Vegetable Science (2016) 43 (1) : 62-69

Study on onion price forecast using time series methods

SP Bhardwaj, Ranjit Kumar Paul and KN Singh

Received: July 2015 / Accepted: May 2016

Abstract

Agricultural production is characterized by risks and uncertainties arising largely due to uncertain yields and relatively low price elasticity of demand, of most commodities. Commodity price movements have a major impact on overall macroeconomic performance. Hence, commodity-price forecasts are a key input to macroeconomic policy planning and formulation. The price volatility in case of Onion is considered to be well known in India. This study has been undertaken to forecast Onion prices before the crop arrival and particularly in the lean periods which witnesses high rise in Onion price. The administration may find enough time period to readjust supply position of Onions in order at avoid high price situation. The study has been illustrated with the time series data on daily Spot price of Onion in Delhi Azadpur Market from 01 January 2009 to 30September 2012. This study was undertaken to obtain a suitable forecast model for forecasting Onion prices. ARIMA (1, 1, 2) model gives reasonable and acceptable forecasts; it does not perform well when there existed volatility in the data series. In this study, GARCH (1, 1) has also been used to forecast prices. The model performs better than ARIMA (0, 1, 1) because of its ability to capture the volatility by the conditional variance of being non-constant throughout the time. Vector Auto Regressive (VAR) a multivariate model for forecast was also attempted but the performance of the model was not improved over GARCH model. The GARCH (1,1) was concluded to be a better model than others in forecasting price of Onion because the values for test statistics calculated using this model were smaller than those calculated using other model and also both the AIC and SIC values from GARCH model were smaller and the percent deviation in forecast price from actual price was comparatively low in GARCH model. Therefore, it showed that GARCH is a better model than ARIMA for estimating daily prices.

Keywords: Price forecast, series methods, onion, ARIMA, GARCH modeal

Indian Agricultural Statistics Research Institute, Library Avenue, Pusa Campus, New Delhi-110012
Emails: bhardwaj@iasri.res.in, ranjitstat@gmail.com

Introduction

Price forecasting has been very important in decision making at all levels and in different sectors of the economy. Agriculture is characterized by risks and uncertainties largely due to uncertain yields and relatively low price elasticity of demand, of most commodities. Decision makers require some information about the future and the likelihood of the possible future outcomes. Price forecasts are critical to market participant for making production and marketing decisions and to policy makers who administer commodity programs and assess the impacts of domestic or international markets. Therefore, commodity price movements have a major impact on overall macroeconomic performance of commodity markets. Hence, commodity-price forecasts are a key input to macroeconomic policy planning and formulation. The price volatility in case of Onion is considered to be notorious one in India.

The literature on price forecasting has focused on two main groups of linear, single-equation, reduced-form econometric models as well as Time Series models. The first group (Financial Models) includes models which are directly inspired by financial economic theory and it is based on the market efficiency hypothesis (MHE), while models belonging to the second group (Structural Models) consider the effects of commodity market agents and real variables on commodity prices. Reza Moghaddasiand et.al (2008) has used annual farm and guaranteed prices of wheat and rice (as a competitive product) and wheat stock for 1966 to 2006 and the findings revealed the superiority of time series models (unit root and ARIMA (3,2,5)) for forecasting of wheat price. ARIMA models outperformed the structural model in predicting the price of wheat for the period 1966-2006. Rangsan Nochai et.al (2006) has studied model of forecasting oil palm price of Thailand in three types of prices as farm price, wholesale price and pure oil price for the period of five years, 2000-2004. The objective of the research was to find an appropriate ARIMA Model for forecasting in three types of oil palm price by considering the minimum of mean absolute percentage error (MAPE). The MAPE for each model was found to be very small.

