Available online at www.ijpab.com

DOI: http://dx.doi.org/10.18782/2320-7051.7133

ISSN: 2320 – 7051

Int. J. Pure App. Biosci. **7 (1):** 333-341 (2019)

Research Article

International Journal of Pure & Applied Bioscience

Price Forecasting of Cluster Bean using Autoregressive Integrated Moving Average Model: Case of Hanumangarh of Rajasthan in India

Vijay Kumar Baldodiya[1*] and P. K. Awasthi[2]

[1]Ph.D. Scholar, [2]Professor,

Department of Agricultural Economics & F. M. College of Agriculture, JNKVV, Jabalpur

*Corresponding Author E-mail: vj071088@gmail.com

Received: 13.12.2018 | Revised: 20.01.2019 | Accepted: 27.01.2019

ABSTRACT

The price behaviour of a commodity plays crucial role in farm level crop production planning. In this paper, an attempt has been made to forecast Cluster bean price using statistical time-series modelling techniques- Autoregressive Integrated Moving Average (ARIMA) Models. The forecasting performance of these models has been evaluated and compared by using common criteria such as: mean absolute percentage error, Akaike Information Criteria (AIC), Schwarz's Bayesian Information criterion (SBC) and Theil U Statistics. The data used in this study include monthly wholesale price of cluster bean from January 2004 to December 2017. Among all the models tried, the Box-Jenkins ARIMA model (4, 1, 4) was best fit with least AIC (-232.14), SBC (-200.96) and MAPE (8.04). ARIMA (4, 1, 4) model is constructed based on autocorrelation and partial autocorrelation. Finally, forecasts were made based on the model developed. On validation of the forecasts from these models, ARIMA (4, 1, 4) model performed better than the others for cluster bean in Hanumangarh market. Forecasted price of cluster bean shows that its market prices would be ruling in the range of Rs 3899 to Rs 4793 per quintal during January 2018 to December 2018. Thus, ARIMA model can be used to predict the future price of cluster bean in Hanumangarh market of Rajasthan.

Key words: *Cluster Bean, Crop, Income, Agricultural export*

INTRODUCTION

India is the world's leading producer of guar bean (cluster bean) largely cultivated in semi-arid tracts of Rajasthan under rain fed conditions. Development of international markets in recent years have transformed guar from a low value crop grown on marginal land to one that can generate substantial income for farmers and other stakeholders. India export of guar products increased rapidly from a mere US$ 84.70 million in 2001-02 to US$ 3930 million in 2012-13 becoming the largest agricultural export in the country. However, two years later the value of India's exports of guar products plunged to US$ 1574 owing to a unit price decline[1]. Volatility in guar prices have been a source of uncertainty and confusion among stakeholders regarding the long term outlook for the product.

Cite this article: Baldodiya, V.K. and Awasthi, P.K., Price Forecasting of Cluster Bean using Autoregressive Integrated Moving Average Model: Case of Hanumangarh of Rajasthan in India, *Int. J. Pure App. Biosci.* **7(1):** 333-341 (2019). doi: http://dx.doi.org/10.18782/2320-7051.7133

In India the production of Guar was 2.75 million metric tons with an area of 5.58 million ha and productivity level of 492 Kg/ha. Rajasthan is the single largest producer and alone contributed for more than 70 percent of India's total output and Gujarat, Haryana, Punjab and Uttar Pradesh contribute for the rest. Rajasthan occupied an area of 4.78 million ha with a production of 2.22 million metric tons and productivity of 465 Kg/ha. Bikaner is a potential guar producing district in the state producing 0.199 million metric tons with an area of 0.53 million ha and productivity level of 376 Kg/ha (Agricultural Statistics 2015-16, Rajasthan). Considering the fact that cluster bean continues to be an important export crop for India and its productivity has shown upward trend, it is an ideal export crop to expand the export of the country. Also there is need to develop a strong market support system through market intelligence. The returns to the producer farmer are not only governed by production but prices at which the produce is marketed. The prices of cluster bean fluctuate to a great extent mainly because of its supply. Thus, the price forecast may help producers in acreage allocation and time of sale. Sowing time of cluster bean is in between start of June to end of July in Rajasthan. The peak time for arrival is November but it starts in small quantity by October. Though gradual, arrivals continue round the year. Thus, the present study was an attempt to identify the best suited model of price forecasting for cluster bean in Hanumangarh market of Rajasthan. Series analysis is the art of saying that what will happen in the future rather than why. Now days there are various forecasting models in use. Forecaster can choose his own method of forecasting based on his knowledge and available external information. As the process goes on, this procedure can be modified to meet the conditions and to satisfy the current situation. Different forecasting models may be fitted more or less equally well to the data, but they forecasts different future values. Thus, the present study was an attempt to identify the best suited model of price forecasting for cluster bean in Hanumangarh market of Rajasthan.

