FORECASTING CEMENT STOCK PRICES USING ARIMA MODEL: A CASE STUDY OF FLYING CEMENT INDUSTRY

Usama Abbasi[1*], Abdullah Almarashi[2], Khushnoor Khan[3], Saman Hanif[4]
[1*]Army Public College of Management Sciences, Department of Commerce, Rawalpindi, Pakistan,
[2,3,4]King Abdulaziz University, Department of Statistics, Faculty of Sciences, Jeddah, Saudi Arabia
[*]Corresponding Author e-mail: khushnoorkhan64@gmail.com

ABSTRACT**:** *This study focuses on forecasting of stock prices for Flying Cement Company using Time Series analysis using Autoregressive Integrated Moving Average (ARIMA) model,. The data for the stock prices were taken from Financial Times website for the period January 2016 to January 2017 (265 days).The data were split into two sets namely–in-sample and out-sample sets where, in-sample set was used for model estimation and out-sample set for cross validation of the results. Eviews and Minitab softwares were used for analyzing the data Graphical representation and statistical tables are also appended. Based on the results of the study ARIMA (1, 2, 1) model was found to be more appropriate for forecasting the time series of the stock prices. The results and the methodology of the current study will assist financial/portfolio managers in forecasting the future trends of stock prices for different companies.*

Keywords: In-sample, Out-sample, ARIMA, cross validation, Forecasting

1. INTRODUCTION:

Cement is the back bone of construction industry. Any country having self sufficiency in cement production has a key role to play in the construction industry not for that particular country but also for the region where that country is located. A case in point is Kingdom of Saudi Arabia and Pakistan which are surrounded by under developed states in dire need of cement in order to complete their infrastructure projects. Like the shares of other companies the share of cement companies are also traded through stock exchanges. Of late there has been a upsurge in megaprojects throughout Pakistan like China Pakistan Economic Corridor (CPEC) which heralded enormous construction activities. Increase in cement purchasing automatically affects the financial position of the cement companies.

Improvement in financial position directly affects the cash flows and in turn improves the stock prices of the commodity. Hence, the need to study the behavior of stock prices of cement industry in Pakistan.Since the prices of any stock are based on time therefore, studying the stock prices entails studying time series data coupled with forecasting. Forecasting involves making estimates of the future values of variables of interest using past and current information. There are a number of methods to generate prediction ranging in intuitive judgments through time series analysis to econometric models [2]. There are many methods that are used to model and forecast time series. The most common classes of time series methods are the autoregressive integrated moving average (ARIMA) models [3].

Two famous econometricians formulated the strategy of forecasting a times series called the Box–Jenkins method named after the statisticians George Box and Gwilym Jenkins, [11] this method applies autoregressive moving average (ARMA) or autoregressive integrated moving average (ARIMA) models to find the best fit of a time-series model to past values of a time series. The Box-Jenkins method (ARIMA) is one of the most widely used time series forecasting methods in practice [4]. One of the objectives of this study is to observe the behavior of stock prices of a cement company in Pakistan – Flying Cement and suggest an appropriate model for forecasting the stock prices.

1.1 Framework of the Paper:

The remainder of the paper is organized into four sections: Section 2 throws light on the methods and material used in the present study; Section 3 discusses the stages in the application of Box-Jenkins technique and presents analysis of data using time series model and selecting the most suitable model to forecast oil production; Section 4 summarizes the results of the current study and draws conclusions; Sections 5 discusses future implications followed by acknowledgements.

2. METHODS AND MATERIAL:

Typically, effective fitting of Box-Jenkins models requires at least a moderately long series. [5] recommends at least 50 observations. Many others would recommend at least 100 observations. For the current study secondary data from the Financial Times website has been selected for analysis. The Flying Cement daily stock prices data were obtained from Financial Times [1] time varying from 1st January 2016 to 30th January 2017. The data are divided into two parts –in sample (from 1st January 2016 – 31st December 2016) and out sample (from 1st January 2107 -20th January 2017). The first data set is used for model estimation and the second set for forecasting and model validation.

3. Stages in Applying Box-Jenkins Method:

The Box-Jenkins (B-J) methodology which calls for the following three steps:

1. Identification
2. Estimation
3. Diagnostic Checking

A point to be kept in mind is that the B-J methodology is applicable only to stationary variables.There are three primary stages in building a Box-Jenkins time series model: model identification, model estimation and model validation. Box-Jenkins [6] uses a statistical procedure to identify a model. The other two steps are quite straightforward.

