1,2)*	362.640	375.740	0.420	23.577	17.077	-0.335	4.719	0.913	4.549	0.919
Maharashtra	ARIMA(3,1,2)	382.950	394.410	0.204	32.526	26.673	-0.067	6.488	0.891	2.726	0.987
	No GARCH										
	ARIMAx(4,1,2)*	364.460	380.570	-1.105	25.777	19.944	-0.404	5.094	0.932	6.305	0.789
Andhra Pradesh	ARIMA(1,1,2)	373.400	381.630	0.560	29.808	23.295	0.091	4.079	0.948	7.522	0.675
	GARCH(1)	370.912	378.966	13.860	40.087	28.832	4.749	1.782	0.888	0.137	0.711
	ARIMAx(2,1,2)*	372.540	384.010	0.410	29.296	23.302	0.176	4.131	0.952	7.600	0.668
Madhya Pradesh	ARIMA(1,1,2)	302.870	311.060	0.246	11.557	8.913	0.013	3.287	0.946	16.231	0.093
	GARCH(1)	325.796	333.984	-2.875	15.672	12.546	4.738	-1.463	0.891	1.186	0.276
	ARIMAx(4,1,1)*	292.940	306.040	-0.004	9.360	7.759	-0.170	2.791	0.964	3.506	0.361
Tamil Nadu	ARIMA(1,1,2)	354.440	362.620	-0.464	23.682	18.741	-0.121	5.255	0.874	6.418	0.779
	No GARCH										
	ARIMAx(1,1,2)*	354.300	364.130	-0.355	22.547	18.476	-0.115	5.125	0.884	7.719	0.656
India	ARIMA(0,1,3)*	283.060	291.120	0.117	9.844	7.683	0.063	1.920	0.972	6.293	0.790
	GARCH(1)	308.301	316.489	1.487	12.482	10.133	2.524	0.335	0.952	0.030	0.862
	ARIMAx(1,1,2)	300.540	311.820	-0.206	10.598	8.168	-0.023	2.011	0.966	2.014	0.570

Note: * indicates the best model and used further for forecasting purpose.

Table 4.4.4.C2: Validation and forecasting of urad productivity (kg/hectare) in India on the basis of selected best model

States	*Model*	*2010*		*2011*		*2012*		*2016*	*2018*	*2020*
		Obser-ved	*Predi-cted*	*Obser-ved*	*Predi-cted*	*Obser-ved*	*Predi-cted*	*Predi-cted*	*Predi-cted*	*Predi-cted*
Uttar Pradesh	ARIMAx (3,1,2)	423.39	500.26	669.06	549.41	668.47	595.41	597.84	593.97	612.13
Maharashtra	ARIMAx (4,1,2)	336.12	343.72	682.57	309.14	684.07	334.41	349.77	422.21	402.01
Andhra Pradesh	ARIMAx (2,1,2)	627.04	629.77	545.26	565.56	681.48	579.50	638.42	651.47	673.80
Madhya Pradesh	ARIMAx (4,1,1)	380.02	396.95	390.23	291.01	270.18	301.46	315.15	313.31	315.30
Tamil Nadu	ARIMAx (1,1,2)	380.06	333.39	406.68	380.72	580.06	406.90	411.64	417.58	423.85
India	ARIMA (0,1,3)	417.77	434.82	541.81	464.31	549.17	461.87	479.86	487.73	495.54

