https://doi.org/10.56093/ijas.v90i5.104384

Indian Journal of Agricultural Sciences **90** (5): 1009–13, May 2020/Article

Forecasting of onion (*Allium cepa*) price and volatility movements using ARIMAX-GARCH and DCC models

SOURAV GHOSH, K N SINGH, A THANGASAMY, DEBARATI DATTA and ACHAL LAMA*

ICAR-Indian Agricultural Statistics Research Institute, New Delhi 110 012, India

Received: 22 August 2019; Accepted: 04 September 2019

ABSTRACT

In the present investigation an attempt has been made to forecast and understand the volatility transmission in onion prices for three vital markets in Maharashtra, viz. Lasalgaon, Pune and Nagpur. The ARIMAX-GARCH model was employed to estimate mean and volatility among the different markets and also examined the nature of dynamic correlation using the DCC model. The quantity arrival of each market was considered as covariate to improve the mean forecast. We have obtained superior results for ARIMAX-GARCH over ARIMAX model in terms of forecasting. Forecasting efficiency of the models was judged in terms of lower RMSE and MAPE values. Presence of volatility was found in and between the markets as well. Lasalgaon market exhibits highest volatility, whereas the combination of Lasalgaon and Nagpur market experiences the largest volatility movement among them. Identification of interdependency of the markets in terms of volatility movement helps the traders as well as policy makers in a large way. The concerned stakeholders can easily anticipate the prices of other dependent market based on the behaviour of one market.

Key words: ARIMAX-GARCH, DCC model, Onion prices, Volatility transmission

Onion (*Allium cepa*) has universal importance worldwide as a commercial crop and vegetable spices and forms a fundamental part of many diets, both vegetarian and non-vegetarian. India, with an area of 1.26 m ha and annual production of 23.3 m tons, is the second largest onion producing country in the world **(**Anonymous 2019**)**. Maharashtra has the highest area under onion (4.7 m ha) and ranks first in onion production (6.5 m tons) with a production share of 30% **(**Anonymous 2019**)**. Among the various constraints faced by the Indian onion farmers, price instability has been a primary concern affecting the income levels of the farmers as well as the pace of agricultural production. Price instability is determined by a number of factors like yearly variation in production, low price elasticity of demand, seasonality, market inefficiencies, weak supply chains and market. The sudden rise in onion market price impacts both producers and consumers through a "spillover effect" to the other onion markets which leads to high economic inflation (Sinha *et al*. 2018).

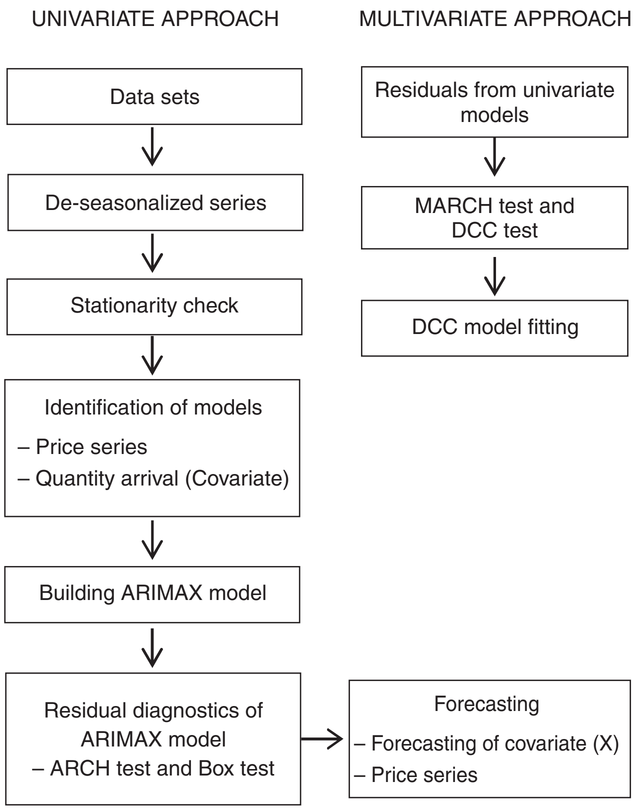
Volatility modelling and estimation in crop yield and agricultural commodity prices can be performed by several applications of Autoregressive Integrated Moving Average Model (ARIMA), Generalized Autoregressive Conditional Heteroskedasticity (GARCH) and its family of models (Paul *et al*. 2014). Lama *et al*. (2015) observed superiority of the GARCH model for modelling domestic and international oil prices and cotton prices highlighted by high forecasting accuracy than corresponding ARIMA model. A new class of multivariate GARCH (MGARCH) model known as Dynamic Conditional Correlation (DCC) model has the flexibility of the univariate GARCH models coupled with parsimonious parametric model for the correlations and can predict the degree of interactions among various volatile commodities and markets (Engle 2002, Lean and Teng 2013, Li and Lin 2015, Lama *et al*. 2016). For the market participants, one of the important tasks is to know about shocks and volatility transmission mechanism which can spread instantaneously from one market to another market for price regulation and policy formulation. In this context it is necessary to know to what extent the prices are being fluctuated and to draw meaningful policy conclusion.