Chakriya Bowman et.al (2004) assessed the accuracy of a number of alternate measures of forecast performance. The analysis indicated that although judgmental forecasts tend to outperform the model-based forecasts over short horizons of one quarter for several commodities, models incorporating futures prices generally yield superior forecasts over horizons of one year or longer. When evaluating the *ex-post* effectiveness of forecasts, standard statistical measures were commonly used. This research focused primarily on *RMSE*, which gives a measure of the magnitude of the average forecast error, as an effectiveness measure. K. Assis et.al (2010) has compared the forecasting performances of different time series methods for forecasting cocoa bean prices. Four different types of univariate time series methods or models were compared, namely the exponential smoothing, autoregressive integrated moving average (ARIMA), generalized autoregressive conditional heteroskedasticity (GARCH) and the mixed ARIMA/GARCH models. The time series data was became stationary after the first order of differencing. Based on the results of the ex-post forecasting (starting from January until December 2006), the mixed ARIMA/GARCH model outperformed the exponential smoothing, ARIMA and GARCH models. Liew Khim Sen et.al (2007) had taken up time series modeling and forecasting of the Sarawak black pepper price. Their empirical results showed that Autoregressive Moving Average (ARMA) time series models fit the price series well and they have correctly predicted the future trend of the price series within the sample period of study. Guillermo Benavides (2009) examined the volatility accuracy of volatility forecast models for the case of corn and wheat futures price returns. The models applied here were a univariate GARCH, a multivariate ARCH (the BEKK model), an option implied and a composite forecast model. The results showed that the option implied model is superior to the historical models in terms of accuracy and that the composite forecast model was the most accurate one (compared to the alternative models) having the lowest mean-square-errors.

Materials and Methods

The study has been illustrated with the time series data on daily Spot price of Onion in Delhi Azadpur Market from 01 January 2009 to 30 September 2012. The time series properties of commodity prices were assessed by performing unit root tests. Rejection of the null hypothesis of a unit root under the Augmented Dickey Fuller (ADF) test was taken as evidence of statioinarity. The forecasting technique used for a time series analysis that contains a trend or seasonal or non-stationary data was Auto Regressive Integrated Moving Average (ARIMA) which was considered to be most suitable model. The minimum mean absolute percentage errors (MAPEs) of forecasting values were used in selecting an adequate model.

Statioinarity Test or Unit Root Test: The most widely used tests for unit roots are Dickey and Fuller (1979) test and the Augmented Dickey Fuller (ADF) test. Both are used to test the null hypothesis that the series has unit root or non stationary. The DF Test is stated as follows:

$$Y_t = \quad + \rho Y_{t-1} + e_t \qquad \ldots\ldots\ldots\ldots\ldots(1)$$

Where μ and ρ are parameters and e_t is random term. Here the null hypothesis is that H_0 : $\rho = 1$ indicating that the series is non-stationary.

$$\Delta Y_t = \quad + \gamma Y_{t-1} + e_t \qquad \ldots\ldots\ldots\ldots.(2)$$

Where $\gamma = \rho - 1$ & $\Delta Y_t = Y_t - Y_{t-1}$

The null hypothesis is H_0 : $\gamma = 0$. The test can be carried out by performing a τ–test on the estimated γ. The τ statistics under the null hypothesis of a unit root does not follow the conventional t distribution. Dickey and Fuller (1979) showed that distribution under null hypothesis is non standard and simulated critical values for selected sample size. If the error term e_t is auto-correlated, the equation (2) is modified as

$$\Delta Y_t = \mu + \gamma Y_{t-1} + \alpha_i \sum_{i=1}^{m} \Delta y_{t-1} + \epsilon_t \ldots\ldots\ldots\ldots\ldots(3)$$

Where m = number of lagged difference terms required so that the error term ϵ_t is serially independent. The null hypothesis is the same as the DF test, i.e., $H_0 : \gamma = 0$, implying that Y_t is non stationary. When DF test is applied to models like the equation (3), it is called Augmented Dickey Fuller (ADF) test.

Time Series Models: The price forecasts based on these models are only the non-structural-mechanical forecasts. Autoregressive integrated moving average (ARIMA) models are a class of linear models that are capable of representing stationary as well as non-stationary time series. This approach to forecasting is based on Box and Jenkins (1970) popularly known as ARIMA model. The methodology refers to the set of procedures for identifying, fitting, and checking ARIMA models with time series data.

Non-seasonal Box-Jenkins Models for Stationary Series:

(1) A pth –order autoregressive model: AR(p), which has the general form:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots\dots + \alpha_p Y_{t-p} + \varepsilon_t \quad \text{———(4)}$$

Y_t = Response (dependent) variable at time t

$Y_{t-1}, Y_{t-2}, \dots\dots, Y_{t-p}$ = Response variable at time lags t-1, t-2, ……., t-p, respectively.

$\alpha_0, \alpha_1, \alpha_2, \dots\dots, \alpha_p$ = Coefficients to be estimated, ε_t = Error term at time t.

