Economic Affairs, Vol. 67, No. 04, pp. 491-496, September 2022
DOI: 10.46852/0424-2513.4.2022.14

Research Paper

Forecasting Potato Prices: Application of ARIMA Model

Prakash Singh Badal[1]*, V. Kamalvanshi[1], Abha Goyal[1], Pramod Kumar[2] and Biswajit Mondal[3]

[1]Department of Agricultural Economics, Institute of Agricultural Sciences, Banaras Hindu University, Varanasi, Uttar Pradesh, India
[2]Division of Agricultural Economics, Indian Agricultural Research Institute, New Delhi, India
[3]National Rice Research Institute, Cuttack, Odisha, India

*Corresponding author: psbadal@bhu.ac.in (**ORCID ID:** 0000-0002-8766-3194)

Received: 20-05-2022 **Revised:** 19-08-2022 **Accepted:** 02-09-2022

ABSTRACT

Price fluctuations in potatoes (*Solanum tuberosum* L.) concern consumers, farmers, and policymakers, and its accurate price prediction is important for all the stakeholders. In India, out of a total of 5.34 million ha of land under vegetables, potato occupies nearly 20.8 per cent of area. India produces 12.3 per cent of world potato production (around 45.34 million tons) and is next only to China. The major potato-producing states are highly concentrated in the Indo-gangetic plains of the country. Uttar Pradesh, West Bengal and Bihar account for 32.4, 26.9 and 14.6 per cent of national production of potato. The present study was designed to forecast the prices of potato in these three major potato-producing states of the country. Autoregressive Integrated Moving Average forecasting models - ARIMA (1,0,1) for Varanasi market, ARIMA (2,0,1) for Kolkata market, and the ARIMA (3,0,1) for Patna market were applied. The performance of the ARIMA models produced reliable forecast of prices of potatoes for all three major producing states.

HIGHLIGHTS

- Price forecasting of potato for all three major potato producing states of India for the year 2022.
- Application of Auto Regressive Integrated Moving Average (ARIMA) model for price forecasting.
- Granger Causality test was applied to know the relationship between arrival and prices of potato for all three major potato producing states of India namely Uttar Pradesh, West Bengal, and Bihar.
- Evaluation of forecasted market price of potato for all its three major potato producing states.

Keywords: *Solanum tuberosum*, ARIMA, forecasting, potato, price, RMSE

In some countries Irish Potato (*Solanum tuberosum* L.) is an important staple crop. According to FAO estimates, over 370 million metric tons of potatoes were produced worldwide in 2020. China and India rank first and second, respectively, in production of potato in world, with about one third of the world's potatoes produced in these two countries. Uttar Pradesh, Bihar and West Bengal are the three major potato producing states in India, which collectively contribute 70 per cent to the national potato production (GOI, 2021). Potato is a short duration crop and provides the highest yield per unit area and time. It also has a high potential to prevent hunger and malnutrition due to its high nutritional value. Potato produces more food, energy and protein per unit area and time compared to any other food crop. It also provides excellent opportunity to increase farmers' incomes as they have the potential to produce 5-10 times more than grains, pulses or oilseeds (Singh and Kumar 2013). Violent price fluctuations in potato are a matter of concern among consumers, farmers, and policy makers, and therefore, its accurate price prediction is important for all the stakeholders.

How to cite this article: Badal, P.S., Kamalvanshi, V., Goyal, A., Kumar, P. and Mondal, B. (2022). Forecasting Potato Prices: Application of ARIMA Model. *Econ. Aff.*, **67**(04): 491-496.

Source of Support: None; **Conflict of Interest:** None

Price behavior of potato and its forecasting in different markets has been examined by many researchers using suitable econometric models. Autoregressive Integrated Moving Average (ARIMA) model has been the most applied method for price forecasting (Dhakre and Bhattacharya, 2014). In other studies Seasonal Autoregressive Integrated Moving Average (SARIMA) has given better results depending upon crop and location (Kumar and Baishya, 2020). This paper applied one of the widely used forecasting models, ARIMA using time series data obtained from different publications of the Government of India. The study was done : (1) to assess the correlation between arrivals and prices of potatoes for its three major producing states namely Uttar Pradesh, West Bengal and Bihar, and (2) to forecast and compare the prices of potato for the selected threes states, using appropriate forms of ARIMA.

