Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 68(1) 2014 85-92

Development of Out-of-Sample Forecasts Formulae for ARIMAX-GARCH Model and their Application*

Ranjit Kumar Paul, Himadri Ghosh and Prajneshu
Indian Agricultural Statistics Research Institute, New Delhi

Received 19 December 2012; Revised 30 January 2014; Accepted 31 January 2014

SUMMARY

Autoregressive integrated moving average with exogenous variable - Generalized autoregressive conditional heteroscedastic (ARIMAX-GARCH) methodology is employed here for describing volatile data by incorporating exogenous variables in the mean-model. Brief description of this model along with its estimation procedure is given. As an illustration, ARIMAX and ARIMAX-GARCH models are employed for modelling and forecasting of wheat yield of Kanpur district of Uttar Pradesh, India. Comparative study of the fitted models is carried out from the viewpoint of dynamic one-step ahead forecast error variance along with Mean square prediction error (MSPE), Mean absolute prediction error (MAPE) and Relative mean absolute prediction error (RMAPE). The formulae for more than one-step ahead out-of-sample forecasts along with forecast error variances for ARIMAX-GARCH model are derived analytically by recursive use of conditional expectation. Its superiority over ARIMAX approach is demonstrated for the data under consideration.

Keywords: ARIMAX, GARCH, Heteroscedasticity, Out-of-sample forecast, SAS and EViews software packages, Wheat yield.

1. INTRODUCTION

Quantitative understanding of crop responses to climate requires development of statistical models for various characteristics of a crop by taking into account its time-series behaviour along with exogenous climate factors. The response of interest (*e.g.* crop yield) as the response variable and various climate related quantities (*e.g.* growing season's rainfall, average monthly temperature) as the predictor variable(s) need to be considered for model building. It is well-known that one of the main factors causing yields to change from year to year is climate variability– no two growing seasons experience exactly the same weather (Lobell *et al.* 2006). When forecasting is carried out for dynamic behaviour of crop yield, it should be able to take advantage not only of historical data, but also of the impact of various driving forces, like temperature, precipitation, and relative humidity from the external environment. The key problem is how to incorporate pertinent external information into the forecasting process and subsequently into the decision making process. Lobell *et al.* (2006) developed weather-based yield forecast model for 12 Californian crops. The authors combined weather and yield data in a linear regression model to test how well yield anomalies could be predicted before harvest based on monthly weather measurements. Lobel and Burke (2010) used both simple regression model as well as panel regression models to predict crop yield responses to climate change. Bratina and Faganel (2008) employed ARIMAX model for forecasting primary demand for a beer brand in Slovenia.

*The paper was presented on 19 December 2012 in Dr. G.R. Seth Memorial Young Scientist Award Session during the 66th Annual Conference of the Indian Society of Agricultural Statistics held at IASRI, New Delhi during 18-20 December 2012.
Corresponding author: Ranjit Kumar Paul
E-mail address: ranjitstat@gmail.com

The second concern regarding forecasting of crop yield is the presence of heteroscedasticity. Most studies on crop yields either tend to ignore it or handle it improperly. Simple linear time trend as well as time-series models usually encounter variances that change over time. The Autoregressive conditional heteroscedastic (ARCH) model was proposed by R.F. Engle in 1982 and was applied to estimate the variance of inflation in U.K. It allows the conditional variance to change over time as a function of squared past errors, leaving the unconditional variance constant. As emphasized by Jaffee (2005), volatility seems to be the norm rather than exception due to variations in climatic conditions, and the rapidity with which producers can respond to price changes, etc. The combination of ARCH specification for conditional variance and Autoregressive (AR) specification for conditional mean has many appealing features, including a better specification of the forecast error variance. Ghosh and Prajneshu (2003) employed AR(p)-ARCH(q)-in-Mean model for carrying out modelling and forecasting of volatile monthly onion price data.