(2) A qth-order moving average model: MA(q), which has the general form:

$$Y_t = \; + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots\dots - \theta_q \varepsilon_{t-q} \quad \text{————(5)}$$

Where, Y_t = Response (dependent) variable at time t

μ = Constant mean of the process, $\theta_1, \theta_2, \dots\dots, \theta_q$ = Coefficients to be estimated

ε_t = Error term at time t., $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots\dots, \varepsilon_{t-q}$ = Errors in previous time periods that are incorporated in the response Y_t.

(3) Autoregressive Moving Average Model: ARMA(p,q), which has general form:

$$Y_t = \psi_0 + \psi_1 Y_{t-1} + \psi_2 Y_{t-2} + \dots\dots + \psi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots\dots - \theta_q \varepsilon_{t-q} \quad \text{————(6)}$$

ARIMA model-building: According to equation (5), a highly useful operator in time-series theory is the lag or backward linear operator (B) defined by $BY_t = Y_{t-1}$

Model for non-seasonal series are called Autoregressive integrated moving average model, denoted by ARIMA (p, d, q). Here p indicates the order of the autoregressive part, d indicates the amount of differencing, and q indicates the order of the moving average part. If the original series is stationary, d = 0 and the ARIMA models reduce to the ARMA models. The difference linear operator (Δ), defined by-

$$\Delta Y_t = Y_t - Y_{t-1} = Y_t - BY_t = (1-B)Y_t \quad \text{————(7)}$$

The stationary series W_t obtained as the dth difference (Δ^d) of Y_t

$$W_t = \Delta^d Y_t = (1-B)^d Y_t, \quad \text{————(8)}$$

ARIMA (p,d,q) has the general form:

$$\psi_p(B)(1-B)^d Y_t = \; + \theta_q(B)\varepsilon_t \text{ or}$$

$$\psi_p(B)W_t = \; + \theta_q(B)\varepsilon_t \quad \text{————(9)}$$

Model Checking: In this step, model must be checked for adequacy by considering the properties of the residuals whether the residuals from an ARIMA model must have the normal distribution and should be random. An overall check of model adequacy is provided by the Ljung-Box Q statistic. The test statistic Q is as follows-

$$Q_m = n(n+2)\sum_{k=1}^{m} \frac{r_k^2(e)}{n-k} \sim X_{m-r}^2$$

Where $r_k(e)$ = the residual autocorrelation at lag k, n= the number of residuals, m= the number of time lags included in the test. If the p-value associated with the Q statistic is small (p-value < α), the model is considered inadequate.

GARCH Method: In econometrics, Auto Regressive Conditional Heteroskedasticity (ARCH) models are used to characterize and model observed time series. They are used whenever there is reason to believe that, at any point in a series, the terms will have a characteristic size, or variance. In particular ARCH models assume the variance of the current error term to be a function of the actual sizes of the previous time periods' error terms: often the variance is related to the squares of the previous innovations. Such models are often called ARCH models (Engle, 1982), although a variety of other acronyms are applied to particular structures of model which have a similar basis. ARCH models are employed commonly in modeling financial time series that exhibit time-varying volatility clustering, i.e. periods of swings followed by periods of relative calm. If an autoregressive moving average model (ARMA model) is assumed for the error variance, the model is a Generalized Autoregressive Conditional Heteroskedasticity (GARCH, Bollerslev (1986)) model.

The GARCH (1, 1) Model Specification**:** To measure the extent of price volatility, GARCH (1, 1) Model has been applied in the study

$$Y_t = X_t' \theta + \epsilon_t \quad \text{————(10)}$$

$$Y_t = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \quad \text{————(11)}$$

The mean equation given in equation (10) is written as a function of exogenous variables with an error term.

Since σ_t^2 is the one-period ahead forecast variance based on past information, it is called the conditional variance. The conditional variance equation specified in equation (11) is a function of three terms: A constant term: ω. , the volatility from the previous period, measured as the lag of the squared residual from the mean equation: ϵ_{t-1}^2 (the ARCH term), Last period's forecast variance: σ_{t-1}^2 (the GARCH term).

The (1, 1) in GARCH (1, 1) refers to the presence of a first-order autoregressive GARCH term (the first term in parentheses) and a first-order moving average ARCH term (the second term in parentheses). An ordinary ARCH model is a special case of a GARCH specification in which there are no lagged forecast variances in the conditional variance equation-*i.e.,* a GARCH (0, 1).