Stage 1: Model Identification

The first stage involves checking the stationarity of the series through visual examination and formal statistical tools. A point to be noted at this point is that stationarity is a prerequisite for applying Box-Jenkins method. For the current study stationarity will be checked through time series plot of the stock prices along with series Correlogram – Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots. As for the formal statistical tool Augmented Dickey Fuller (ADF) test is carried out. Figure 1 presents the time series, ACF and PACF plots for the stock prices of the flying cement from January 2016 to December 2016 (245 days).Through visual inspection Figure 1 it is clear that the data values do not have a constant mean and variances. Also autocorrelations are declining gradually as the number of lags such a property is common in non-stationary processes. Formal statistical tool for checking stationarity of the time series data is to go for Augumented Dickey Fuller (ADF) test it is used to know when to .

difference time series data to make it stationary. The null hypothesis of the Augmented Dickey-Fuller t-test is

H_0: θ = 0(i.e. the data needs to be differenced to make it stationary) versus the alternative hypothesis of H1: θ < 0 (i.e. the data is stationary and doesn't need to be differenced).

.Regarding formal statistical tests the p-value for the ADF test is 0.532 which suggests that the hypothesis of non-stationarity is not rejected hence the data for the stock prices is non-stationary at the level form

In order to attain stationarity condition the data is subjected to log differencing (first log differencing). The results are shown in Figure1