However, ARCH model has some drawbacks. Firstly, when the order of ARCH model is very large, estimation of a large number of parameters is required. Secondly, conditional variance of ARCH(q) model has the property that unconditional autocorrelation function (ACF) of squared residuals, if it exists, decays very rapidly compared to what is typically observed, unless maximum lag q is large. To overcome these difficulties, Bollerslev (1986) proposed the Generalized ARCH (GARCH) model in which conditional variance is also a linear function of its own lags. This model is also a weighted average of past squared residuals, but it has declining weights that never go completely to zero. It gives parsimonious models that are easy to estimate and, even in its simplest form, has proven surprisingly successful in predicting conditional variances. The most popular GARCH model in applications is GARCH (1,1) model. Paul *et al.* (2009) compared the efficiency of ARIMA-GARCH model vis-à-vis ARIMA model in respect of modelling and forecasting of India's volatile monthly spices export data. A good description of nonlinear time-series models and their applications is given by Prajneshu (2012).

In this paper, we have studied ARIMAX and ARIMAX-GARCH models along with procedures for estimation of their parameters. Formulae for more than one-step ahead out-of-sample forecasts along with forecast error variances for ARIMAX-GARCH model are derived analytically by recursive use of conditional expectation. As an illustration, these are applied for modelling and forecasting of wheat yield data of Kanpur District of Uttar Pradesh, India.

2. DESCRIPTION OF MODELS

2.1 ARIMAX Model

ARIMAX model (Bierens 1987) is a generalization of ARIMA model and is capable of incorporating an external input variable (X). Given a (k + 1) - time-series process $\{(y_t, x_t)\}$, where y_t and k components of x_t are real valued random variables, ARIMAX model assumes the form

$$\left(1-\sum_{s=1}^{p}\alpha_s L^s\right)y_t = \mu+\sum_{s=1}^{q}\beta_s' L^s x_t+\left(1+\sum_{s=1}^{r}\gamma_s L^s\right)e_t, \quad (1)$$

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, e_t's are 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.

2.2 Estimation of Parameters of ARIMAX Model

Nonlinear least squares method is employed to estimate the parameters of ARIMAX model. Following Bierens (1987), equation (1) can be written as follows:

$$y_t = \mu\Big/\left(1+\sum_{s=1}^{r}\gamma_s\right)+\left[\left(\sum_{s=1}^{p}\alpha_s L^s+\sum_{s=1}^{r}\gamma_s L^s\right)\Big/\left(1+\sum_{s=1}^{r}\gamma_s L^s\right)\right]y_t$$
$$+\left[\left(1+\sum_{s=1}^{q}\beta_s' L^s\right)\Big/\left(1+\sum_{s=1}^{r}\gamma_s L^s\right)\right]x_t+e_t,$$

Denoting $\theta_0 = (\mu, \alpha_1, ..., \alpha_p, \beta_1', ..., \beta_q', \gamma_1, ..., \gamma_r)'$,

$z_t = (y_t, x_t')'$, $\phi(\theta_0) = \mu\Big/\left(1+\sum_{s=1}^{r}\gamma_s\right)$ we get

$$\sum_{s=1}^{\infty}\eta_s(\boldsymbol{\theta_0})L^s = \left(1+\sum_{s=1}^{r}\gamma_s L^s\right)^{-1}\left[\frac{\sum_{s=1}^{p}\alpha_s L^s+\sum_{s=1}^{r}\gamma_s L^s}{\sum_{s=1}^{q}\beta_s L^s}\right].$$

Now equation (1) can be written as an ARX(∞) model:

$$y_t = \varphi(\boldsymbol{\theta})+\sum_{s=1}^{\infty}\eta_s(\boldsymbol{\theta})' z_{t-s} + e_t . \quad (2)$$

Assuming that only z_1, ..., z_n are observed, nonlinear least squares estimator of θ_0 can be obtained as follows:

