



<http://www.bomsr.com>
Email:editorbomsr@gmail.com

RESEARCH ARTICLE

A Peer Reviewed International Research Journal



AN IMPROVED MODEL FOR FORECASTING OF VALUE OF GOLD: A HYBRID MODEL APPROACH

JYOTHI U¹, SURESH KK²

¹Research and Development Centre, Bharathiar University, Coimbatore, Tamil Nadu

²Professor & Head, Department of Statistics, Bharathiar University, Coimbatore Tamil Nadu

¹Department of Statistics, KristuJayanti College, Bangalore, Karnataka



ABSTRACT

The price of gold in India has undergone unprecedented fluctuations in the recent past. As India is the second largest imported of gold and its price is a matter of concern. Moreover, though the demand of gold in India as jewelry and for industrial purpose is prevalent, the major demand is as an investment since it has always proved to be a safe haven in times of financial turbulences. The study attempts to find an econometric model which will be able to forecast a financial time- series accurately for long term trend by capturing the amplitudes of volatility clustering. This paper aims at building the best fit ARIMA model for the given data and hence to capture the volatility of the residuals by means of the best fit GARCH model. The prediction power and validity of a hybrid model of ARIMA and GARCH is analysed in the study.

Key words: Time Series, Forecasting, ARIMA, GARCH, Volatility clustering, Hybrid model

©KY PUBLICATIONS

I. INTRODUCTION

Gold is considered to be a hedge as it assures maximum returns at minimum risk. Due to these attributes the demand of the gold is increasing day by day in India. India is one of the largest consumers of gold. Nearly 800 tons of gold is imported every year. Indian accounts for 23 percent of the world's total annual demand for gold [1]. The efficiency of these models in capturing the fluctuations in gold is the interest of the study. Price of gold has been going through unpredictable ups and downs. The rise of price from 2008 took it to almost 200 percent increase in 2013 (all time high of Rs. 32943 per 10 gram on 29th August 2013) and an almost equal fall of around 20 percent in 2015 (Rs 26645 per 10 grams on September 2015).

Among the various econometric models, the most widely applicable ones are ARIMA as a linear model which allows one to focus on the dynamics of the series and GARCH to capture volatility. ARIMA is extensively used in the area of short term forecasting of stationary time series. The highlight of this model is the assumption of uniform variance of the residuals. When this is violated, the model seems to be weak. Hence an attempt is made to construct an autoregressive model of the series and capture the variances of the residuals obtained using the most widely used Heteroskedastic model Generalised Heteroskedastic Model (GARCH).

II. Methodology

Autoregressive Integrated Moving Average Model ARIMA(p, d, q) developed by Box-Jenkins[2] is a combination of Autoregressive model, AR(p) and Moving Average model, MA(q) of a time series which is stationary at level d. AR model describes the model as a linear model regressed on its past values at p time points while the MA model represents the residuals as a model of its past values at q time points.

The ARIM (p,d,q) model may be represented as

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \epsilon_t + \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + \dots + \beta_q \epsilon_{t-q} + V_t \quad (1)$$

To develop the ARIMA model, we analyse the components of the series, i.e, trend and seasonality. This is to detrend and deseasonalise the data if there is any pattern observed. This is followed by stationarity test which can be obtained by any standard tests like Dickey –Fuller test or Phillips – Perron test. If the series is found to be non-stationary, we try to achieve stationarity by either differencing or log transformations. The value of d in ARIMS is the level at which a given series becomes stationary and is determined through Unit- root test. Once we achieve a stationary time series, the parameters of the AR and MA (p, q) are determined using Correlogram which will depict Autocorrelation and Partial Autocorrelation plots. The value of p is determined from the PCF plot while that of q is determined from the ACF plot. The significance level of individual coefficients is measured by Box- Pierce Q statistics and for all the coefficients jointly together by Ljung- Box (LB) statistic. The Box- Pierce Q statistics is given by

$$Q = \sum_{k=1}^m \rho_k^2 = \chi^2 \quad (2)$$

And LB statistic is defined as

$$LB = n(n+2) \sum_{k=1}^m \frac{\rho_k^2}{n-k} \quad (3)$$

where n is the sample size and m is the lag length.