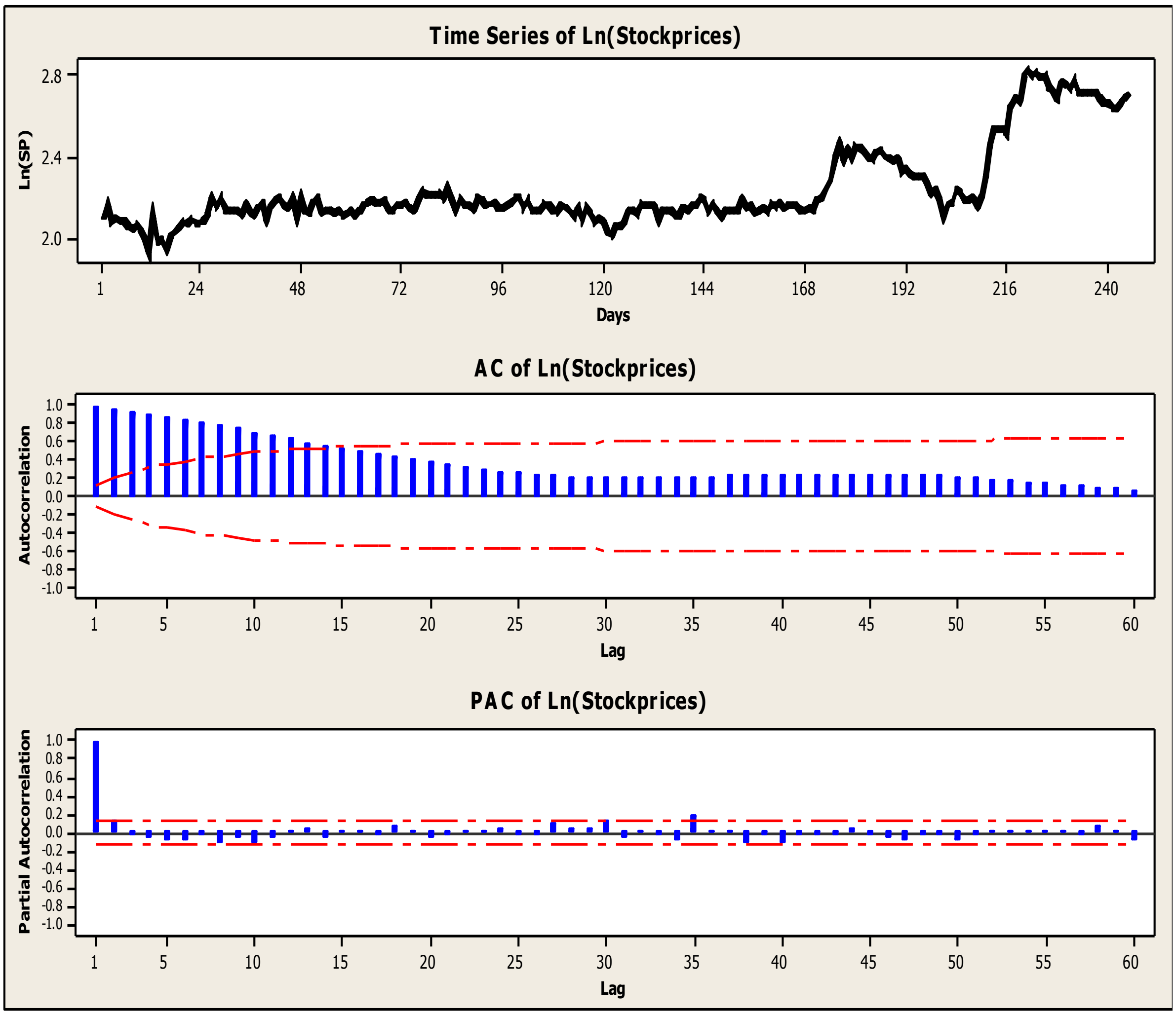

Figure 1: Time Series, ACF and PACF Plots of Flying Cement Stock Shares

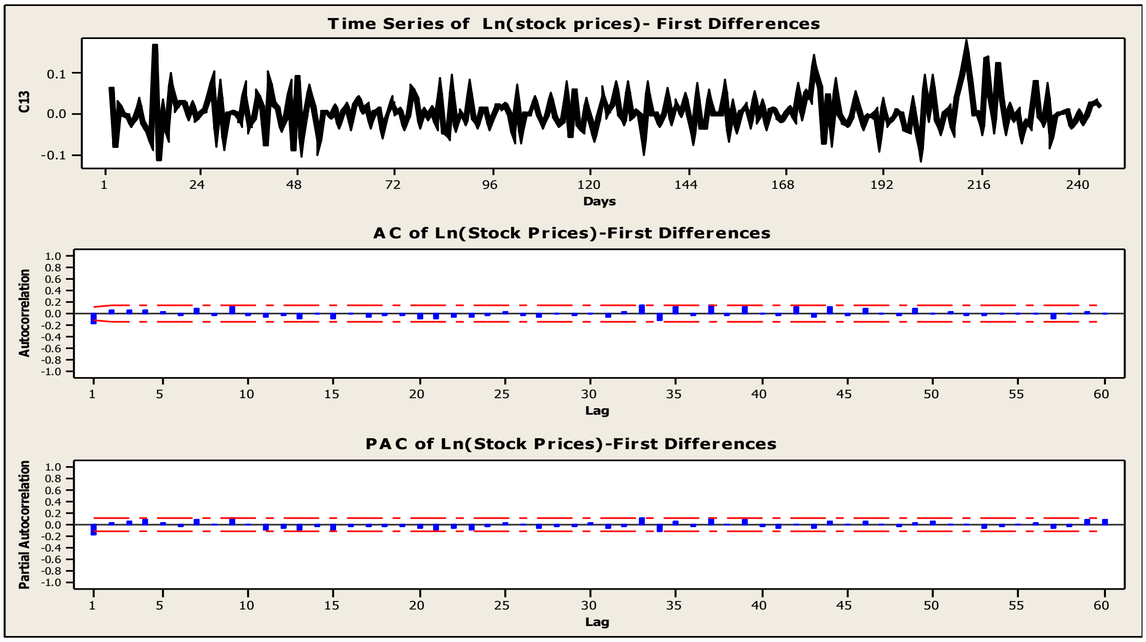

Figure 2: Time Series, ACF and PACF Plots for First Differences

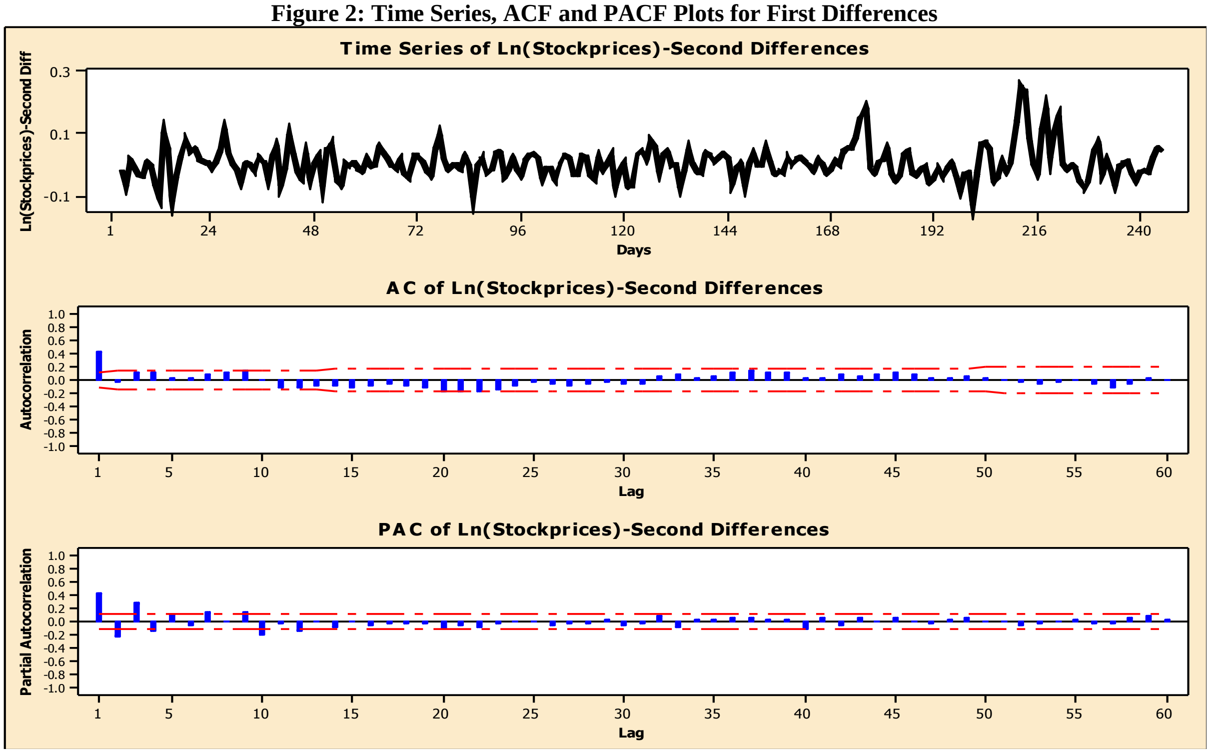

Figure 3: Time Series, ACF and PACF Plots for Second Differences

Though the p-value of the ADF test is 0.01 suggesting a stationary times series but he patterns of the ACF and PACF of the first log differences in Figure 2 are relatively small and mostly lie within the confidence intervals. Therefore, no ARIMA model can be identified from the first log differences of the stock prices of the flying cement as AR(1) and MA(1) were both insignificant with p-values > 0.05. A second order lagged difference from the original series is obtained and the visual results are shown in Figure 3 along with the Correlogram.

General guidelines for selecting the appropriate model are given in Table 1.

The general model introduced by Box and Jenkins (1970) summarized the forecasting model as ARIMA (p,d,q) where p refers to autoregressive parameter, d refers to differencing and q refers to moving average parameter. The values of the parameter are determined by observing the patterns in Correlogram i.e. ACF and PACF plots. The following table gives general rule for identification of the likely model.

Table 1: Model Identification using ACF and PACF

Model	ACF	PACF
AR ()	Geometric Decay	Significant till *p* lags
MA	Significant till *p* lags	Geometric Decay