Let

$$\tilde{g}_t(\boldsymbol{\theta}) = \varphi(\boldsymbol{\theta})+\sum_{s=1}^{t-1}\eta_s(\boldsymbol{\theta})' z_{t-s}, \quad if \quad t\geq 2$$

$$= \varphi(\boldsymbol{\theta}), \quad if \quad t\leq 1$$

and denote

$$\hat{Q}(\boldsymbol{\theta}) = (1/n)\sum_{t=1}^{n}\left(y_t - \tilde{g}_t(\boldsymbol{\theta})\right)^2 .$$

Then the least squares estimator $\hat{\boldsymbol{\theta}}$ of θ_0 is a solution of $\hat{Q}(\hat{\boldsymbol{\theta}}) = \inf_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \hat{Q}(\boldsymbol{\theta})$. Under e_0 : $E(e_t|(y_{t-1}, x_{t-1}), (y_{t-2}, x_{t-2}), ...) = 0$, $\sqrt{n}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\right)\rightarrow N_m\left(\boldsymbol{0}, \boldsymbol{\Omega}_1^{-1}\boldsymbol{\Omega}_2\boldsymbol{\Omega}_1^{-1}\right)$ in distribution, where $\boldsymbol{\Omega}_1$ is the probability limit of $\hat{\boldsymbol{\Omega}}_1 = (1/n)\sum_{t=1}^{n}\left((\partial/\partial\boldsymbol{\theta}')\,\tilde{g}_t(\hat{\boldsymbol{\theta}})\right)\left((\partial/\partial\boldsymbol{\theta})\,\tilde{g}_t(\hat{\boldsymbol{\theta}})\right)$, and $\boldsymbol{\Omega}_2$ is the probability limit of

$$\hat{\boldsymbol{\Omega}}_2 = (1/n)\sum_{t=1}^{n}(y_t - \tilde{g}_t((\hat{\boldsymbol{\theta}}))^2((\partial/\partial\boldsymbol{\theta}'))\tilde{g}_t(\hat{\boldsymbol{\theta}}))((\partial/\partial\boldsymbol{\theta})\tilde{g}_t(\hat{\boldsymbol{\theta}})).$$

After estimating the parameters of ARIMAX model, residuals are used for testing for ARCH effects by employing ARCH-Lagrange Multiplier (ARCH-LM) test due to Engle (1982), as described below.

2.3 Testing for ARCH Effects

Let ε_t be the residual series. The squared series $\{\varepsilon_t^2\}$ is then used to check for conditional heteroscedasticity, which is also known as the ARCH effects. To this end, two tests, briefly discussed below, are available. The first one is to apply the usual Ljung-Box statistic $Q(m)$ to the $\{\varepsilon_t^2\}$ series. The null hypothesis is that the first m lags of autocorrelation functions of the $\{\varepsilon_t^2\}$ series are zero. The second test for conditional heteroscedasticity is the LM test, which is equivalent to usual F-statistic for testing $e_0 : \alpha_i = 0$, $i = 1, 2, ..., q$ in the linear regression

$$\varepsilon_t^2 = a_0 + a_1\varepsilon_{t-1}^2 + ... + a_q\varepsilon_{t-q}^2 + e_t, \; t = q + 1, ..., \; T \quad (3)$$

where e_t denotes error term, q is prespecified positive integer, and T is sample size. Let $SSR_0 = \sum_{t=q+1}^{T}(\varepsilon_t^2-\varpi)^2$, where $\varpi = \sum_{t=q+1}^{T}\varepsilon_t^2/T$ is sample mean of (ε_t^2), and $SSR_1 = \sum_{t=q+1}^{T}\hat{e}_t^2$, where $\hat{e}_t$ is least squares residual of (3). Then, under e_0:

$$F = \frac{(SSR_0 - SSR_1)/q}{SSR_1(T-q-1)}$$

is asymptotically distributed as chi-squared distribution with q degrees of freedom. The decision rule is to reject e_0 if $F > \chi_q^2(\alpha)$, where $\chi_q^2(\alpha)$ is the upper $100(1-\alpha)^{th}$ percentile of χ_q^2 or, alternatively, the p-value of F is less than α.