Diagnostic test can also be carried out with Akaike's Information Criterion (AIC) and Schwartz Information Criterion (SIC). The least values of AIC and SIC will be yield by the best fit models. Mean absolute Percent Error (MAPE) is also widely utilized as model diagnostic.

The basic assumption that the variance of residuals in this model is heteroskedastic is violated most of the times. Hence in this paper we extract the residuals to fit a GARCH model which will explain the long term patterns of heteroskedasticity.

Among the multitude of non- linear models, the most popular models are Autoregressive Conditionally Heteroskedastic (ARCH) model and Generalized Autoregressive Conditionally Heteroskedastic (GARCH) models[6]. They are expedient in modeling and forecasting volatility. The assumption of CLRM is that the variance of the errors is constant, which is termed homoscedasticity. ARCH models do not assume that variance is constant. ARCH models are preferred to CLRMs since it is unlikely to financial time series to have constant variance for the errors over time. Another important feature that motivates the use of ARCH/ GARCH models is volatility clustering or volatility pooling. Volatility clustering describes the tendency of large changes in assets prices to follow large

changes and small changes to follow small changes. Knowledge of volatility is important to make effective financial planning.

If the first difference has varying variance instead of a constant, it is termed Autoregressive Heteroskedastic (ARCH) models. ARCH models are capable of modeling and capturing volatility clustering. An ARCH(q) model is usually represented as

$$Y_t = f(Y_{t-p}) + u_t; \quad (4)$$

$u_t \sim N(0, h_t)$ where

$$h_t = \alpha_0 + \alpha_1 u_{t-1} + \alpha_2 u_{t-2} + \alpha_3 u_{t-3} + \dots + \alpha_q u_{t-q} \quad (5)$$

Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model is one of the most popular ARCH model. It allows the conditional variance to be dependent upon previous own lags, so that the conditional variance equation can be represented as

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (6)$$

Eq (6) represents the GARCH (1, 1) model.

GARCH (p, q) model may be represented as follows:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1} + \alpha_2 u_{t-2} + \alpha_3 u_{t-3} + \dots + \alpha_q u_{t-q} + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 + \dots + \beta_p \sigma_{t-p}^2 \quad (7)$$

Before estimating a GARCH type model, first it must be ensured that this model is appropriate. The most widely used test for non-linearity is BDS test. This test is also a model diagnostic. It has a null hypothesis that the data are pure noises. If a proposed linear model is adequate, then the standardized residuals should be white noises, while if the postulates model is insufficient to capture all the relevant features of the data, BDS statistic for the standardized residuals will be statistically significant. Once ARCH effect is found then we are required to find the specification of the model using Akaike Information Criterion (AIC) and Schwartz Bayesian Information Criterion (SBIC).

III. Review of Literature

Research papers in Econometric modeling of financial time – series has no dearth. But hybrid models which combine two or more models to improve the accuracy of estimation and prediction are not very popular as the former.

There are a few researches that has taken place with a similar concept of hybrid of two or more models. Sallemet.al [7] in an attempt to model the time series of rainfall, findsthere is volatility clustering in the series. The focus was on modeling and predicting the mean behavior of the time series through conventional methods of an Autoregressive Moving Average (ARMA) modeling proposed by the Box Jenkins methodology. The conventional models operate under the assumption that the series is stationary that is: constant mean and either constant variance or season-dependent variances, however, does not take into account the second order moment or conditional variance, but they form a good starting point for time series analysis. The residuals from preliminary ARIMA models derived from the daily rainfall time series were tested for ARCH behavior. The autocorrelation structure of the residuals and the squared residuals were inspected, the residuals are uncorrelated but the squared residuals show autocorrelation, the Ljung-Box test confirmed the results. McLeod-Li test and a test based on the Lagrange multiplier (LM) principle were applied to the squared residuals from ARIMA models. The results of these auxiliary tests show clear evidence to reject the null hypothesis of no ARCH effect. Hence indicates that GARCH modeling is necessary. Therefore the composite ARIMA-GARCH model captures the dynamics of the daily rainfall series in study areas more precisely. On the other hand, Seasonal ARIMA model became a suitable model for the monthly average rainfall series of the same locations treated.