MATERIAL AND METHODS

As the aim of the study was to forecast prices of cluster bean, various forecasting techniques were considered for use. ARIMA model, introduced by Box and Jenkins[4], was frequently used for discovering the pattern and predicting the future values of the time series data. Hanumangarh Market was purposively selected for the price forecasting study of cluster bean on the basis of second highest cluster bean arrivals. The various price forecasting Autoregressive Integrated Moving Average (ARIMA) models were tried to identify the best-fit model for market price of cluster bean. The secondary data of monthly wholesale cluster bean prices were collected for the study from the AGMARKNET. The data of cluster bean price in Hanumangarh market for the period from January 2004 to December 2017 was utilized for model fitting and data from January 2017 to December 2017 were used for validation. The details of various price forecasting ARIMA models are as follows:

ARIMA Models

Box-Jenkins (ARIMA) model was used to measure the relationships existing among the observations within the series. Box-Jenkins time series model written as ARIMA (p, d, q) was first popularized by Box-Jenkins (1976). The acronym ARIMA stands for "Auto-Regressive Integrated Moving Average". Lags of the differenced series appearing in the forecasting equation are called "auto-regressive (AR)" terms, lags of the forecast errors are called "moving average (MA)" terms and a time series, which need to be differenced to make the data stationary, is said to be an "integrated" version of a stationary series. This model amalgamates three types of processes, namely, auto-regressive of order p, d is differencing to make to the series stationary of order d and moving average of order q.

This method applied only to a stationary time series data and if the data are non-stationary, it has to be brought into stationary by method of differencing i.e.

$W_t = Yt – Y_{t-1}$

The series Wt is called the first difference of Yt and the second difference of the series is $Vt=Wt – W_{t-1}$. The ARIMA modeling consists of following four operational steps.

Identification of the model is concerned with deciding appropriate value of p, d, q,

P, D and Q, where,

p= order of the non-seasonal AR terms

d= non-seasonal differencing

q= order of the non- seasonal MA terms

P= order of the seasonal AR term

D= seasonal differencing

Q= order of the seasonal MA term

Before identifying the model, identify the characteristics of time series for stationarity and seasonality as said above and the same must be removed. After the time series has been stationarized by differencing, the next step in fitting an ARIMA model is to determine whether auto regressive (AR) or moving average (MA) terms are needed to correct any autocorrelation that remains in the differenced series. Of course, with software like Stat graphics you could just try same different combinations of terms and see what works the best. But there is a more systematic way to do this. By looking at the autocorrelation function (ACF) and partial autocorrelation (PACF) plots of the differenced series, we can tentatively identify the numbers of AR and MA terms that are needed. ACF plot is merely a bar chart of the coefficients of correlation between a time series and lag of itself. The PACF plot is a plot of the partial correlation coefficients between the series and lags of itself. After identifying the suitable model, the next step is to obtain the minimum the squares. Estimating the parameters of the Box-Jenkins model is a quite complicated non-linear estimation problem. For this reason, using much commercial statistical software like SAS, SPSS, Minitab, etc. was used for the estimation of parameters. After having estimated the parameters of a tentatively identified ARIMA model, it is necessary to do diagnostic checking to verify that the chosen model is adequate. Considerable skills are required to choose the right ARIMA model. Diagnostic checking helps us to identify the differences in the model, so that the model could be subjected to the modification if needed. After satisfying about the adequacy of the model, it can be used for forecasting. One of the reasons for the popularity of ARIMA modeling is its success in forecasting. In many cases, forecasting's obtained by Box-Jenkins method are reliable than those obtained from traditional econometrics modeling.

RESULT AND DISCUSSIONS

ARIMA Model Forecasting:.