2.4 GARCH Model

The ARCH(q) model (Engle 1982) for the series $\{\varepsilon_t\}$ is defined by specifying the conditional distribution of ε_t given information available up to time $t-1$. Let ψ_{t-1} denote this information. Then, ARCH (q) model for the series $\{\varepsilon_t\}$ is given by

$$\varepsilon_t \mid \psi_{t-1} \sim N(0, h_t), \quad (4)$$

$$h_t = a_0 + \sum_{i=1}^{q} a_i\varepsilon_{t-i}^2, \quad (5)$$

where $a_0 > 0$, $a_i \geq 0$ for all i and $\sum_{i=1}^{q} a_i < 1$ are required to be satisfied to ensure nonnegativity and finite unconditional variance of stationary $\{\varepsilon_t\}$ series. Bollerslev (1986) proposed the Generalized ARCH (GARCH) model has the following form

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

$$h_t = a_0 + \sum_{i=1}^{q} a_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} E_j h_{t-j}, \qquad (6)$$

where $\xi_t \sim$ IID (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

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

As (6) is a more parsimonious model of the conditional variance than a high-order ARCH model, most users prefer it to the simpler ARCH alternative.

2.5 Estimation of Parameters for GARCH Model

Similar to the estimation of parameters of ARMA models, most frequently used estimators for ARCH/GARCH models are those derived from a (conditional) Gaussian likelihood function. The loglikelihood function of a sample of T observations, apart from constant, is

$$L_T(\theta) = T^{-1} \sum_{t=1}^{T} \left(\log h_t + \varepsilon_t^2 h_t^{-1} \right),$$

where

$$h_t = a_0 + \sum_{i=1}^{q} a_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} E_j h_{t-j} .$$

For a general GARCH model, conditional variance (h_t) cannot be expressed in terms of a finite number of past observations. Some truncation is inevitable. By induction, it is possible to derive

$$h_t = a_0/(1 - \sum_{i=1}^{q} a_i) + \sum_{i=1}^{q} a_i \varepsilon_{t-i}^2$$

$$+ \sum_{i=1}^{q} a_i \sum_{k=1}^{\infty} \sum_{j_1=1}^{p} \cdots \sum_{j_k=1}^{p} b_{j_1} \ldots b_{j_k} \varepsilon_{t-i-j_1-\ldots-j_k}^2 .$$

where the multiple sum vanishes if $q = 0$. It may be noted that the multiple sum above converges with probability 1 since each a_i and b_i is nonnegative, and expected value of the multiple series is finite. In practice, above expression of h_i, $t > q$, is replaced by the truncation version

$$\tilde{h}_t = a_0/(1 - \sum_{i=1}^{q} a_i) + \sum_{i=1}^{q} a_i \varepsilon_{t-i}^2$$

$$+ \sum_{i=1}^{q} a_i \sum_{k=1}^{\infty} \sum_{j_1=1}^{p} \cdots \sum_{j_k=1}^{p} b_{j_1} \ldots b_{j_k} \varepsilon_{t-i-j_1-\ldots-j_k}^2 ,$$

$$\left(t - i - j_1 - \ldots - j_k \geq 1 \right).$$

In general, suppose that f (.) is the probability density function of ε_t. However, generally, maximum likelihood estimators are derived by minimizing

$$L_T(\theta) = T^{-1} \sum_{t=v}^{T} \left(\log \sqrt{\tilde{h}_t} - \log f\left(\varepsilon_t / \sqrt{\tilde{h}_t} \right) \right),$$

where $\tilde{h}_t$ is the truncated version of h_t (Fan and Yao 2003). For heavy tailed error distribution, Peng and Yao (2003) proposed Least absolute deviations estimation (LADE) which minimizes $\sum_{t=v}^{T} \left| \log \varepsilon_t^2 - \log(h_t) \right|$, where $v = p + 1$, if $q = 0$ and $v > p + 1$, if $q > 0$. Fan and Yao (2003) and Straumann (2005) have given a good description of various estimation procedures for conditionally heteroscedastic time-series models.