Babu et.al [1] has used the hybrid approach to model Stock market data for multi-step or N-step ahead forecasting. For one-step ahead prediction ARIMA or GARCH can be employed, but for multi-step ahead prediction, the prediction accuracy decreases and data dynamics is lost over the complete prediction horizon. In the study they have decomposed the series into two using Moving Average (MA) filter, one of them is modelled using ARIMA and the other using GARCH aptly. Indian Stock market data is considered in order to evaluate the accuracy of the proposed model. The performance of this model is compared with traditional models, which reveals that for multi-step ahead prediction, the proposed model outperforms the others in terms of both prediction accuracy and preserving data trend.

Maizah[4] et.al.proposed hybrid model of linear ARIMA and non-linear GARCH to be better in forecasting the price of Malaysian gold. They have used the daily selling prices of the 1 oz Malaysian gold recorded from 18th July 2001 until 15th April 2014. ARIMA (1, 1, 1) model is found to be the most appropriate by comparing the AIC and MAPE values of prediction. The residuals of this model on analysis by ARCH LM test for hetroskedasticity proved to have ARCH effect. Hence a GARCH model is proposed to handle the hetroskedasticity of the series. ARIMA(1, 1, 1) along with GARCH (2,1) proved to be the most appropriate combination. Based on the AIC values, the model that minimizes the estimated information loss more is ARIMA-GARCH. ARIMA is 0.46 times as probable as ARIMA-GARCH to minimize the information loss. The bias proportion, the variance proportion, and the covariance proportion sum up to 1. While the bias proportion measures how far the mean of the forecast is from the mean of the actual series, the variance proportion measures how far the variation of the forecast is from the variation of the actual series. The remaining unsystematic forecasting errors are measured by the covariance proportion measures. They conclude that the most effective way to improve forecasting accuracy is using hybrid model.

Maizah [5] et.al used a four year daily data to find an accurate model for prediction and concludes that ARIMA- GJR is a better model which is providing lower values of MAPE than ARIMA only. Using ordinary least squares method to estimate the parameters, an appropriate ARIMA model for this series is ARIMA (2, 1, 2) with an AIC value of 10.88681. When the model was used for forecasting, the MAPE value for in-sample forecast is 0.759026. the residuals were not following Normal distribution, according to JarqueBera test result and also existence of c =volatility clustering was evident among the residuals as per LM test result. Comparing the various variants of GARCH models, it is concluded that ARIMA(2, 1,,2)- GJR(1,1) gives the best results with least MAPE values even though non-normality is prevalent in the residuals.

IV. Result and Discussion

The value of gold traded in MCX in lakhs of rupees for the past 5years (April 1st, 2010 to November 30th, 2015) is the data used in the present study. A total of 1204 sample data from April 1st, 2010 to March 31st 2014 is used as in-sample for modeling and the remaining data of 427sample points from April 1st. 201 to November 30th, 2015 are used for forecasting as out-sample. The analysis is carried out using the econometric software Eviews.

The general descriptive statistics as provided in Fig1 indicates average value of five years as Rs. 768436.20 with a very high standard deviation of Rs. 564844.70. The series is positively skewed ($1.674 > 0$) indicating majority of the days price was below the average price and the series is leptokurtic (Kurtosis= $8.22 > 3$) indicating a clustering of price around the average. The series seem to vary significantly from Normal distribution as the null hypothesis of Normality in JarqueBera test for is rejected. (Jarque – Bera value = 2620.364 with p- value < 0.05).