1: Identification of the model:

A model was identified for the wholesale prices of cluster bean in Hanumangarh market. The monthly wholesale prices from January-2004 to December-2017 were used to estimate the ACF and PACF. The price data were tested for stationarity. It can be seen from Table 1 and Fig. 1 that the auto correlation function (ACF) declined very slowly from 0.932 to 0.182 and as many ACFs were significantly different from zero and fell outside 95 per cent confidence interval, indicating the price of Cluster bean to be non-stationary. The analysis of partial autocorrelation coefficients of cluster bean prices are depicted in table 5.28.

The graphical presentation of ACF and PACF of table 1 are given in Fig. 1. The partial autocorrelation function (PACF) dropped steeply after the first lag period from 0.932 to 0.015, which also indicated the non-stationarity of the price series.

Table 1: Autocorrelation and Partial Correlation Coefficient of Cluster bean price in Hanumangarh market (Level)

Autocorrelations				Partial Autocorrelation	
Lag	Autocorrelation	Std. Error	Box-Ljung Statistic Value	Value	S.E.
1	0.932	0.076	148.582	0.932	0.077
2	0.815	0.076	262.925	-0.408	0.077
3	0.678	0.076	342.571	-0.102	0.077
4	.0569	0.076	398.923	0.236	0.077
5	0.505	0.076	443.645	0.181	0.077
6	0.466	0.075	481.868	-0.124	0.077
7	0.462	0.075	519.679	0.224	0.077
8	0.455	0.075	556.667	-0.104	0.077
9	0.437	0.075	590.912	-0.096	0.077
10	0.400	0.074	619.762	0.011	0.077
11	0.360	0.074	643.277	0.159	0.077
12	0.318	0.074	661.769	-0.174	0.077
13	0.279	0.074	676.134	0.016	0.077
14	0.240	0.073	686.828	-0.044	0.077
15	0.205	0.073	694.694	0.016	0.077
16	0.182	0.073	700.930	0.015	0.077

The underlying process assumed is independence (white noise)

The table showed that the autocorrelation and partial autocorrelation functions at lag 16 were significantly different from zero and fell outside the 95 % confidence interval. Differencing of price data of Cluster bean was done to make the series stationary. The Augmented Dickey Fuller based unit root test procedure was done to check whether the price series of Cluster bean were stationary or not. From Table 2, it could be inferred that Augmented Dickey Fuller test values are above the critical (1%) given by MacKinnon statistical table at levels implying that the series are non- stationary indicating the existence of unit root. After taking the first difference, the series become stationary which means that the calculated values for the market is less than the critical value (1%) and free from the consequence of unit root.

Table 2: ADF unit root test for prices of Cluster Bean markets of Hanumangarh

ADF test value	Augmented Dickey-Fuller (ADF)	Critical value (1%)
Level	-2.2390(0.0267)	-3.470679
1st difference	-7.484515(0.0000)	

**significant at 1% level

Note: Figures in parentheses indicate Mackinnon (1996) p-value

The results of the Augmented Dickey-Fuller (ADF) unit root test for cluster bean showed that the level data were non-stationary but their first differences were stationary. The graphical presentation of ACF and PACF diagnostics is given in Fig. 2 which also confirms the above results. Hence, the value of d in the ARIMA model was unity (1) because the differencing was carried out only once to arrive at stationary series.

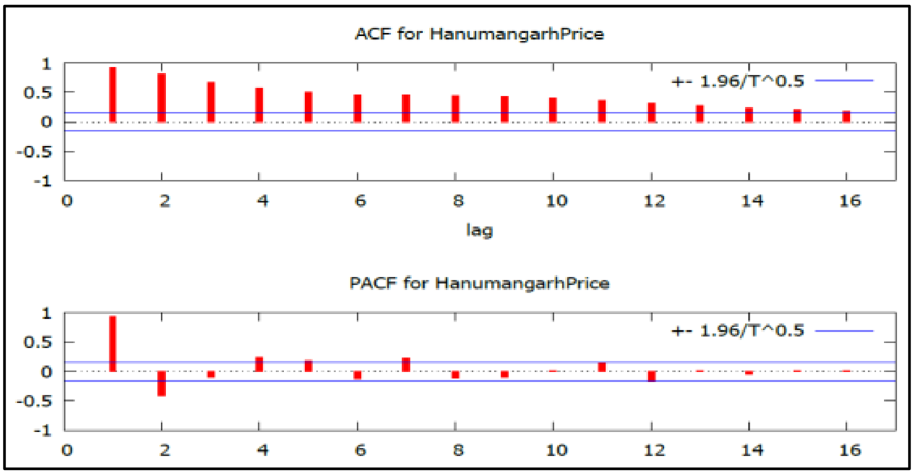

Fig. 1: Autocorrelation and Partial Correlation Coefficient of Guar price in Hanumangarh market