MATERIALS AND METHODS

In this study time series data of onion price (₹/Quintal) and arrival quantity (tonnes) of three important markets of Maharashtra namely Lasalgaon, Pune and Nagpur was collected from the AGMARKNET website (https://agmarknet.gov.in). Out of total 175 data points (January 2005 to July 2019), 169 points (January 2005 to December 2018) were used for model building and remaining 6 points

*Corresponding author e-mail: chllm6@gmail.com

(January 2019 to July 2019) were used for validation purpose. A two stage modelling approached was employed, viz. modelling of the mean dependency of the series using ARIMAX model and thereafter modelling of the residuals using GARCH and MGARCH-DCC models. The GARCH model is capable of capturing volatility in the series (univariate approach), but cannot account for the movement of volatility between series and hence the MGARCH-DCC model was used. The methodology implemented is depicted through the following flowchart below:



ARMAX model: The use of auxiliary variable in the time series model was first studied by Hurvich and Tsai (1989) in ARMA model. They proposed the Autoregressive Moving Average with exogenous variable (ARMAX) model as:

$$\varphi(L)\,(y_t - x'_t\beta) = \theta\,(L)e_t$$

where, $\varphi(L) = (1+\sum_{i=1}^{q}\theta_i L^i$ and θ (L) = $\varepsilon_t = \xi_t h_t^{1/2}$

Here L is the lag operator and x_t is the vector of exogenous variables. The parameter estimation and inferences can be drawn in the same way of ARMA model.

Autoregressive Integrated Moving Average with exogenous variable (ARIMAX) model: Similar to that of ARMA model, exogenous variable (X) can be incorporated in ARIMA model. The ARIMAX model is given by:

$$\varphi(L)\,\Delta^d\,(y_t - x'_t\beta) = \theta\,(L)e_t \quad (1)$$

where $\Delta = (1 - L)$ and d is the order of differencing require for making non-stationary data stationary and the remaining notations are same as previous. In this model auxiliary variable may be of stationary or non-stationary. Parameter estimation and inferences can be drawn in the same way of ARIMA model.

GARCH model: Bollerslev (1986) and Taylor (1986) proposed the GARCH model independently of each other, in which conditional variance is also a linear function of its own lags and has the following form:

$$\varepsilon_t = \xi_t h_t^{1/2}$$

$$h_t = a_0 + \sum_{i=1}^{a} a_i\varepsilon_{t-i}^2 + \sum_{j=1}^{p} b_j h_{t-j} \quad (2)$$

where $\xi \sim N(0,1)$. A sufficient condition for the conditional variance to be positive is

$$a_0 > 0,\ a_i \geq 0,\ i = 1, 2, ..., q.\ b_j \geq 0, j = 1, 2, ..., p$$

The GARCH (p, q) process is weakly stationary if and only if

$$\sum_{i=1}^{q} a_i + \sum_{j=1}^{p} b_j < 1$$

The conditional variance defined by (2) has the property that the unconditional autocorrelation function of ε_t^2 ; if it exists, can decay slowly.