This specification is often interpreted in a financial context, where an agent or trader predicts this period's variance by forming a weighted average of a long term average (the constant), the forecasted variance from last period (the GARCH term), and information about volatility observed in the previous period (the ARCH term). If the asset return was unexpectedly large in either the upward or the downward direction, then the trader will increase the estimate of the variance for the next period. This model is also consistent with the volatility clustering often seen in financial returns data, where large changes in returns are likely to be followed by further large changes. There are two equivalent representations of the variance equation that may aid you in interpreting the model:

(1) If we recursively substitute for the lagged variance on the right hand side of equation(11) we can express the conditional variance as a weighted average of all the lagged squared residuals:

$$\sigma_t^2 = \frac{\omega}{(1-\beta)} + \alpha \sum_{j=1}^{\infty} \beta^{j-1} \epsilon_{t-j}^2 \quad (12)$$

We can see that the GARCH (1, 1) variance specification is analogous to the sample variance, but it down-weights more distant lagged squared errors.

(2) The error in the squared returns is given by $v_t = \epsilon_t^2 - \sigma_t^2$. Substituting for the variance in the variance equation and rearranging terms we can write our model in terms of the errors:

$$\epsilon_t^2 = \omega + (\alpha+\beta)\,\epsilon_{t-1}^2 + v_t - \beta\, v_{t-1} \quad (13)$$

Thus, the squared error follow a hetroskedastic ARMA (1,1) process. The autoregressive root which governs the persistence of volatility shocks is the sum of α plus β. The ARCH parameters corresponds to α and GARCH parameters to β. If the sum of ARCH and GARCH coefficients close to 1, indicating that volatility shocks are quite persistent.

Vector Autoregressive (VAR) process: Let us consider a univariate time series y_t, t=1,2,…,T arising from the model

$$y_t = \upsilon + \Xi_1\, y_{t-1} + \Xi_2\, y_{t-1} + ... + \phi_k y_{t-k} + u_t,\ \ u_t \sim IN(0,\sigma) \quad (14)$$

where, u_t is a sequence of uncorrelated error terms and Ξ_j, j=1,…,k are the constant parameters. This is a sequentially defined model; y_t is generated as a function of its own past values. This is a standard autoregressive framework or AR(k), where k is the order of the autoregression.

If a multiple time series $\mathbf{y}_t$ of n endogenous variables is considered, the extension of (1) will give the VAR(k) model (VAR model of order k), i.e. it is possible to specify the following data generating procedure and model $\mathbf{y}_t$ as an unrestricted VAR involving up to k lags of $\mathbf{y}_t$,

$$y_t = \upsilon + A_1\, y_{t-1} + ... + A_k\, y_{t-k} + u_t,\ \ u_t \sim IN(0,\Sigma) \quad (15)$$

where, $y_t = (y_{1t}, y_{2t}, \ldots, y_{nt})'$ is (n×1) random vector, each of the A_i is an (n×n) matrix of parameters, υ is a fixed (n×1) vector of intercept terms. Finally, $u_t=(u_{1t},u_{2t},\ldots,u_{nt})$ is a n-dimensional white noise or innovation process, i.e., $E(u_t)=0$, $E(u_t u'_t)=\Sigma$ and $E(u_t u'_s)=0$ for s≠t. The covariance matrix Σ is assumed to be non-singular. Using lag operator (L) (2) can be written as, $(I_n - A_1 L - ... - A_k L^k)\, y_t = \upsilon + u_t$.

The process y_t is said to be stable if the roots of the polynomial, $|I_n - A_1 L - ... - A_k L^k| = 0$ lie outside the complex unit circle i.e. have modulus greater than one.

Diagnostic Measures

Information criteria: In statistics, the **Bayesian information criterion** (**BIC**) or **Schwarz criterion** (also **SBC**, **SBIC**) is a criterion for model selection among a finite set of models. It is based, in part, on the likelihood function, and it is closely related to Akaike information criterion (AIC). While fitting a model, it is possible to increase the likelihood by adding parameters, but doing so may result in over fitting. The BIC resolves this problem by introducing a penalty term for the number of parameters in the model. The penalty term is larger in BIC than in AIC. The BIC was developed by Gideon E. Schwarz, who gave a Bayesian argument for adopting it. It is closely related to the Akaike information criterion (AIC). In fact, Akaike was so impressed with Schwarz's Bayesian formalism that he developed his own Bayesian formalism, now often referred to as the ABIC for "a Bayesian Information Criterion" or more casually "Akaike's Bayesian Information Criterion".