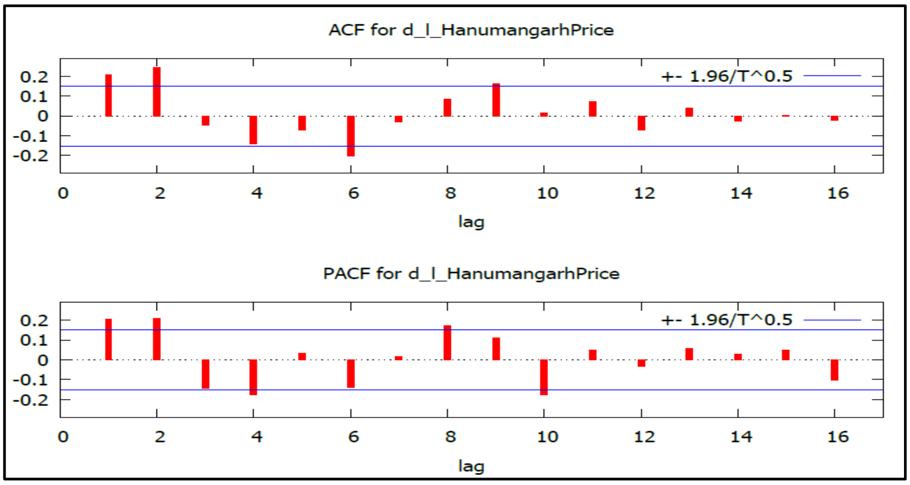

Fig. 2: Autocorrelation and Partial Correlation Coefficient of 1st difference series of Guar price in Hanumangarh market

2: Model estimation:

It was found from the above plots of ACF and PACF (Fig 2), that the values of ACF and PACF were significant with differentiated series. It helps in obtaining the combinations by observing the lags of ACF and PACF, Lag of AR can be worked out through PACF and lag of MA can be worked out through ACF. Most probable Combinations were ARIMA (0,1,2), ARIMA (0,1,1), ARIMA (1,1,1), ARIMA (1,1,2), ARIMA (4,1,4).

Table 3: Probable Combination of models in Hanumangarh Markets

ARIMA (p, d, q)	Logliklihood	Akaike Information Criteria (AIC)	Schwarz Bayesian Criteria (SBC)
0,1,2	113.88	−219.7562	−207.2843
0,1,1	106.05	−206.0946	−196.7406
1,1,1	108.52	−209.0475	−196.5756
1,1,2	114.17	−218.3408	−202.7509
4,1,4	**126.07**	**-232.14**	**-200.96**

These are the most probable combinations but we can test for all the possible combinations up to lag observed and select the best combination for forecasting. The best combination can be selected on the basis of minimum values of Akaike Information Criteria (AIC) and Schwarz Base Criteria (SBC).

On the basis of minimum values of AIC and SBC it was found that AR(4) MA(4) i.e. ARIMA (4, 1, 4) model selected for the forecasting. Therefore, for one step ahead forecast differenced prices for the years 2017 was considered and forecast for the years of 2018.

3: Diagnostic checking:

The model verification is concerned with checking the residuals of the model to see if they contained any systematic pattern which still could be removed in order to improve the chosen ARIMA, which has been done through examining the autocorrelations and partial autocorrelations of the residuals of various orders.

Table 4: Autocorrelation and Partial Correlation Coefficient of residuals of ARIMA(4, 1, 4) model for cluster bean price in Hanumangarh market