MGARCH – DCC model:

For a multivariate time series $y_t = (y_{1t}, y_{2t},..., y_{Nt})'$ the MGARCH model is given by:

$$y_t = h_t^{1/2}\varepsilon_t \quad (3)$$

where, $H^{1/2}_t$ is N x N positive-definite matrix and of the conditional variance of y_t. N is the number of series and t=1,2,…,n (number of observations). It is with the specification of conditional variance that the MGARCH model changes. The core issues in MGARCH model is to construct the conditional variance-covariance matrix H_t A relatively easy estimation approach is the CCC model introduced by Bollerslev (1990). This model assumes the conditional correlations to be constant. This restriction strongly reduces the number of unknown parameter and thus simplified the estimation. In case of CCC model the H_t represented as follows:

$$H_t = D_t R D_t \quad (4)$$

where, $D_t = diag(h_{11,t}^{\frac{1}{2}},...,h_{NN,t}^{\frac{1}{2}})$ and R is a symmetric positive-definite matrix whose elements are (constant) conditional correlations ρ_{ij}, $i, j = 1, 2,..., N$ ($\rho_{ij} = 1$, $i = j$).

In case of DCC the R matrix is also time varying thus making it dynamic. The representation of the model is as follows:

$$H_t = D_t R_t D_t \quad (5)$$

where, $R_t = \text{diag}\,(Q_t)^{-1/2} Q_t\, \text{diag}\,(Q_t)^{-1/2}$ and

$$Q_t = (1-\alpha-\beta)R + au_{t-1}u'_{t-1} + \beta Q_{t-1} and\ u_t = D_t^{-1}y_t$$

R is the unconditional covariance matrix of u_t. And the conditional covariances are given by:

$$h_{ij,t} = q_{ij,t}\sqrt{h_{ii,t}h_{jj,t}}\ /\ \sqrt{q_{ii,t}q_{jj,t}} \quad (6)$$

Q_t is written as GARCH($1,1$) type equation and then transformed to get R_t.

DCC model was estimated using ML (maximum likelihood) method. The above mentioned methodology is implemented to the data sets using R software.

RESULTS AND DISCUSSION

In the present investigation we have modelled and forecasted the volatile onion prices of three markets namely Lasalgaon, Pune and Nagpur using their arrival quantity as exogenous variables with the help of ARIMAX-GARCH model. Further we studied the movement of volatility between the markets by using MGARCH-DCC model. At the beginning, descriptive statistics of the selected onion market prices were computed. Pune followed by Lasalgaon showed the highest price instability/volatility indicated by high coefficient of variation (79.07% and 77.36% respectively). Among the markets highest price was recorded in Pune (₹ 4704.80/q) in September, 2015 and lowest in Lasalgaon (₹ 215.50/q) in September 2008. Time plots of price (Fig 1) series indicated the presence of seasonality, which was further ascertained by the autocorrelation function (ACF) and partial autocorrelation (PACF) plots. The seasonal factors were thereby extracted using multiplicative model for price series and additive model for the arrival series. Average price across markets were higher in the months of August to January and lower in February to July. Following the basic procedure of time series analysis, we tested both the seasonally adjusted series for their unit root behaviour. All the three price series were found to be stationary after first differencing. On similar pattern, arrival series was also found to be stationary after first differencing. In order to explore the movement of volatility, among the markets, it is important to understand the linear dependency among them. The correlation among the markets with respect to its prices were found to be as high as 0.963 (Lasalgaon/Nagpur) and lowest between Nagpur/Pune (0.935). These high values of correlation strongly suggest the linear dependency among the markets. In order to employ ARIMAX model understanding the behaviour of the covariate series (arrival) is paramount, and in the present study covariate series (arrival) follow ARIMA class of model. In the ARIMAX model, the covariate coefficients for all the three series were found to be significant at 5%. To check the adequacy of the model, residuals obtained after fitting ARIMAX were tested using Box-Pierce and ARCH-LM test. The results of both the test indicated the non-normality and heterocedasticity of the residuals respectively. With confirmation of presence of volatility in all the three series, we proceeded further with GARCH model for analysing the residuals. The estimates obtained from GARCH models revealed that the persistence of volatility was highest in Lasalgaon market with high value of β (0.61) which indicates that that the present price shock will have prolonged impact in the following months. Similar trend was also observed for Pune market ($\beta > \alpha$). But in case of Nagpur, the market seemed to absorb the fluctuation in prices faster as compared to the other two markets as the value of α (0.48) is greater than β (0.38). In the present context, sharing information among the markets is a common phenomenon. Hence, any exploration will paint an incomplete picture of the markets if they are analysed in isolation. Thus to better understand the movement of volatility (price instability) among the markets, we made use of MGARCH-DCC model for the