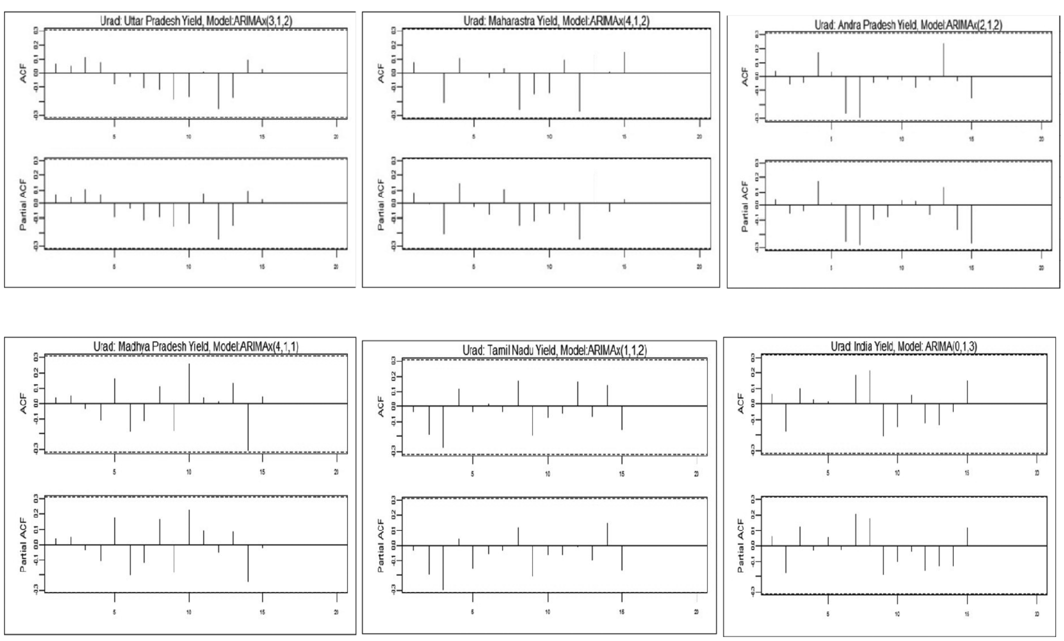

Figure 8: ACF and PACF graphs of residuals for the best fitted models of urad productivity in India

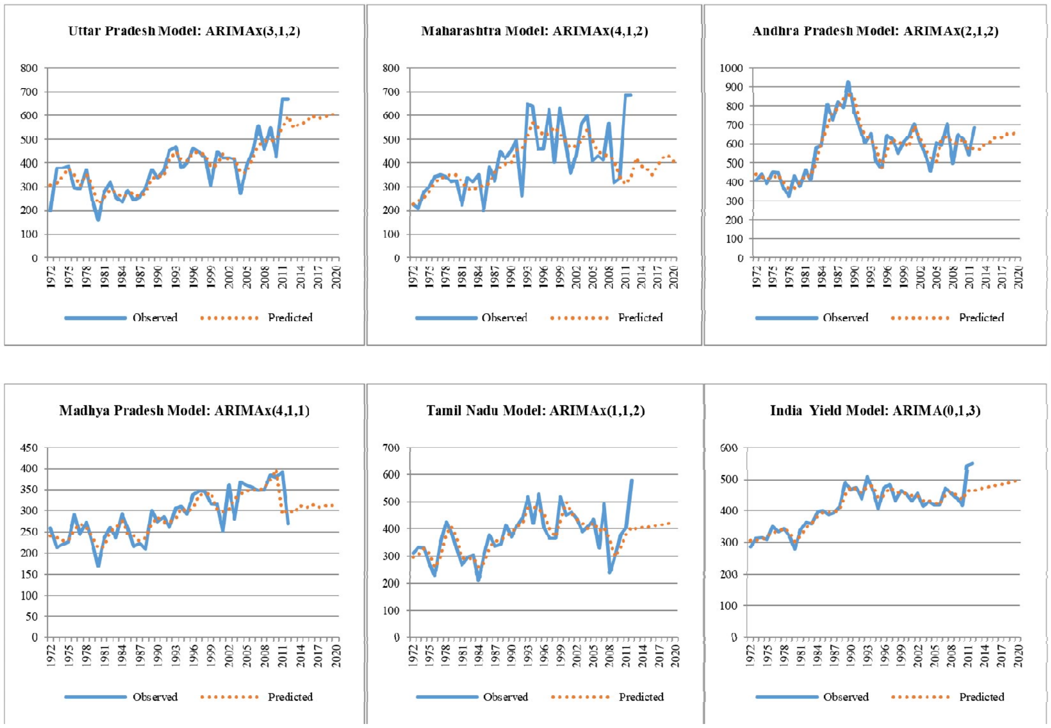

Figure 9: Observed and forecasted urad productivity (kg per hectare) using best selected model in India