Lag	Auto correlation Coefficient		Box-Ljung Statistic	Sig.	Partial Autocorrelation cofficient	
	Value	Std. Errora			Value	Std. Error
1	-0.019	0.077	0.064	0.800	-0.019	0.077
2	-0.017	0.076	0.111	0.946	-0.017	0.077
3	-0.027	0.076	0.235	0.972	-0.027	0.077
4	-0.072	0.076	1.135	0.889	-0.074	0.077
5	0.025	0.076	1.244	0.941	0.021	0.077
6	-0.207	0.076	8.779	0.186	-0.211	0.077
7	-0.060	0.075	9.412	0.224	-0.075	0.077
8	0.135	0.075	12.642	0.125	0.124	0.077
9	0.143	0.075	16.303	0.061	0.146	0.077
10	-0.045	0.075	16.669	0.082	-0.075	0.077
11	0.047	0.074	17.072	0.106	0.057	0.077
12	-0.077	0.074	18.148	0.111	-0.096	0.077
13	0.048	0.074	18.57	0.137	0.032	0.077
14	-0.015	0.074	18.61	0.180	0.025	0.077
15	-0.004	0.073	18.612	0.232	0.092	0.077
16	0.000	0.073	18.612	0.289	-0.044	0.077

For this purpose, autocorrelation and partial autocorrelation functions of various orders of the residuals of ARIMA (4, 1, 4) up to 16 lags were computed and are presented in Table 4. The value of ACF and PACF varied from -0.019 to -0.044 which showed that the autocorrelation and partial correlation functions at lag 16 were significantly different from zero and fell within the 95% confidence interval. The graphical presentation of ACF and PACF diagnostics is given in Fig. 3 which also confirms the above results. Thus it could be concluded that the selected ARIMA (4,1,4) model was appropriate for forecasting the price of Guar during the period under study.

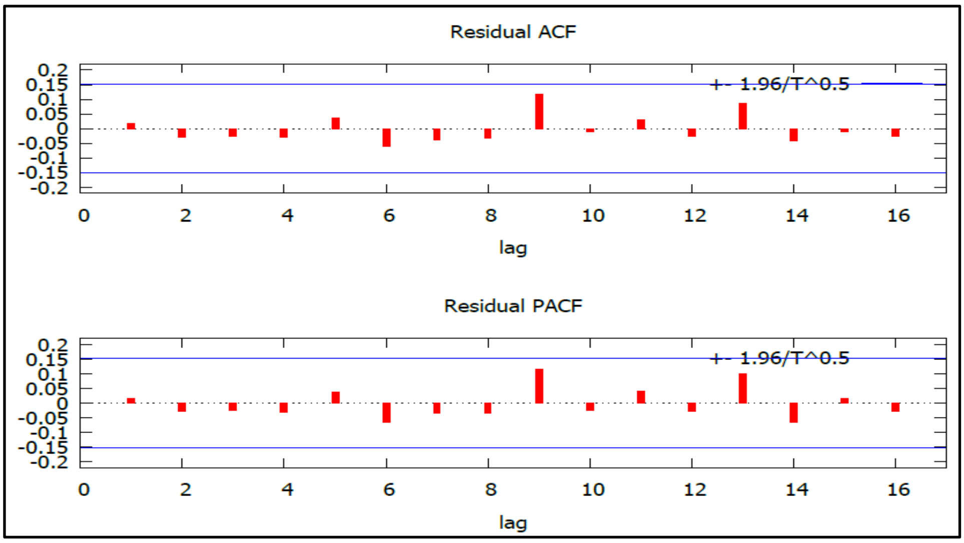

Fig. 3: Autocorrelation and Partial Correlation Coefficient of residuals of ARIMA(4, 1, 4) model for cluster bean price in Hanumangarh market

Table 5: One step ahead Forecasting for prices of the Guar in Hanumangarh market:

Months	Actual Price	Predicted Price
Dec-16	3128.61	
Jan-17	3147.60	3171.22
Feb-17	3145.99	3043.54
Mar-17	3656.06	3387.42
Apr-17	3814.40	3792.84
May-17	3443.53	3811.28
Jun-17	3211.50	3550.33
Jul-17	3162.36	3142.31
Aug-17	3563.64	3015.26
Sep-17	3513.37	3717.54
Oct-17	3484.29	3800.07
Nov-17	3519.50	3445.11
Dec-17	3723.49	3637.99
MAPE(2004-2017)		**8.04**
MAPE(Jan-Dec)		**5.70**
U1		**0.04**
U2		**1.00**

It was found from the table that the model is valid by observing the Mean Absolute Percentage (MAPE) and Theil U statistics. Average MAPE was 8.04 per cent and 5.70 per cent for the year 2017 and Theil U was less than one, which showed that model is best for further forecasting.