The statistical measures of fit called information criteria. Let: **n** = number of observations (e.g. data values, frequencies), **k** = number of parameters to be estimated (e.g. the Normal distribution has 2: mu and sigma), $\mathbf{L}_{max}$ = the maximized value of the log-Likelihood for the estimated model (i.e. fit the parameters by MLE and record the natural log of the Likelihood.)

SIC (Schwarz information criterion, aka Bayesian information criterion BIC)

$$SIC = ln[n]k - 2ln\,[L_{max}]$$

AIC (Akaike information criterion)

$$A_{IC} = \left[\frac{2n}{n-k-1}\right]k - 2\,ln\,[L_{max}]$$

The aim is to find the model with the lowest value of the selected information criterion.

Absolute Accuracy Performance Measures of Forecast: The absolute accuracy analysis is the statistic, mean squared error (MSE), defined as: MSE = $\sum(\hat{y}_t - y_t)^2$, Where y_t and $\hat{y}_t$ are the actual and forecast values, respectively. MSE is considered as a "non-parametric" statistic that indicates the size of the individual forecast errors from actual values. The square root of MSE, called the root mean squared error (RMSE) represents the mean size of forecast error, measured in the same units as the actual values

$$MSE = \sum_{t=T+1}^{T+h}(\hat{y}_t - y_t)^2$$

$$RMSE = \sqrt{\sum_{t=T+1}^{T+h}(\hat{y}_t - y_t)^2}$$

The absolute size of the errors the mean absolute forecast error (MAE) is used:

$$MAE = \sum_{t=T+1}^{T+h}|\hat{y}_t - y_t|/h$$

The RMSE is similar to MAE. The MAE and RMSE depend on the scale of the dependent variable. These should be used as relative measures to compare forecasts for the same series across different models.

The relative mean absolute prediction error (RMAPE) is calculated using the following formula

$$\mathbf{RMAPE = 100}\sum_{t=T+1}^{T+h}\left|\frac{\hat{y}_t - y_t}{y_t}\right|/h$$

The RMAPE calculates the forecast error as a percentage of actual value.

Results and Discussion

Unit Root Test: Augmented Dickey Fuller (ADF) test was applied to the Spot price series data to test the null hypothesis that the series has unit root or non stationary. The results are given in Table-1. The 'τ -Statistics' obtained for all the price series is significant and greater than at 1 percent level, the null hypothesis of series has unit root or non stationary data series cannot be rejected. The alternative hypothesis is true. Thus data series is subjected to first differencing to make the data stationary. The results of differenced series indicated that the 'τ - Statistic' obtained for price series is not significant and less than at 1 percent level, we are bound to reject the null hypothesis and the alternative hypothesis of stationary series and no unit root is true. The data series became stationary at one differencing and the data is now ready for further econometric analysis. In Table - 2 Augmented Dickey Fuller Test for Quantity Arrival of Onion Delhi Market showed that the series is stationary at current level.

Estimation equation of ARIMA (1, 1,2): Model for non-seasonal series are called Autoregressive integrated moving average model, denoted by ARIMA (p, d, q).

Table 1: Augmented Dickey Fuller Test for spot market price of onion Delhi market

	Level Data		At First Difference	
	t-Statistic	Prob.*	t-Statistic	Prob.*
ADF Test value	-3.05347	0.1182	-41.3835	0.00
1% level	-3.96613		-3.96613	
5% level	-3.41377		-3.41377	
10% level	-3.12895		-3.12895	

Table 2: Augmented Dickey Fuller Test for quantity arrival of onion Delhi market

	Level Data	
	t-Statistic	Prob.*
ADF Test value	-6.2229	0.00
1% level	-3.43593	
5% level	-2.86389	
10% level	-2.56807	

Here p indicates the order of the autoregressive part, d indicates the amount of differencing, and q indicates the order of the moving average part. If the original series is stationary, d = 0 and the ARIMA models reduce to the ARMA models. Estimate the parameters for a tentative model has been selected on the basis of significance level of AR and MA terms as given in Table-3. In this particular case both moving average term and autoregressive terms was found statistically significant.