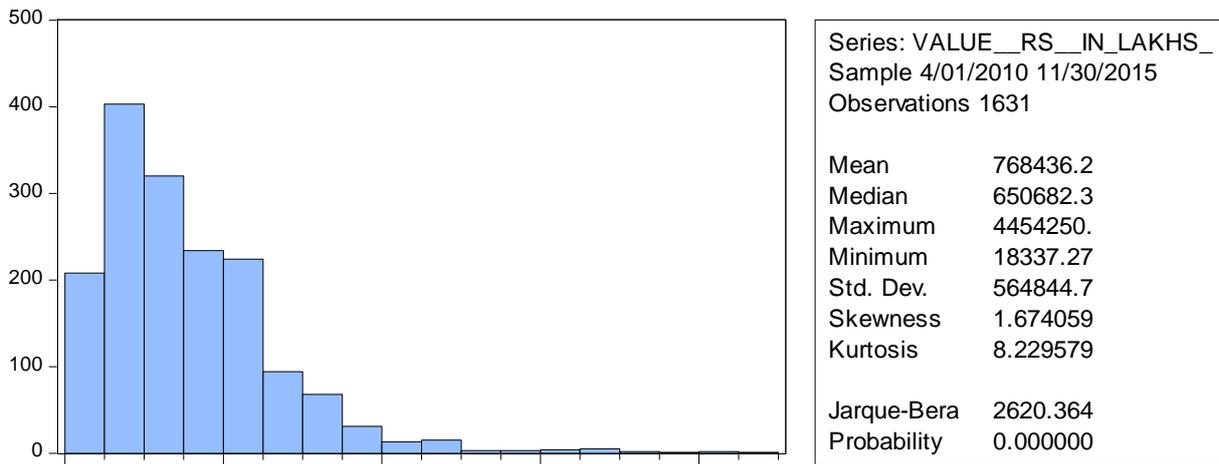


Figure:1: Daily Value of gold traded in MCX in Lakhs of rupees.

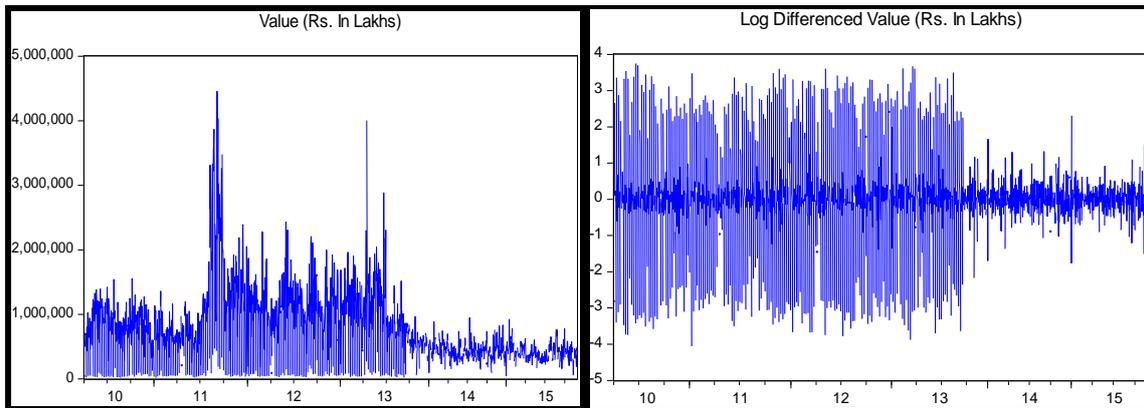


Figure 2. Plot of Daily Value of gold traded in MCX in Lakhs of rupees and Plot of Log differenced Value.

The line graph (Fig 2) indicates a typical stationary time-series. This is confirmed by Augmented Dickey-Fuller test ($t = -3.0959$ with $p\text{-value} < 0.05$), indicating the absence of Unit root.

Table 1: Augmented Dickey- Fuller test

Lag Length: 11 (Automatic - based on SIC, maxlag=24)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.095984	0.0271
Test critical values:		
1% level	-3.434184	
5% level	-2.863120	
10% level	-2.567659	

*MacKinnon (1996) one-sided p-values.

The parameters of ARIMA model is determined on the basis of the Correlogram and confirmed by AIC and MAPE values

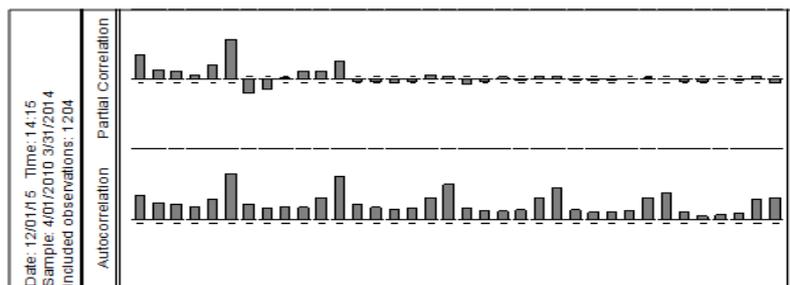


Figure 3: Correlogram of the Log differenced Value of gold traded in MCX

The Correlogram hints AR(1) and MA(1). However the different parameters of p and q in ARIMA(p,d,q) is checked with d = 0, since the series is stationary at level. The criterion to select the best model is least values of AIC and MAPE value for the in- sample forecast.