ARMA	Geometric Decay	Geometric Decay

The p-value of the ADF test for second differences is 0.01 indicating that there is no presence of unit root in the series thus suggesting that the series is now stationary and thus can be used for estimation the model and forecasting. The ACF and PACF plots in Figure 3 suggest that AR(1) and MA(1) will be best suited to the time series plot. Thus the p and q values for the ARIMA (p,2,q) model are set at 1, respectively. Therefore ARIMA model is set to be ARIMA (1,2,1) see table 2. In general the ARIMA model can be expressed as:

$$\mathbf{y_t = 3.49x10^{-5} - 0.1633y_{t-1} + \varepsilon_t - 0.993\varepsilon_{t-1}}$$

Stage 2:Model Estimation

The model estimation was carried using Eviews software the results of the analysis are shown in Table 2
and also we see that the parameters of ARIMA(1,2,1) model are significant.

Table2: Estimation Equation of ARIMA (1, 2, 1)

Variable	Coefficient	SE	t-Statistic	Prob.
C	3.49E-05	3.37E-05	1.037665	0.3005
AR(1)	-0.163363	0.063441	-2.575037	0.0106
MA(1)	-0.993944	0.004474	-222.1717	0.0000
R-squared	0.584324	Akaike info criterion		-3.613289
Durbin-Watson stat	1.997421	Schwarz criterion		-3.570038

Stage 3: Model Validation and Forecasting. This stage entails validation using out of sample data in order to see whether the suggested model fits the data well and also forecasting the stock prices for some future days and comparing them with the actual prices. Both the numerical values and graphical methods are shown for practitioner purposes.