Parameter Estimation GARCH (1, 1) Model: In Table 4, the conditional mean equation, the parameter found is ω = -50.5779 and one statistically significant AR term (-0.07402). While the conditional variance equation gives ω = 169872.4 α_1 = 0.32953and a high value of β_1 = 0.563632A which implied that volatility is persistent and it takes a long time to change.

Table-3 Parameter estimation of ARIMA (1,1,2)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-4.94758	24.11265	-0.20519	0.8375
AR(1)	-0.22063	0.029654	-7.44018	0
MA(2)	-0.06963	0.03033	-2.29573	0.0219
R-squared	0.045312	Mean dependent var		-4.8594
Adjusted R-squared	0.04363	S.D. dependent var		1091.123
S.E. of regression	1067.055	Akaike info criterion		16.78583
Sum squared resid	1.29E+09	Schwarz criterion		16.7991
Log likelihood	-9548.14	Hannan-Quinn criter.		16.79084
F-statistic	26.93512	Durbin-Watson stat		1.994544
Prob(F-statistic)	0			
Inverted AR Roots	-0.22			
Inverted MA Roots	0.26	-0.26		

Table 4: Parameter estimation of GARCH (1, 1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-50.5779	29.88697	-1.69231	0.0906
AR(1)	-0.07402	0.060618	-1.2211	0.222
	Variance Equation			
C	169872.4	15899.79	10.68394	0
RESID(-1)^2	0.32953	0.04783	6.889595	0
GARCH(-1)	0.563632	0.039367	14.31728	0
R-squared	0.022527	Mean dependent var		-4.8594
Adjusted R-squared	0.021666	S.D. dependent var		1091.123
S.E. of regression	1079.238	Akaike info criterion		16.14087
Sum squared resid	1.32E+09	Schwarz criterion		16.163
Log likelihood	-9179.15	Hannan-Quinn criter.		16.14923
F-statistic	6.544996	Durbin-Watson stat		2.262935
Prob(F-statistic)	0.000033			
Inverted AR Roots	-0.07			

Parameter estimation of Vector Autoregressive (VAR) Model: In Table-5 the coefficient of price variable (-0.20834) and for quantity (-0.13796) both the variables used in the model are statistically significant as evident from t value. The lag quantity arrival and lag prices of onion in the mandi influence the forecasts of onion prices to some extent.

Evaluation forecast Performances of forecast Models

Information criterion: The AIC and SIC values are obtained from equation estimation from both ARIMA and GARCH models using E-Views and given in Table-6. We found that both the AIC and SIC values from GARCH model are smaller than that from ARIMA model. Therefore, it shows that GARCH is a better model than other models for estimating daily prices

Forecast Performance: In the forecasting stage, we calculate RMSE, MSE and MAE and RMAPE values from different models. These are tabulated in Table-7. If the actual values and forecast values are closer to each other, a small forecast error will be obtained. Thus,

Table 5: Parameter estimation of Vector Autoregressive (VAR) Model

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	136.0968	64.91865	2.09642	0.0363
DPRICERSPERTONE(-1)	-0.20834	0.029069	-7.16727	0
QTYARRIVTONE(-1)	-0.13796	0.055088	-2.50443	0.0124
R-squared	0.046421	Mean dependent var		-4.8594
Adjusted R-squared	0.044741	S.D. dependent var		1091.123
S.E. of regression	1066.435	Akaike info criterion		16.78466
Sum squared resid	1.29E+09	Schwarz criterion		16.79794
Log likelihood	-9547.47	Hannan-Quinn criter.		16.78968
F-statistic	27.62633	Durbin-Watson stat		2.023037
Prob(F-statistic)	0			

Table 6: Information criterion for different models

Model	AIC	SIC
ARIMA	16.78583	16.7991
GARCH	16.14087	16.163
VAR	16.78466	16.79794

Table 7: Forecast performance of different forecast methods.