Table 2: Comparing the AIC and MAPE values to choose (p, q) values

AR(p)	MA(q)	AIC	MAPE
2	0	9.4732	34.1045
1	1	9.1825	31.1192
2	1	9.1815	31.7624
2	2	9.1810	30.1062
1	0	9.6104	31.8164
1	2	9.1820	31.4600
1	3	9.1752	31.8154

The most appropriate model seems to be ARIMA(2, 0, 2) with least values of AIC as 9.181 and MAPE as 30.1062. The fitted ARIMA model is represented as

$$Y_t = 0.7191 * Y_{(t-1)} + 0.2789 * Y_{(t-2)} - 0.5514 \epsilon_{(t-1)} - 0.3208485 \epsilon_{(t-2)},$$

where Y_t is the value of gold and ϵ is the residuals of the model.

IV.1. Residual Diagnosis of the model ARIMA (2,0,2)

The residuals of the fitted model is supposed to be $N(0, \sigma^2)$. Figure 4 provides the descriptive analysis of the residuals and result of normality test. The residuals are not $N(0, \sigma^2)$ and hence it is worth to model the residuals for its non- linearity.

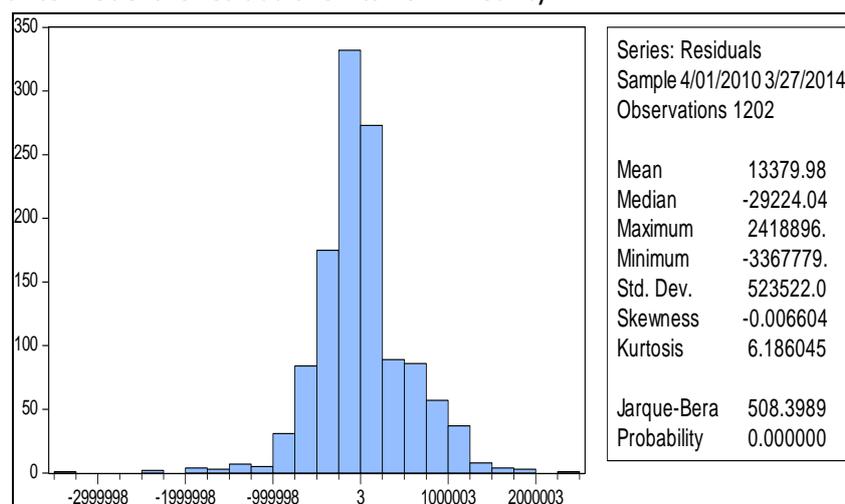


Figure 4: Descriptive statistics of residuals of ARIMA(2, 0,2)

IV.2. Test for Serial correlation

The residuals are tested for their serial correlation by Breusch – Godfrey LM test. The null hypothesis of no serial correlation is rejected (F statistic value=5.03, p- value =0.0067 <0.05). Hence the residuals seem to have autocorrelation.

Table 3: Result of test of serial correlation carried out on residuals of ARIMA(2, 0 2)

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	5.030966	Prob. F(2,1196)	0.0067
Obs*R-squared	9.248808	Prob. Chi-Square(2)	0.0098

IV.3. Test for Heteroskedasticity

Test of heteroskedasticity by Breusch-Pagan-Godfrey with null hypothesis that ARCH effect does not exist is carried out on the residuals. The test result is provided in table 4; it suggests the presence of ARCH effect (F- statistic= 111.01, p- value < 0.05 . This is ascertained from the line graph of the residuals