MATERIALS AND METHODS

The study was undertaken during March 2022, using weekly data on potato prices at Varanasi market for Uttar Pradesh, Kolkata market for West Bengal and Patna market for Bihar. Potato is usually sown in Uttar Pradesh from mid-October to early November and harvested from late February to the first week of March. In West Bengal sowing begins in late October and harvesting is done between late December to early January. In Bihar potato sowing starts from late October to early November and is harvested during February to March. The daily time series data on potato prices were obtained from the websites http://www.nhb.gov.in and agmarknet.gov.in for the period of January 2017 to December 2021. Daily data were transformed into weekly data by averaging of data for available days. The data were analyzed to assess the correlation between arrivals and prices of potato for different states separately and used to fit the best time series model using R programming software.

Correlation analysis was carried out to determine the relationship between arrivals and prices of potato for all the three states. Granger Causality test was performed to show the causal relationship between arrival and prices. The time series ARIMA model is the generalization of Autoregressive Moving Average (ARMA) model. An ARIMA model describes the correlation between data points and takes into account the difference of the values. This model is a combination of Auto-regressive (AR) process, Moving Averages (MA) process and its Integration (I). The ARIMA (p,d,q) model can be represented by the general forecasting equation known as Box-Jenkins equation (Box and Jenkins; 1970).

$$Y_t = \mu + \sum_{i=1}^{p} \varphi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_i$$

Where,

Y_t is price

μ is the mean of series,

φ_1 ……φ_p are the parameters of AR model,

θ_1,…………θ_q are the parameters of the MA model

ε_t, $\varepsilon_{(t-1)}$,………$\varepsilon_{(t-q)}$are the error term.

'p' stands for the order of autoregressive process,

'd' is the order of differencing, and

'q' is the order of moving average process.

Model building of ARIMA for analyzing time series and for forecasting is characterized by the following steps: (a) Identification of model, (b) Estimation of model, (c) Validation or Diagnostic checking, and (d) Forecasting. In the identification stage, the time series data was tested for stationarity using the Augmented Dickey Fuller (ADF) test and the estimated Autocorrelation Functions (ACF) and Partial Autocorrelation Functions (PACF) were compared to find a match. In the present study, data were found to be stationary and there was no need of making it stationary by first order differencing or second order differencing. Atthe estimation stage, the values of autoregressive term (p), moving average term (q) and the order of differencing (d) were estimated with the ACF and PACF graph to fit the model. The next step namely validation was necessary to test the adequacy of the model. Model selection was made based on the lowest values of the criteria namely log likelihood, Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), Schwarz-Bayesian Information Criteria (SBC), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and highest value of R-squared (Badhusha *et al.* 2017). Following the model selection, it was necessary to verify the appropriateness of the selected model with the help of ACF and PACf graph of residuals to verify whether the series of residuals was or random or not. It was found that ACF was effective in measuring the overall adequacy of the chosen model by examining a quantity *Q* known as L-jung-Box statistic. On the basis of L-jung Box statistic, it was found that the residuals of time series data were independently distributed and the model was found adequate.

RESULTS AND DISCUSSION

The Granger Causality test found that there was a positive correlation between market arrival and prices of potato in all three selected markets. However there was no causal relationship found between arrival and prices in West Bengal and Bihar. On the contrary arrival caused price movement in Uttar Pradesh, whereas price movement didn't cause arrival in the state (Table 1).

Augmented Dickey Fuller (ADF) test was applied to check the stationarity of time series data and found that the data was stationary for all three states and first and second order differencing was not required.

For development of this time series model *t series* and *forecast* packages available in R software were used. After analysis three different ARIMA models were generated for the selected three states. ARIMA (1,0,1) for Varanasi (Uttar Pradesh), ARIMA (2,0,1) for Kolkata (West Bengal) and ARIMA (3,0,1) for Patna (Bihar), were found suitable for forecasting. In the figure 1-3 there are total three sub figures, in which the sub figure (a) is the plotting of residual of that particular time series data to see the pattern of residuals; the sub figure (b) is the ACF plot of residuals which indicates the autocorrelation of residuals and the sub figure (c) is the histogram of residuals for the interpretation of normality of residuals. There are three separate residual plots for the three markets in Uttar Pradesh, West Bengal and Bihar (Fig. 1-3).