Thus, from the study of modeling and forecasting of area, production and productivity of urad in major growing states and whole India following findings emerged out:

1. The area, production and productivity series of urad for all selected states and whole India, none of the series is found stationary and hence first order differencing is done to achieve stationarity.
2. In case of area under urad cultivation, ARIMA model outperformed GARCH in all the states including whole India.
3. Inclusion of independent variable has improved the model efficiency in modeling urad production and productivity in maximum cases.
4. The forecasted figures indicate that area under urad would increase marginally in Uttar Pradesh, Maharashtra, Andhra Pradesh and whole India; whereas in Madhya Pradesh and Tamil Nadu area would decrease in 2020 as compared to 2012 figures.
5. The forecasted figures for urad production would increase marginally in 2020 as compared to 2012 in Maharashtra, Madhya Pradesh and whole India whereas Uttar Pradesh, Andhra Pradesh and Tamil Nadu have the tendency to decrease its production in future.

6. The forecasted figures indicate that, urad productivity in Madhya Pradesh would increase to 315.30 kg per hectare in 2020 as compared to 270.18 kg per hectare in 2012, whereas all other states including whole India would decrease productive capacity in future as compared to 2012.

Conclusion

The above discussion highlighted the fact that the production scenario of urad is concerned, Madhya Pradesh ranks first in average area under urad but Andhra Pradesh was the highest urad producing associated with highest productivity during the study period. Area in case of Madhya Pradesh; production and productivity in maximum states have changed randomly during the study period. Area under urad in Uttar Pradesh, Madhya Pradesh and Tamil Nadu has increasing trend whereas production and productivity of urad in all the states including whole India except Uttar Pradesh and Madhya Pradesh have shown decreasing trend during recent period under study.

All the data series of uradare non-stationary and hence first order differencing is made to achieve stationarity before modeling. For modeling and forecasting area under urad, ARIMA model outperformed GARCH model in all the states under study, whereas inclusion of auxiliary variables improve the model accuracy for production and productivity in maximum cases. The forecasted figures indicate that, urad productivity would in Madhya Pradesh would increase to 315.30 kg per hectare in 2020 as compared to 270.18 kg per hectare in 2012, whereas all other states including whole India would decrease productive capacity in future as compared to 2012.

Reference

Jain, C. L. (2012). Present scenario and future prospectus of rabi pulses production in Chhattishargh. Manual of model training course on production techniques in rabipulses. Pulished by directorate of extension services. Indira Gandhi KrishiViswavidhyalaya. pp 7 to 13.

Rahman, N.M.F., M.A. Aziz, M.M. Rahman and N. Mohammad (2013). Modeling on Grass Pea and Mung Bean Pulse production in Bangladesh using ARIMA model. *IOSR Journal of Agriculture and Veterinary Science,* 6(1), 20-31.

Sahu, P.K. (2006). Forecasting yield behaviour of potato, mustard, rice and wheat under irrigation. *J. Veg. Sci.,* 12,81-99.

Vishwajith, K.P., B.S. Dhekale, P.K. Sahu, P. Mishra and Md. Noman (2014). Time series modelling and forecasting of pulses production in India. *J. Crop Weed.,* 10, 147-154.

Vishwajith, K.P., P.K. Sahu, B.S. Dhekale, Md. Noman and P Mishra (2016). Comparison of ARIMA, ARIMAx and GARCH model for forecasting the gram production scenario of India. RASHI, 1(1), 24-38

Yaziz, S.R., M.H. Ahmad, L.C. Nian & N. Muhammad (2011). A comparative study on Box-Jenkins and Garch models in forecasting crude oil prices. *Journal of Applied Sciences,* 11(7), 1129-1135.