Table 6: Forecasting of Cluster bean price:

Months	ARIMA(4,1,4)		
	Forecast	**UCL**	**LCL**
Jan 2018	3899.46	4865.42	3083.91
Feb 2018	3742.00	5229.00	2594.93
Mar 2018	3763.81	6027.61	2202.41
Apr 2018	4018.68	7091.53	2065.79
May 2018	4114.42	7875.00	1885.85
Jun 2018	4221.03	8677.89	1740.04
Jul 2018	4406.52	9477.08	1694.47
Aug 2018	4423.44	9859.73	1607.22
Sep 2018	4425.40	10277.43	1503.89
Oct 2018	4563.30	11041.46	1447.56
Nov 2018	4680.48	11793.62	1384.13
Dec 2018	4792.68	12586.97	1316.28

Table 6 shows out a sample of forecasted cluster bean prices in the coming months of 2018 in Hanumangarh. The forecasted values are the point predicted values in the months. The upper and lower boundary values indicates, our expectation of cluster bean prices not to fall below the lower boundary or go above the upper boundary for the coming four months.

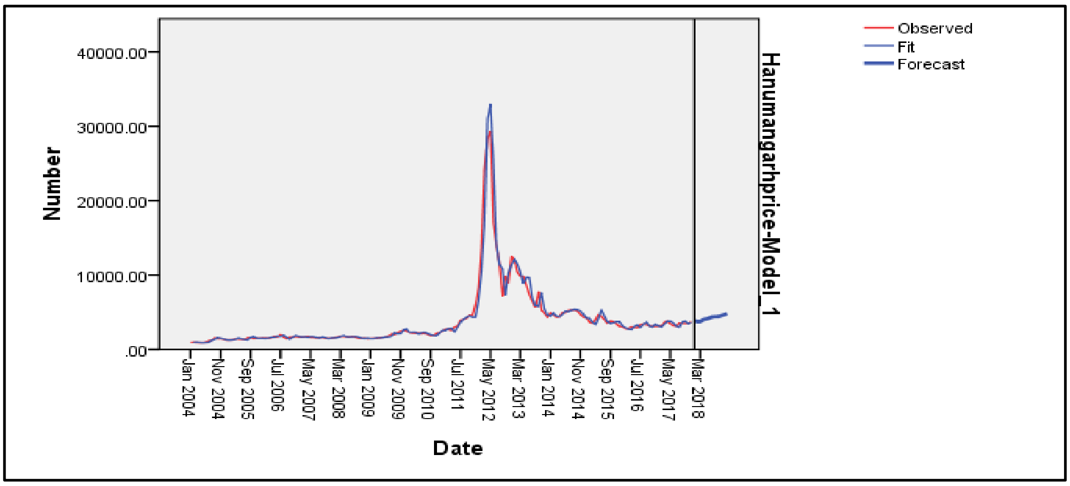

Fig. 5.30: Forecasting of Cluster bean price in Hanumangarh market:

CONCLUSION

In the present study the best fitted model was ARIMA (4,1,4).On comparing the alternative models, it was observed that AIC (-232.14), SBC (-200.96) and MAPE (8.04) were least for ARIMA (4, 1, 4) model was considered the most representative model for the price of cluster bean in Hanumangarh market of Rajasthan. The validity of the forecasted values can be checked when the data for the lead periods become available. The developed model can be used as a policy instrument of the producers and sellers. The limitation of the ARIMA model is that it requires a long time series data. Like any other method, this technique also does not guarantee perfect forecasts. Nevertheless, it can be successfully used for forecasting long time series data.

REFERENCES

1. A.P.E.D.A.,Retrievedfromhttp://agriexchange.apeda.gov.in/IntTrade/TopDestinationAPEDA.aspx on 08/2017 (2015).
2. Directorate of Economics and Statistics www. http://eands.dacnet.nic.in/
3. Commissionerate of Agriculture, Rajasthan. www.krishi.rajasthan.gov.in.
4. Box, G.E.P. & Jenkins, G.M., *Time series analysis: Forecasting and control, 2nd edition. Holden Day*. (1976).
5. Burark, S.S. and Sharma, H., *Agricultural Economics Research Review*, 25 (Conference Number), 530 (2012).
6. Burark, S.S., Pant, D.C., Sharma, H. and Bheel, S., *Indian Journal of Agricultural marketing*, (Conference special) **25(3):** (2011).
7. Karim, M.R., Awal, M.A., Akter, M., *Bangladesh Journal of Agricultural Research*, **(35:1),** 17-28.19 (2010).