Test statitics	Forecast Models	Forecast Days 5	10	15	20	30	45	60
MAE	ARIMA	41.28	39.28	43.12	41.66	60.32	46.21	43.63
	GARCH	64.68	53.52	55.47	54.59	51.55	57.03	53.10
	VAR	80.80	73.70	78.06	78.42	87.37	85.31	78.86
	ARIMA	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	GARCH	0.01	0.00	0.00	0.00	0.00	0.01	0.00
RMAPE	VAR	0.01	0.01	0.01	0.01	0.01	0.01	0.01
MSE	ARIMA	2300.92	2274.23	2653.59	2487.82	8958.32	6138.25	5169.01
	GARCH	4306.60	3203.16	4007.69	3730.37	3606.98	4213.95	3665.52
	VAR	8010.80	7481.28	7992.94	7792.23	12274.14	10901.32	9650.13
RMSE	ARIMA	47.97	47.69	51.51	49.88	94.65	78.35	71.90
	GARCH	65.62	56.60	63.31	61.08	60.06	64.91	60.54
	VAR	89.50	86.49	89.40	88.27	110.79	104.41	98.24

Table 8: Difference in actual and forecast price of onion

Forecast	Actual	Forecast Price in different models			Percent deviations from actual price		
Day	Price	ARIMA	GARCH	VAR	ARIMA	GARCH	VAR
1	7000	6974.15	6945.68	7098.97	0.37	0.78	1.41
2	7080	7051.53	7019.76	6995.67	0.74	0.28	0.06
3	7100	7082.18	7044.20	7082.63	1.17	0.63	1.18
4	7100	7090.59	7045.68	7211.28	1.29	0.65	3.02
5	7040	7045.96	6990.12	7020.20	0.66	0.14	0.29
6	7130	7107.63	7069.02	7068.93	1.54	0.99	0.98
7	7130	7118.11	7075.68	7165.80	1.69	1.08	2.37
8	6900	6943.15	6862.70	6990.24	0.81	1.96	0.14
9	6980	6971.50	6919.76	6899.96	0.41	1.15	1.43
10	6980	6971.39	6925.68	7060.66	0.41	1.06	0.87
11	7170	7121.45	7101.61	7040.87	1.73	1.45	0.58
12	7170	7150.13	7115.68	7116.30	2.14	1.65	1.66
13	7170	7160.58	7115.68	7055.24	2.29	1.65	0.79
14	7170	7162.58	7115.68	7097.88	2.32	1.65	1.40
15	7150	7147.72	7097.16	6981.61	2.11	1.39	0.26
16	7150	7144.84	7095.68	7219.64	2.07	1.37	3.14
17	7380	7323.06	7308.65	7272.66	4.62	4.41	3.90
18	7380	7357.59	7325.68	7335.42	5.11	4.65	4.79
19	6920	7011.49	6899.73	7029.33	0.16	1.43	0.42
20	6830	6874.29	6782.34	6841.01	1.80	3.11	2.27
21	6830	6836.60	6775.68	6804.87	2.33	3.20	2.79
22	6600	6647.79	6562.70	6745.50	5.03	6.25	3.64
23	7090	6992.32	6999.41	6937.09	0.11	0.01	0.90
24	6630	6694.66	6609.73	6710.80	4.36	5.58	4.13
25	6820	6797.27	6751.61	6800.21	2.90	3.55	2.85
26	6800	6789.65	6747.16	6764.50	3.01	3.61	3.36
27	6800	6793.77	6745.68	6801.91	2.95	3.63	2.83
28	6800	6793.24	6745.68	6862.91	2.95	3.63	1.96
29	6520	6575.30	6486.40	6493.66	6.07	7.34	7.23
30	6800	6800.00	6800.00	6800.00	2.86	2.86	2.86

smaller RMSE, MAE and RMAPE values are preferred. Most of the forecast errors from GARCH model are smaller than that from other model. Therefore, we can conclude that GARCH model performs better than other two models. In other words, GARCH is a better forecast model for daily prices of onion.

This study was undertaken to obtain a suitable models for forecasting Onion prices. In this study, the model that has been selected for forecasting onion prices is ARIMA (1, 1, 2). This model gave reasonable and acceptable forecasts; it did not perform very well when there exists volatility in the data series. In this study, GARCH (1, 1) has also been used to forecast prices.

In Vector Autoregressive (VAR) Model, the lag quantity arrival and lag prices of onion in the mandi influence the forecasts of onion prices to some extent. The GARCH(1,1) was concluded to be a better model than other models in forecasting spot price of Onion because the percent deviation in forecast values from actual values were smaller in GARCH model. Also both the AIC and SIC values from GARCH model were smaller than that obtained from other model. Therefore, it showed that GARCH is a better model for estimating daily prices.