The Akaike information criterion (AIC) and Bayesian information criterion (BIC) values for GARCH model with Gaussian distributed errors are computed by

$$AIC = \sum_{t=1}^{T} \left(\log \tilde{h}_t + \varepsilon_t^2 \tilde{h}_t^{-1} \right) + 2(p + q + 1)$$

and

$$BIC = \sum_{t=1}^{T} \left(\log \tilde{h}_t + \varepsilon_t^2 \tilde{h}_t^{-1} \right) + 2(p + q + 1) \log(T - v + 1) \qquad (7)$$

where T is the total number of observations. Evidently, likelihood equations are extremely complicated. Fortunately, parameter estimates can be obtained by using a software package, like EViews, SAS, and MATLAB.

3. AN ILLUSTRATION

Annual wheat yield data of Kanpur district of Uttar Pradesh, India during 1972 to 2011 comprising 40 data points are obtained from Directorate of Economics and Statistics, Government of India. The first 36 observations, *i.e.* the data from 1972 to 2007 are used for model building and remaining 4 data points, *i.e.* the data from 2008 to 2011 are used for validating the model. Daily climate data on maximum temperature for the same time period, obtained from India Meteorological Department, Government of India, is first converted to weekly data. It may be noticed that wheat yield data vary from a minimum of 9.78 quintals per hectare in 1974 to a maximum of 34.80 quintals per hectare in 2011. It may be noted that the yield was 29.12 quintals per hectare in 1990 and in the very next year, it suddenly got down to 21.25 quintals per hectare and again in the next year, it jumped to the value of 30.62 quintals per hectare. All this implies possible presence of conditional heteroscedasticity in the yield data. On exploratory data analysis, it is seen that wheat yield is significantly negatively correlated with maximum temperature at the Critical Root Initiation (CRI) stage of wheat crop every year with correlation coefficient –0.49. In India, the CRI stage comes roughly 21 days after sowing of wheat crop, *i.e.* in the third week of November every year. SAS software Ver. 9.3 and EViews software Ver. 7 are used for data analysis.

3.1 Fitting of ARIMAX Model

Evidently, data set of wheat yield is not stationary and so it requires to be differenced. In order to select order of the ARIMA model, unit root test proposed by Dickey and Fuller (1979) is applied for parameter ρ in the auxiliary regression

$$\Delta_1 y_t = \rho y_{t-1} + \alpha_1 \Delta_1 y_{t-1} + \varepsilon_t. \tag{8}$$

The relevant null hypothesis is $e_1 = \rho = 0$ and the alternative hypothesis is $e_1 : \rho < 0$.

In the present situation, estimate of ρ is –0.64 with calculated *t*-statistic as –4.62, which is less than the critical value of *t* at 5% level of significance, *i.e.* –1.95 (Franses 1998, Page 82). Therefore, e_0 is not rejected at 5% level and so $\rho = 0$. Thus, there is presence of one unit root and so differencing is required. On the other hand, time-series data of maximum temperature at third week of November is found to be stationary. On the basis of minimum AIC and BIC values, best ARIMAX(1, 1, 1) model is given by

$$\Delta_1 y_t = 17.36 - 0.61\, x_t - 0.23\, \Delta_1 y_{t-1} - 0.78\varepsilon_{t-1} + \varepsilon_t.$$
$$\quad (2.84)\ \ (0.10)\qquad (0.17)\qquad\qquad (0.09)$$

where the values in parentheses denote standard errors of the corresponding parameter estimates. For visual display, fitted model along with data points is exhibited in Fig. 1.