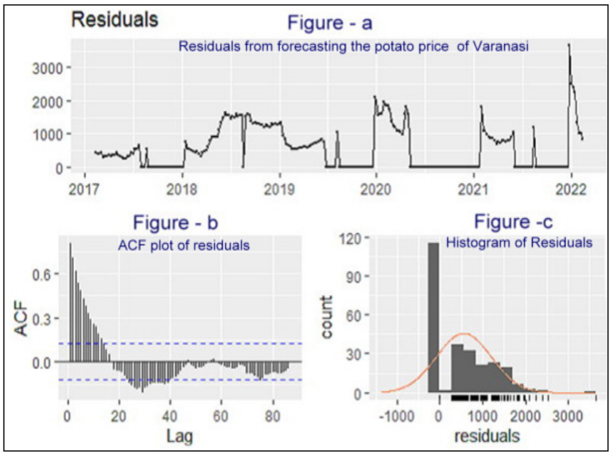

Fig. 1: Residual plot from ARIMA (1,0,1) for Varanasi (Uttar Pradesh)

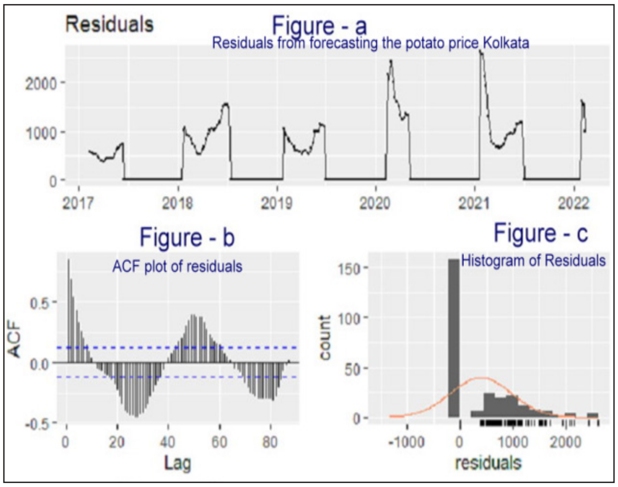

Fig. 2: Residual plot from ARIMA (2,0,1) for Kolkata (West Bengal)

Table 1: The p-value from Granger Causality test and the value of correlation coefficient for the state of Uttar Pradesh, West Bengal and Bihar

Year	Uttar Pradesh (Varanasi market)			West Bengal (Kolkata market)			Bihar (Patna market)		
	P value		Correlation Coefficient (r)	P value		Correlation Coefficient (r)	P value		(r)
	Prices ~ Arrival	Arrival ~ Prices		Prices ~ Arrival	Arrival ~ Prices		Prices ~ Arrival	Arrival ~ Prices	
2017	0.79	0.01	0.66	0.86	0.62	0.63	0.04	0.96	0.59
2018	0.19	-7.23	-0.066	0.30	0.69	0.72	0.02	0.74	0.60
2019	0.60	0.002	0.66	0.46	0.09	0.68	0.008	0.73	0.65
2020	0.89	-4.81	0.95	0.42	0.003	0.75	0.0001	0.49	0.97
2021	0.72	-0.055	0.59	0.29	0.72	0.69	0.186	0.33	0.90
Overall	0.66	-3.12	0.64	0.74	0.12	0.56	15.48	0.53	0.73

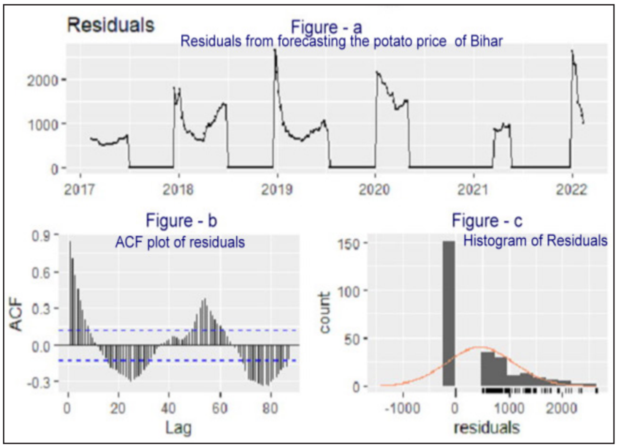

Fig. 3: Residual plot from ARIMA (3,0,1) for Patna (Bihar)

The three fitted models for the selected three states were valid and could be used for forecasting and the residuals were normally and independently distributed as evident from the ACF and PACF residual plots. The results of Ljung-Box Q statistic are given in Fig. 4 through 6. Values of R- squared, AIC, BIC, RMSE and MAE for the models are given in Table 2.