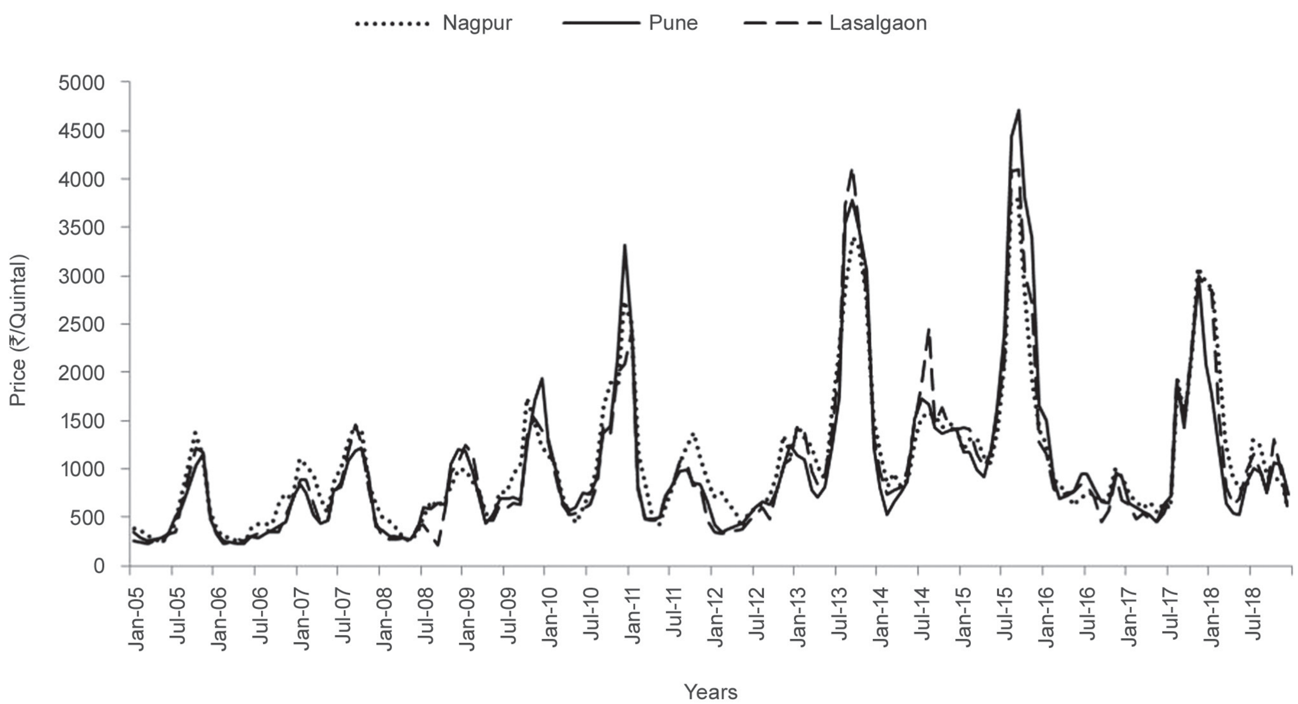


Fig 1 Time plot of onion prices (₹/Quintal) at Nagpur, Pune and Lasalgaon market of Maharashtra.

Table 1 Estimates of DCC model

Estimate	Lasalgaon/Pune	Pune/Nagpur	Lasalgaon/Nagpur
μ_1	0.94	0.95	0.92
μ_2	0.97	0.96	0.97
ω_1	12.50	9.596	1.41
ω_2	7.76	9.432	27.00
α_1	0.43	0.44	0.55
α_2	0.17	0.33	0.50
β_1	0.53	0.50	0.41
β_2	0.68	0.62	0.42
δ_{DCC1}	0.06	0.09	0.12
δ_{DCC2}	0.33	0.38	0.29

All the parameter estimates were found to be significant at 5% level of significance.