Table 4: Result of heteroskedasticity test carried out
Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	111.0110	Prob. F(2,1199)	0.0000
Obs*R-squared	187.8017	Prob. Chi-Square(2)	0.0000
Scaled explained SS	483.2872	Prob. Chi-Square(2)	0.0000

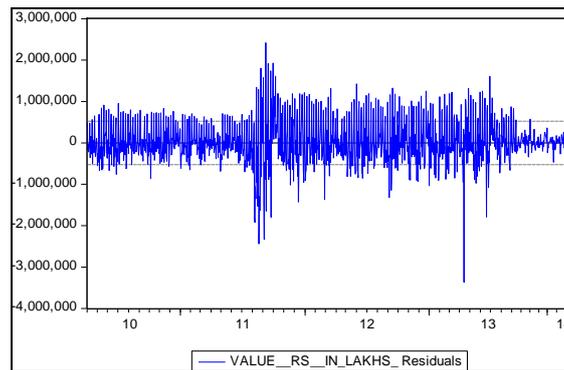


Figure 5: Volatility clustering of residuals of ARIMA(2, 0, 2)

Thus it is imperative to develop better model that will describe the ARCH/GARCH effect. Once again based on the minimum values of AIC and MAPE, ARIMA(2, 0, 2) – GARCH (1,1) model seems to be the best fit.

Table 5: Coefficients of ARIMA (2, 0, 2)- GARCH (1, 1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
VALUE_RS_IN_LAKHS_(1)	0.254167	0.016036	15.84986	0.0000
VALUE_RS_IN_LAKHS_(2)	0.747085	0.016043	46.56687	0.0000
MA(1)	-0.184986	0.027246	-6.789453	0.0000
MA(2)	-0.675880	0.024923	-27.11833	0.0000
Variance Equation				
C	2.34E+10	5.73E+09	4.090749	0.0000
RESID(-1)^2	0.219515	0.042196	5.202271	0.0000
GARCH(-1)	0.703056	0.046804	15.02141	0.0000

The results show both ARCH and GARCH effects are significant. They are the internal causes of volatility in the model. The AIC value of this hybrid model is 7.335 and MAPE value is 26.333 which is less than the ARIMA model developed earlier. The residuals of the is hybrid model is tested for ARCH effect. The Breusch-Pagan-Godfrey test accepted at 5% level the null hypothesis of no ARCH effect (F Statistic = 2.987, P value =0.076 > 0.05); this indicates that the model is better than ARIMA model. The test for normality and serial correlation are also in favour of the hybrid model developed.

IV.4. Model Validation

Sample data from April 1st, 2014 to November 30th 2015 is used to validate the model. The 427 observations are stationary at level (ADF Test, $t = -14.9157$, $p\text{-value} = 0.00 < 0.05$, hence the null hypothesis of unit root is rejected.). The series is slightly positively skewed and leptokurtic (Skewness coefficient = $0.73435 > 0$, Kurtosis coefficient = $4.343 > 3$). Test of Normality rejects that the series is distributed normal (Jarque-Bera = 71.025 , $p\text{-value} = 0.00 < 0.05$)

ARIMA(2,0,2)- AGRCH(1,1) model of the out- of - sample data

Table 6: Hybrid Model developed on the out- sample data

GARCH = C(5) + C(6)*RESID(-1)^2 + C(7)*GARCH(-1)				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
Y'(1)	0.449175	0.031502	14.25841	0.0000
Y'(2)	0.549875	0.031300	17.56777	0.0000
MA(1)	-0.176629	0.040009	-4.414709	0.0000
MA(2)	-0.590296	0.040448	-14.59382	0.0000
Variance Equation				
C	9.45E+08	5.56E+08	1.700870	0.0890
RESID(-1)^2	0.064384	0.031661	2.033512	0.0420
GARCH(-1)	0.880263	0.054993	16.00678	0.0000

It may be observed that at 5% level of significance ARIMA as well as both ARCH and GARCH are significant. The AIC value and MAPE value are 6.37 and 34.23. The Heteroskedasticity of the residuals of the out- of sample data provided in table rejects ARCH effect in the residuals

Table 7: Test of heteroskedasticity on the Out- sample data

Heteroskedasticity Test: ARCH			
F-statistic	2.831119	Prob. F(1,422)	0.0932
Obs*R-squared	2.825580	Prob. Chi-Square(1)	0.0928

The descriptive statistics is provided in Fig.6. shows the residuals are normally distributed (Jarque- Bera = 5.81 , $p\text{-value} = 0.054 > 0.05$) with mean approximately 0.