Table 2: Residual analysis of ARIMA (1,0,1), ARIMA (2,0,1), ARIMA (3,0,1) for UP, WB and Bihar, respectively

Sl. No.	ARIMA Model	ARIMA (1,0,1)	ARIMA (2,0,1)	ARIMA (3,0,1)
1	AICs	3364.5	3353.86	3365.13
2	BICs	3378.49	3371.15	3385.89
3	RMSE	304.9275	296.714	301.2455
4	MAE	146.3311	129.5838	128.4233
5	R-squared	0.72	0.73	0.74

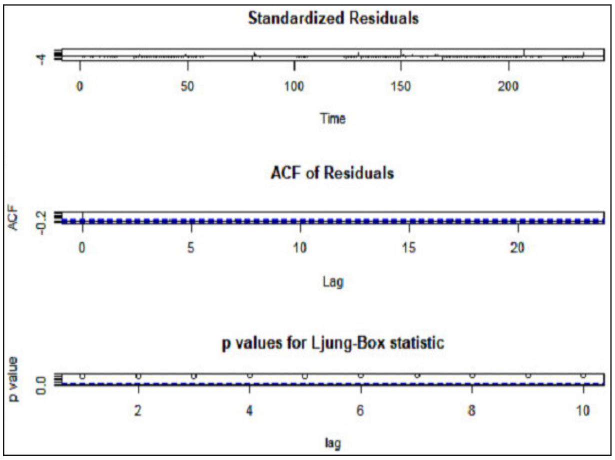

Fig. 4: Graph of Ljung-Box Q statistic for Model ARIMA (1,0,1) For Varanasi (Uttar Pradesh)

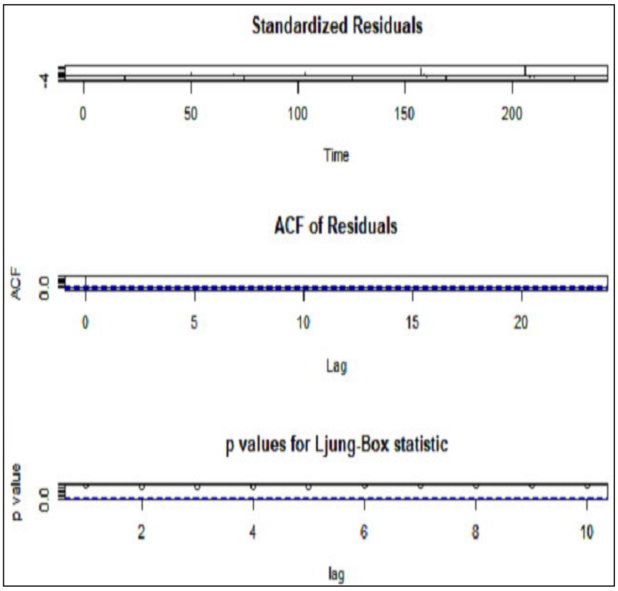

Fig. 5: Graph of Ljung-Box Q statistic for Model ARIMA (2,0,1) for Kolkata (West Bengal)

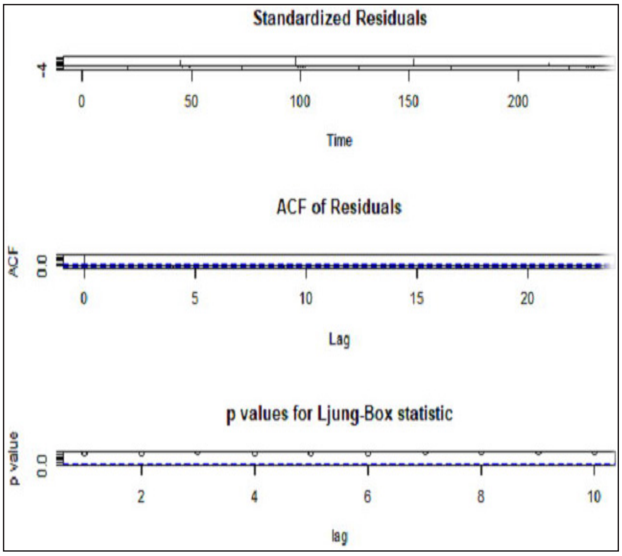

Fig. 6: Graph of Ljung-Box Q statistic for Model ARIMA (3,0,1) for Patna (Bihar)