सारांश

उत्पादन को जोखिम एवं आनिश्चितताओं से चरित्रीकरण किया जाता है जो मुख्यतः निश्चित उपज तथा मांग के अनुरूप कम कीमत लोच वाली अधिकांश वस्तुएं होती है। वस्तु मूल्य गतिशिलता का मुख्य प्रभाव सम्पूर्ण व व्यापक आर्थिक प्रदर्शन पर पड़ता है। अतः वस्तु मूल्य पुर्वानुमान व्यापक आर्थिक निति योजना तथा कार्यान्वयन का एक महत्वपूर्ण निवेश है। प्याज में मूल्य अस्थिरता भारत में पूरी तरह ज्ञात है। यह अध्ययन प्याज का बाजार में पहुँचने से पूर्व मूल्य निर्धारण करना है तथा विशेषतः कम आमद अवधि जबकि प्याज का ज्यादा मूल्य का साक्षी बनता है। शासन प्रबंध को प्रचुर समय आपूर्ति को समायोजित करने के लिए मिल जाता है जिससे अधिक मूल्य की स्थिति से बचा जा सकता है। अध्ययन समय, शृंखला आंकड़ा के साथ सचित्र प्रतिदिन आजादपुर मार्केट, नई दिल्ली में 1 जनवरी 2009 से 30 सितम्बर 2012 तक लिया गया। एक उपयुक्त पुर्वानुमान आदर्श अध्ययन प्याज कीमत ज्ञात करने के लिए किया गया। ए आर आई एम ए (1,1,2) आदर्श ने उचित तथा स्वीकार्य पूर्वानुमान आंकड़ा क्रम की अस्थिरता की वजह से यह उचित निष्पादन नहीं दिया। इस अध्ययन में जी ए आर सी एच (1,1) का भी उपयोग मूल्य पूर्वानुमान के लिए किया गया। यह आदर्श ए आर आई एम ए (0,1,1) की तुलना में अच्छा था क्योंकि सशर्त विचरण की अस्थिरता को पकड़ने की क्षमता अच्छी थी जो अस्थिर रूप से सम्पूर्ण समय तक था। वेक्टब आटो रिग्रेसिव (वी ए आर) मल्टीवेरियेट माडल का भी उपयोग पूर्वानुमान के लिए किया गया लेकिन आदर्श का निस्पादन जी ए आर

सी एच के ऊपर सुधरा हुआ नहीं पाया गया। निष्कर्ष के तौर पर जी ए आर सी एच (1,1) की अन्य की तुलना में प्याज के मूल्य के पूर्वानुमान के लिए उत्तम पाया गया क्योंकि परीक्षण सांख्यिकीय संगणक हेतु यह आदर्श नियम था। अन्य माडल तथा ए आई सी व एस आई सी जो जी ए आर सी एच माडल से प्राप्त हुए वे छोटे थे तथा पूर्वानुमान मूल्य से वास्तविक मूल्य के तुलनात्मक रूप से जी ए आर सी एच माडल में कम विविधता प्रतिशत थी। इस प्रकार यह स्पष्ट होता है कि प्रतिदिन मूल्य निर्धारण में ए.आर.आई.एम.ए. की तुलना में जी.ए.आर.सी.एच एक उत्तम माडल है।

Reference

Assis K, Amran A, Remali Y and Affendy H (2010) A comparison of univariate time series methods for forecasting cocoa bean prices. Trends Agric Eco 3: 207-215.

Moghaddasiand R and Badr BR (2008) An econometric model for wheat price forecasting in Iran. Paper presented in International Conference on Applied Economics–ICOAE.

Sen LK, Shitan M and Hussain H (2007) Time Series Modelling and Forecasting of Sarawak Black Pepper Price. MPRA Paper No. 791, posted 07 November, pp16.

Nochai R and Nochai T (2006) Arima model for forecasting oil palm price. Paper presented in IMT- GT regional conference on Mathematics. Statistics and Applications. Universiti Sains Malasyia, Penang June 13-15.

Bowman C and Aasim M. Husain (2004) Forecasting commodity prices: futures versus judgment. IMF Working Paper. WP/ 04/41.March.

Bollerslev T (1986) Generalised autoregressive conditional heteroskedasticity. J Econometrics 31: 307-327.

Engle R (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica 50: 987–1007.

Dickey DA and Fuller WA (1979) Distribution of the estimators for autoregressive time series with a unit root. J Am Statis Assoc 74: 427-431.

Box G and Jenkins G (1970) Time series analysis: forecasting and control. San Francisco: Holden-Day.