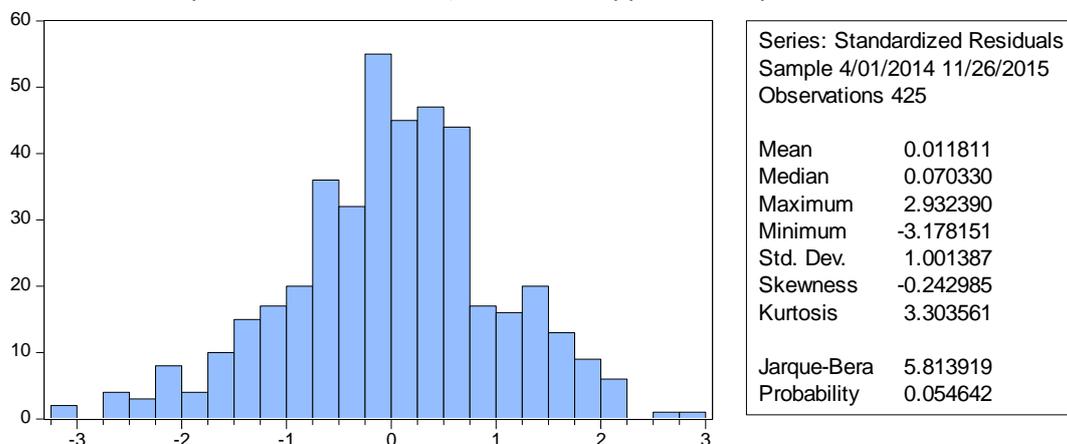


Figure 6: Descriptive statistics of the out- sample data

The test for serial correlation carried out by Ljung- Box Q- test does not reject the possibility serial correlation. The test result is provided in table 8. The results prove that there is no significant serial correlation among the residuals.

Table 8: Ljung- Box Q-Statistic on squared residuals for ARIMA(2, 0, 2)- GARCH (1,1)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob...	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob...			
		1	0.082	0.082	2.8479	0.091				2...	0.022	0.029	16.661	0.675
		2	0.025	0.019	3.1186	0.210				2...	0.009	-0.00...	16.697	0.729
		3	-0.07...	-0.08...	5.6707	0.129				2...	0.002	-0.00...	16.700	0.780
		4	-0.00...	0.010	5.6730	0.225				2...	0.033	0.057	17.182	0.800
		5	-0.07...	-0.07...	7.9923	0.157				2...	0.055	0.047	18.544	0.776
		6	0.043	0.049	8.7851	0.186				2...	0.010	-0.00...	18.587	0.816
		7	0.005	0.001	8.7947	0.268				2...	0.020	0.016	18.777	0.846
		8	-0.05...	-0.07...	10.217	0.250				2...	-0.01...	-0.02...	18.861	0.875
		9	-0.04...	-0.02...	10.993	0.276				2...	-0.02...	-0.01...	19.249	0.890
		1...	-0.04...	-0.03...	11.689	0.306				2...	0.017	0.030	19.387	0.911
		1...	0.011	0.017	11.745	0.383				3...	0.056	0.044	20.815	0.894
		1...	0.018	0.012	11.887	0.455				3...	-0.02...	-0.02...	20.991	0.912
		1...	0.014	-0.00...	11.974	0.530				3...	-0.01...	-0.00...	21.130	0.929
		1...	0.020	0.023	12.158	0.594				3...	-0.05...	-0.04...	22.298	0.921
		1...	0.073	0.071	14.520	0.486				3...	-0.00...	0.010	22.318	0.938
		1...	-0.02...	-0.03...	14.839	0.536				3...	0.077	0.082	25.090	0.892
		1...	-0.05...	-0.05...	16.233	0.507				3...	0.162	0.134	37.389	0.405
		1...	0.003	0.019	16.238	0.576								
		1...	-0.02...	-0.02...	16.452	0.627								

The graph of forecasting of out-sample values and its volatility is provided in Fig: 7.