Table 2 Forecasts of onion price for different markets

Month (2019)	Lasalgaon			Pune			Nagpur		
	Actual value (₹/q)	ARIMAX	ARIMAX-GARCH	Actual value (₹/q)	ARIMAX	ARIMAX-GARCH	Actual value (₹/q)	ARIMAX	ARIMAX-GARCH
Jan	446.55	403.73	387.89	566.67	589.59	512.27	718.70	1011.23	784.39
Feb	331.21	390.68	383.53	366.67	285.36	309.71	683.56	492.57	615.50
Mar	405.86	626.96	346.32	427.08	252.79	328.60	709.38	594.73	539.88
Apr	548.78	764.72	380.27	590.00	322.16	374.72	839.29	586.07	518.39
May	793.65	969.16	618.64	772.73	455.88	530.81	858.33	446.59	466.22
June	1015.78	1231.48	1140.20	1170.83	978.99	1263.75	1145.65	869.98	1114.14

combination Lasalgaon/Pune, Pune/Nagpur and Lasalgaon/Nagpur markets. Results of DCC model (Table 1) clearly indicate that volatility is higher among the markets and it indicated positive movement of volatility among the markets. This finding implies that fluctuation in prices of one market will trigger fluctuation in the other market as well, for instance an increase in volatility in Lasalgaon market will simultaneously increase the volatility in Pune market and vice versa. Dynamic conditional correlations among the markets were computed to understand the dynamic relationship that exists. The highest dynamic conditional correlation was found between Lasalgaon/Nagpur (0.390) which can be explained by their highest linear dependency (0.963). The time varying correlation coefficients do not vary much among the different considered market combinations.

Another important aspect of time series analysis is to forecast the series with the best identified model. To achieve this objective we forecasted the prices of three onion markets using ARIMAX and ARIMAX-GARCH model. The forecasts obtained for Lasalgaon, Pune and Nagpur markets are reported in Table 2. The values reported are indicative of the fact that ARIMAX-GARCH model have forecasts close to the actual prices. Further, the superiority of ARIMAX-GARCH over ARIMAX was evident with lower RMSE and MAPE values for all the three selected markets. The RMSE values for ARIMAX-GARCH model were 172.34, 146.86 and 221.89 respectively for Lasalgaon, Pune and Nagpur markets. The MAPE values using ARIMAX-GARCH model were 20.03, 20.65 and 21.61 respectively for Lasalgaon, Pune and Nagpur markets.

This study has targeted two vital areas of time series analysis, forecasting price volatility and understanding the movement of it among different markets. From, the present investigations we can infer that onion prices in three major markets of Maharashtra are volatile and they interact among them in terms of volatility movement. Identification of interdependency of the markets in terms of volatility movement helps the traders as well as policy makers in a large way. The concerned stakeholder can easily anticipate the price movement of a market based on the behaviour of other correlated markets. This study can be extended to other volatile agricultural commodities across different markets of India.

ACKNOWLEDGEMENTS

Authors are grateful to ICAR-Indian Agricultural Statistics Research Institute, New Delhi and ICAR-Directorate of Onion and Garlic Research, Pune for providing the necessary facilities.

REFERENCES

Anonymous 2019. Onion statistics 2019. Ministry of Agriculture and Farmers Welfare, Government of India.

Bollerslev T. 1986.Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics* **31**(3): 307–27.

Bollerslev T. 1990. Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *Review of Economic and Statistics* **72**(3): 498–505.

Engle R F. 2002. Dynamic conditional correlation- a simple class of multivariate GARCH models. *Journal of Business and Economic Statistics* **20**(3): 339-50.

Hurvich C M and Tsai C L. 1989. Regression and time series model selection in small samples. *Biometrika* **76**(2): 297–307.

Lama A, Jha G K and Paul R K. 2015. Modelling and forecasting of price volatility: An application of GARCH and EGARCH models. *Agricultural Economics Research Review* **28**(1): 73–82.

Lama A, Jha G K, Gurung B, Paul R K and Sinha K. 2016. VAR-MGARCH models for volatility modelling of pulses prices: An application. *Journal of the Indian Society of Agricultural Statistics* **70**(2): 145–51.

Lean H H and Teng K T. 2013. Integration of world leaders and emerging powers into the Malaysian stock market: A DCC-MGARCH approach. *Economic Modelling* **32**(C): 333–42.

Lin B and Li J. 2015. The spillover effects across natural gas and oil markets: Based on the VEC-MARCH framework. *Applied Energy* **155**(1): 229–41

Paul R K, Ghosh H and Prajneshu. 2014. Development of out-of-sample forecast formulae for ARIMAX-GARCH model and their application. *Journal of the Indian Society of Agricultural Statistics* **68**(1): 85–92.

Sinha K, Panwar S, Alam W, Singh K N, Gurung B, Paul R K and Mukherjee A. 2018. Price volatility spillover of Indian onion markets: A comparative study. *Indian Journal of Agricultural Sciences* **88**(1): 114–20.