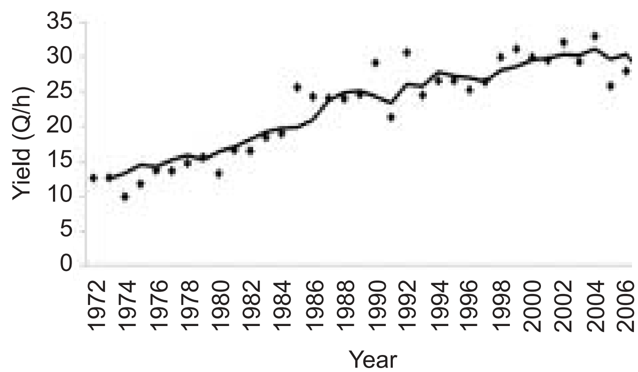

Fig. 1. Fitted ARIMAX model along with data points

3.2 Fitting of ARIMAX-GARCH Model

Residuals of fitted ARIMAX model are used for testing the presence of ARCH effect. The autocorrelation function (ACF) and partial autocorrelation function (PACF) of squared standardized residuals of fitted ARIMAX model are reported in Table 1. A perusal indicates that squared standardized residuals at lags 3 and 11 are significant. Accordingly, ARCH-LM test is applied to test the presence of ARCH effect and it is found that the test is significant at 5 % level at both the lags. Because of

Table 1. ACF and PACF of Squared standardized residuals of fitted ARIMAX model

Lag	ACF	PACF	Lag	ACF	PACF
1	0.066	0.066	9	–0.109	–0.147
2	0.186	0.182	10	0.030	0.087
3	**–0.379**	**–0.209**	**11**	**–0.261**	**–0.267**
4	–0.101	–0.115	12	–0.042	–0.100
5	–0.145	–0.061	13	0.033	0.142
6	–0.005	0.018	14	0.042	–0.075
7	0.025	0.029	15	–0.063	–0.199
8	–0.046	–0.111	16	–0.001	–0.027

parsimony property of GARCH model, instead of ARCH model, ARIMAX-GARCH model is fitted. The estimates of parameters of ARIMAX-GARCH model are presented in Table 2 and fitted model along with data points is exhibited in Fig. 2.

Table 2. Parameters estimates for fitted ARIMAX-GARCH model.

Variable	Coefficient	Standard Error	*t*-Statistic	Probability
Mean Equation				
C	18.79	8.40	2.235	0.025
MAXT_NOV3	–0.67	0.33	–2.207	0.027
AR(1)	–0.66	0.14	–4.676	<0.001
Variance Equation				
C	12.00	5.201	2.307	0.021
ARCH(1)	0.31	0.142	2.120	0.028
GARCH(1)	0.58	0.249	2.320	0.009

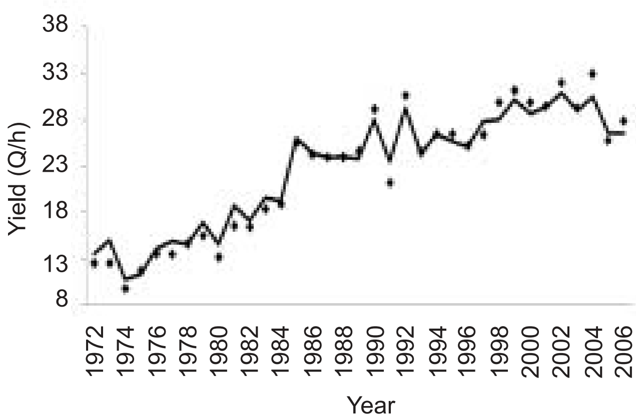

Fig. 2. Fitted ARIMAX-GARCH model along with data points.

4. OUT-OF-SAMPLE FORECASTS

Suppose ε_t is the residual of the fitted ARIMAX-GARCH model, Y_t is differenced series and X_t is the exogenous variable. Consider ARIMAX (1,0,1) model:

$$Y_t = \rho_0 + \rho_1 Y_{t-1} + \rho_2 X_t + \varepsilon_t, \tag{8}$$

where

$$\varepsilon_t = h_t^{1/2}\eta_t,\; h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}. \tag{9}$$