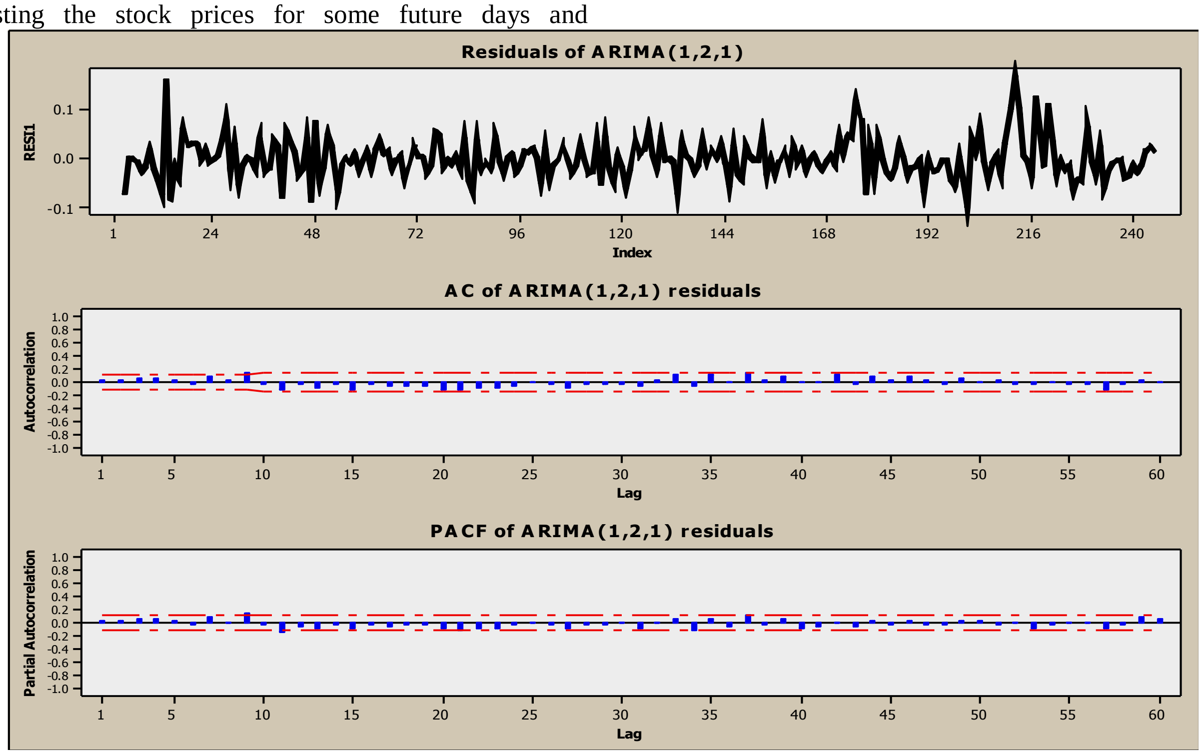

Figure4: ACF and PACF for Residuals of second differences

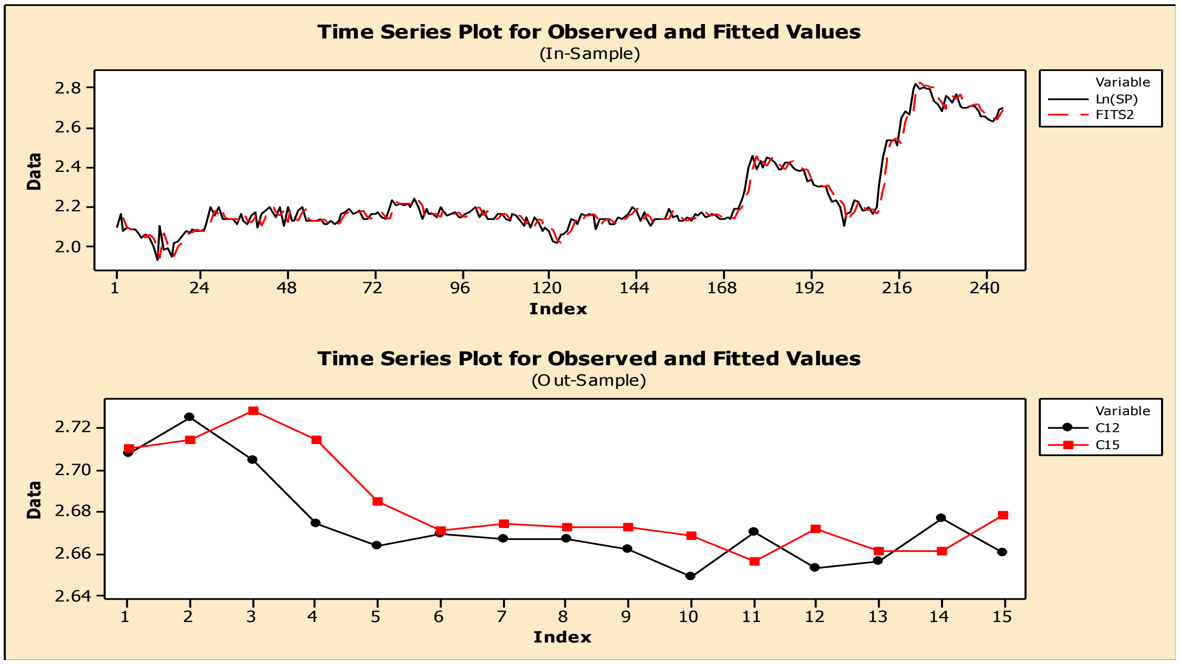

Figure 5 : Time Series Plot for Observed and Fitted Values

Table 3: Cross validation Out Sample (Closing Price) for ARIMA (1, 2, 1)

Dates	Observed	Ln(Observed)	Predicted	Ln(Predicted)
January 02/17	15.00	2.70805	15.05	2.711072
January 03/17	15.26	2.725235	15.10	2.714471
January 04/17	14.95	2.704711	15.32	2.728889
January 05/17	14.50	2.674149	15.10	2.714748
January06/17	14.35	2.66375	14.66	2.685393
January 09/17	14.44	2.670002	14.46	2.671098
January 10/17	14.40	2.667228	14.50	2.674384
January 11/17	14.40	2.667228	14.49	2.673132
January 12/17	14.33	2.662355	14.48	2.672569
January 13/17	14.14	2.649008	14.42	2.668432
January 16/17	14.45	2.670694	14.24	2.656326
January 17/17	14.20	2.653242	14.47	2.671998
January 18/17	14.25	2.656757	14.31	2.661264
January 19/17	14.54	2.676903	14.31	2.660993
January 20/17	14.30	2.66026	14.56	2.678429

4. CONCLUSIONS: The present study was conducted to obtain a suitable forecasting model for stock prices of Flying Cement. ARIMA model has been selected for forecasting stock prices. Table 3exhibits the observed and fitted stock prices which show very little difference in the observed and fitted model.It is suggested that in future for forecasting the time series a Hybrid method which used. But for the scope of the present study ARIMA (1, 2, 1) is the most appropriate model for forecasting purposes.