Weekly forecasted prices for the upcoming month for all states are given in Table 3. In the forecasted plot of potato price (Fig. 7-9), three lines are shown. Black line shows the actual values of potato prices, the red lines indicate the fitted values of prices. These two lines appeared very close to each other which indicated that the fitted values were very close to the actual values. Dark blue line shows the pattern of forecasted price of potatoes in upcoming months of year 2022 in different markets.

Table 3: Forecated price of Uttar Pradesh, West Bengal and Bihar state for the year 2022

Month	Week	Uttar Pradesh (Varanasi market)		West Bengal (Kolkata market)		Bihar (Patna Market)	
		Past price	Forecasted price	Past price	Forecasted price	Past price	Forecasted price
January	1st	963	905	941	70	943	145
January	2nd	904	872	771	138	911	270
January	3rd	867	842	727	200	896	381
January	4th	880	815	676	252	858	475
February	1st	704	790	671	295	829	561
February	2nd	746	768	643	330	813	631
February	3rd	713	748	685	357	848	691
February	4th	793	730	740	377	929	741
March	1st	792	714	832	392	967	782
March	2nd	771	670	913	402	929	814
March	3rd	800	686	935	409	917	839
March	4th	806	674	956	414	950	858
April	1st	895	663	928	416	981	869
April	2nd	1067	654	1010	417	1042	876
April	3rd	0	645	1160	418	1163	878
April	4th	0	637	1159	418	1258	877
November	1st	3660	568	1185	414	2519	510
November	2nd	2520	568	1218	413	2618	504
November	3rd	2392	568	1276	413	2342	448
November	4th	2250	567	1389	414	2300	494
December	1st	1358	567	1493	414	1458	490
December	2nd	1008	566	1610	414	1408	486
December	3rd	1015	566	1551	413	1190	483
December	4th	842	566	978	414	1000	480

**past price of year 2021 in ₹/Quintal; *forecasted price of year 2022 in ₹/Quintal.*

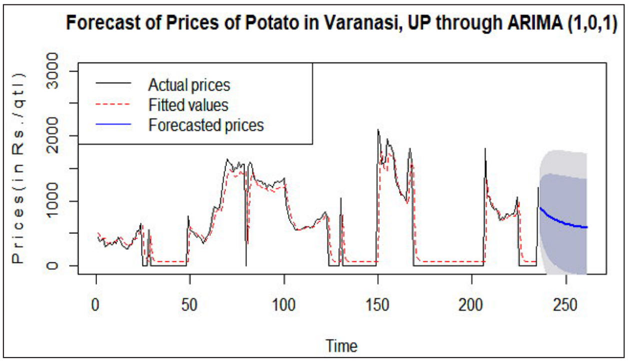

Fig. 7: Forecasting plot of price of potato for Varanasi (Uttar Pradesh)

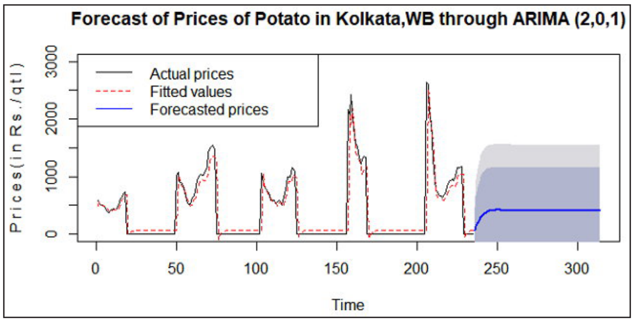

Fig. 8: Forecasting plot of price of potato for Kolkata (West Bengal)