Let F_T denote the information up to time T. Then, optimal one-step ahead forecast of Y_{T+1} given F_T is

$$\hat{Y}_{T+1|F_T} = E[Y_{T+1}|F_T] = \rho_0 + \rho_1 Y_T + \rho_2 X_{T+1} \tag{10}$$

where estimated one-step ahead forecast error variance is given by

$$h_{T+1} = \alpha_0 + \alpha_1 \varepsilon_T^2 + \beta_1 h_T. \tag{11}$$

From equation (10), one-step ahead predictor $\hat{Y}_{T+2|F_{T+1}}$ of Y_{T+2} given F_{T+1} can be estimated by

$$\hat{\hat{Y}}_{T+2|F_{T+1}} = \rho_0 + \rho_1 \hat{Y}_{T+1|F_T} + \rho_2 X_{T+2}. \tag{12}$$

Now, one-step ahead prediction error variance of Y_{T+2} given F_{T+1} can be estimated by taking conditional expectation of ε_{T+1}^2 given F_T. A straightforward algebra, using equation (11), yields

$$\hat{h}_{T+2} = \alpha_0 + (\alpha_1 + \beta_1) h_{T+1}. \tag{13}$$

Proceeding along similar lines as above, optimal one-step ahead forecast of Y_{T+3} given F_{T+2} is obtained as

$$\hat{Y}_{T+3|F_{T+2}} = E[Y_{T+3}|\, F_{T+2}] = \rho_0 + \rho_1 Y_{T+2} + \rho_2 X_{T+3}, \tag{14}$$

which can be estimated by

$$\hat{\hat{Y}}_{T+3|F_{T+2}} = \rho_0 + \rho_1 \hat{Y}_{T+2|F_T} + \rho_2 X_{T+3}, \tag{15}$$

where $\hat{Y}_{T+2|F_T}$ is the conditional expectation of Y_{T+2} given F_T . It can be obtained by using recursive conditional expectation approach as

$$\hat{Y}_{T+2|F_T} = \rho_0 + \rho_1 \hat{Y}_{T+1|F_T} + \rho_2 X_{T+2}.$$

Proceeding along similar lines as above, expression for $\hat{h}_{T+3}$ is derived as

$$\hat{h}_{T+3} = \alpha_0(1 + \alpha_1 + \beta_1) + h_{r+1}(\alpha_1 + \beta_1)^2. \tag{16}$$

In general, the optimal one-step ahead forecast is given by

$$\hat{Y}_{T+i|F_{T+i-1}} = \rho_0 + \rho_1 Y_{T+i-1} + \rho_2 X_{T+i}, \tag{17}$$

which can be estimated by

$$\hat{\hat{Y}}_{T+i|F_{T+i-1}} = \rho_0 + \rho_1 \hat{Y}_{T+i-1|F_T} + \rho_2 X_{T+i}, \tag{18}$$

$$\hat{Y}_{T+i|F_T} = \rho_0 + \rho_1 \hat{Y}_{T+i-1|F_T} + \rho_2 X_{T+i}. \tag{19}$$

Further, the estimated one-step ahead forecast error variance is given by

$$\hat{h}_{T+i} = \alpha_0(1+(i-2)(\alpha_1 + \beta_1)) + h_{T+1}(\alpha_1 + \beta_1)^{i-1} \quad (20)$$

Using above equations, one-step ahead forecasts of wheat yield along with their corresponding standard errors within brackets ( ) for the hold-out data, *i.e.* the data during 2008-'11 in respect of above fitted models are reported in Table 3. A perusal indicates that, for fitted ARIMAX-GARCH model, all the forecast values lie within one standard error of forecasts. However, this attractive feature does not hold for fitted ARIMAX model.

Table 3. One-step ahead forecasts of wheat yield (Quintal/Hectare) for hold-out data

		Forecasts by models	
Years	Actual yield	ARIMAX	ARIMAX-GARCH(1,1)
2008	26.66	26.72 (2.19)	26.63 (2.23)
2009	29.40	27.76 (2.19)	28.38 (3.91)
2010	31.50	29.43 (2.19)	29.37 (4.12)
2011	33.80	30.80 (2.19)	32.35 (4.96)

The Mean square prediction error (MSPE), Mean absolute prediction error (MAPE) and Relative mean absolute prediction error (RMAPE) values for fitted ARIMAX-GARCH model are respectively computed as 1.76, 0.90, and 2.76%, which are found to be much lower than the corresponding ones for fitted ARIMAX model, viz. 4.67, 1.44, and 4.38% respectively. Lower values of all the three statistics reflect the superiority of ARIMAX-GARCH approach for forecasting purposes also. Finally, using equations (19) and (20), out-of sample forecasts of wheat yield along with corresponding forecast error variances for the years 2012-2015 are computed and the same are reported in Table 4.