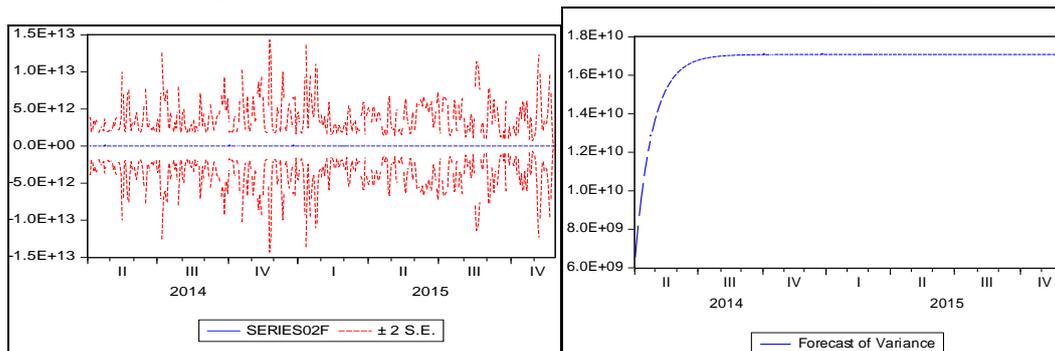


Figure 7: forecasting the values and volatility of out- sample data

From the graph it can be observed that the forecast precision appears to be increased and there is no clustering of volatility. Thus the hybrid ARIMA(2, 0, 2) GARCH(1,1) can be considered as an improved model to forecast the value of gold.

V. Conclusion

This study attempted to capture the volatility of a time series while developing a linear model. Daily value of gold traded through Multi Commodity Exchange (MCX), India, for the past 5 years (April 1st, 2010 to November 30th, 2015) is the data used in the study. A total of 1204 sample data is used as in-sample for modeling and the remaining data of 427 sample points are used for forecasting and testing the validity of the model. The study concludes that ARIMA (2, 0, 2) – GARCH (1, 1) is an appropriate model for forecasting the value of gold in India. The financial time- series of Value of gold traded in MCX, India, is used to construct the model. The model was developed on a sample of 1204 observations and it is validated on an out- sample series of 427 observation which are the daily data over a period of 6 years where there were significant fluctuations in the value of the commodity. The analysis carried out find that the hybrid model is more effective than the usually used ARIMA model for forecasting. ARIMA(2, 0, 2)-GARCH (1,1) is found to be the most appropriate model for the data. It can be concluded that the major drawbacks of ARIMA model is rectified by the hybrid model and it out-performs the usual ARIMA model.

Reference

- [1]. Babu C.N, Reddy B.E, (2015) Selected Indian stock predictions using a hybrid ARIMA-GARCH model, Applied Computing and Information, 11(2), 130-143.
 - [2]. FadhilahYusof, Ibrahim Lawal Kane, ZulkifliYusop; (2013), Hybrid of ARIMA-GARCH Modeling in Rainfall Time Series;JurnalTeknologi, 63(2), 27-34.
 - [3]. G.E.P Box, G. M.Jenkins, (1971),Time Series Analysis: Forecasting and Control Holden-Day, San Francisco, 22(2), 199-201.
 - [4]. MaizahHura Ahmad, Pung Yean Ping SitiYAsir, Nor Miswan (2014), A Hybrid model for improving Malaysian gold forecast accuracy, Int. Journal of Math. Analysis, 8 (28), 1377 - 1387.
 - [5]. MaizahHura Ahmad, Pung Yean Ping SitiYAsir, Nor Miswan, (2015),Forecasting Malaysian Gold using a Hybrid of ARIMA and GJR-GARCH Models,, Int. Journal of Math. Analysis, 9(30),1491 -1501.
 - [6]. R. F. Engle, (2001), An Introduction to the Use of ARCH/GARCH models in Applied Econometrics, Journal of Business, 15(4), 157-168.
 - [7]. Sallem& Khan (2013),The Overview of Gold ETFs and its various positive features; International Journal of Marketing, Financial Services & Management Research; 2(4), 125-135.
-