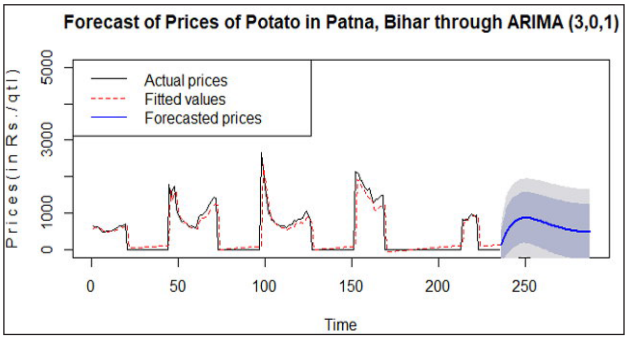

Fig. 9: Forecasting plot of price of potato for Patna (Bihar)

CONCLUSION

Forecasted prices of potato for Uttar Pradesh, West Bengal and Bihar, exhibited different trends for the forecasted months. Prices fluctuated widely over weeks and regions. The forecast plots of potato prices depicted the tentative prices that would prevail in future and how it would be different in all the three states in future. Potato prices were highest in January-February 2022 in Uttar Pradesh; prices peaked in April-May 2022 in Bihar and November-

December 2022 in West Bengal. This study identifies opportune time for inter-state movement of potato in order to stabilize prices.

ACKNOWLEDGEMENTS

Financial support from Indian Council of Agricultural Research- National Agricultural Science Fund (ICAR-NASF), New Delhi is thankfully acknowledged.

REFERENCES

Agricultural Statistics at a glance, 2021. Ministry of Agriculture and Farmer's welfare, Government of India, pp. 165.

Badhusa, M.A.A. and Kaliyaperumal, K.S. 2021. Forecasting national stock price using ARIMA model. *Global and Stochastic Analysis*, **4**(1): 77-81.

Dhyani, B., Kumar, M., Verma, P. and Jain, A. 2020. Stock market forecasting technique using ARIMA model. *Int. J. Recent Techno. and Engine*., **8**(6): 2694-2697.

FAO. 2021. *World Food and Agriculture – Statistical Yearbook 2021.* Rome.

Goyal, A., Badal, P.S., Kamalvanshi, V., Kumar, P. and Mondal, B. 2022. Forecasting of pea prices of Varanasi market Uttar Pradesh, India through ARIMA Model. *Agro Economist- An Int. J*., **9**(1): 49-54.

Gupta, A.K., Singh, A.K. and Rao, V.S. 2018. Forecasting of arrivals and prices of major pulses in Chhattisgarh using ARIMA Models. *The Andhra Agril. J*., **65**(4): 990-993.

Kumar, R.R. and Moumita, B. 2020. Forecasting of Potato prices in India: An application of ARIMA model. *Econ. Aff.*, **65**(4): 473-479.

Kumar, P., Badal, P.S., Paul, R.K.R., Jha, G.K., Venkatesh, P., Kamalvanshi, V., Balasubramaniam, M., Anubukkani, P. and Patel, P. 2021. Forecasting onion price for Varanasi market of Uttar Pradesh, India. *Indian J. Agril. Sci*., **91**(2): 249-253.

Kumar, P., Badal, P.S., Paul, R.K.R., Jha, G.K., Venkatesh, P., Kingsley, I.T., Kamalvanshi, V., Balasubramaniam, M. and Anubukkani, P. 2020. Empowering farmers through future price information: A case of price forecasting of brinjal in eastern Uttar Pradesh. *Indian J. Econ. and Dev.*, **16**(4): 479-488.

Ozer, O.O. and Ilkdogan, U. 2013. The world cotton price forecasting by using the Box-Jenkins model. *J. Tekirdag Agril. Faculty*, **10**(2):13-20.

Sharma, H. 2015. Applicability of ARIMA models in wholesale wheat market of Rajasthan: An investigation. *Econ. Aff.*, **60**: 687- 691.

Shukla, M. and Jharkharia, S. 2011. Applicability of ARIMA models in the wholesale vegetable market, an investigation. *International Conference on Industrial Engineering and Operations Management*. Kuala Lumpur, Malaysia, January, 2011, pp. 22-24.

Singh, D.K., Pynbianglang, K. and Pandey, N.K. 2017. Market arrival and price behaviour of potato in Agra district of Uttar Pradesh. *Econ. Aff.*, **62**: 341-345.

Singh, M.K. and Kumar. A., 2013. Development of potato in Bihar: Issues and strategies. MPRA Paper 51862, revised 03 Nov 2013, Munich, Germany.