Table 4. One-step ahead out-of-sample forecasts of wheat yield (Quintal/Hectare) for fitted ARIMAX-GARCH model

Years	Forecast	Standard error
2012	34.03	5.64
2013	34.24	6.02
2014	34.60	6.75
2015	35.21	7.23

5. CONCLUDING REMARKS

In this article, heteroscedastic time-series data of wheat yield of Kanpur district of Uttar Pradesh, India is modelled by considering most important factor, *i.e.* Maximum temperature at critical root initiation stage of wheat crop, which occurs at 3rd week of November every year as explanatory variable in ARIMAX-GARCH model. Formulae for i^{th}-step ahead out-of-sample forecasts along with forecast error variances are derived analytically and are applied to real data. It is found that ARIMAX-GARCH model outperforms ARIMAX model so far as modelling and forecasting of wheat yield is concerned. This type of study would go a long way in helping the planners to make optimum decisions about export and import policies. Finally, the methodology advocated here is very general and can be used for modelling and forecasting of any crop yield exhibiting volatility by appropriately identifying the exogenous variables.

ACKNOWLEDGEMENTS

Authors are grateful to the referee for valuable suggestions.

REFERENCES

Bierens, H.J. (1987). ARMAX model specification testing, with an application to unemployment in the Netherlands. *J. Econ.*, **35**, 161-190.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *J. Econ.,* **31**, 307-327.

Bratina, D. and Faganel, A. (2008). Forecasting the primary demand for a beer brand using time-series analysis. *Organizacija*, **41**, 116-123.

Dickey, D.A. and Fuller, W.A. (1979). Distribution of the estimators for the autoregressive time series with a unit root. *J. Amer. Statist. Assoc*., **74**, 427-431.

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation. *Econometrica,* **50**, 987-1008.

Fan, J. and Yao, Q. (2003). *Nonlinear Time Series: Nonparametric and Parametric Methods.* Springer, U.S.A.

Franses, P.H. (1998). *Time Series Models for Business and Economic Forecasting*. Cambridge University Press, U.K.

Ghosh, H. and Prajneshu (2003). Nonlinear time series modelling of volatile onion price data using AR(p)-ARCH(q)-in-Mean. *Cal. Stat. Assoc. Bull.*, **54,** 231-247.

Jaffee, S. (2005). Delivering and Taking the Heat: Indian Spices and Evolving Product and Process Standards. *Agriculture and Rural Development Discussion Paper 19.* The World Bank, U.S.A.

Lobell, D.B., Field, C.B., Cahill, K.N. and Bonfils, C. (2006). Impacts of future climate change on California perennial crop yields: Model projections with climate and crop uncertainties. *Agri. For. Met.*, **141**, 208-218.

Lobell, D.B. and Burke, M.B. (2010). On the use of statistical models to predict crop yield responses to climate change. *Agri. For. Met.,* **150**, 1443-1452.

Paul, R.K., Prajneshu and Ghosh, H. (2009). GARCH nonlinear time-series analysis for modelling and forecasting of India's volatile spices export data. *J. Ind. Soc. Agril. Statist.*, 6**3**, 123-131.

Peng, L. and Yao, Q. (2003). Least absolute deviations estimation for ARCH and GARCH models. *Biometrika,* **90**, 967-975.

Prajneshu (2012). Some nonlinear time series models and their applications. *J. Ind. Soc. Agril. Statist.*, **66**, 243-250.

Straumann, D. (2005). *Estimation in Conditionally Heteroscedastic Time Series Models*. Springer, Germany.