5. Future Implications:ARCH/GARCH models thus far have ignored information on the direction of returns; only the magnitude matters. However, there is very convincing evidence that the direction does affect volatility. There is now a variety of asymmetric GARCH models, originally developed and suggested by [10] and also including the EGARCH model of [7] the TARCH model— threshold ARCH—attributed to [8] and a collection and comparison by [9].

Acknowledgement: The writers are thankful to Dr. Qaiser Shahbaz for his valuable inputs in improving the manuscript. The writers are also thankful to Deanship of Scientific Research King Abdulaziz University, Jeddah, Saudi Arabia for their guidance and support.

6. REFERENCES

[1].https://markets.ft.com/data/equities/tearsheet/forecasts?s=FLYNG:KAR

[2]. Mad Nasir Shamsudin and Fatimah Mohd Arshad. (2009), Short Term Forecasting of Malaysian Crude Palm Oil Prices. Available:http: // www. econ.upm.edu.my/fatimah/pipoc. Html

[3].Saud M. Al-Fattah (2006), Time Series Modeling for U.S. Natural Gas Forecasting, E-Journal of Petroleum Management and Economics, ISSN: 1718 – 7559.

[4].Bowerman, Bruce L. and O'Connel, Richard T. (1993), Forecasting and Time Series: An Applied Approach, 3rd Ed, Duxbury Press, Wadsworth, Inc.

[5].Chatfield, C. (1996). The Analysis of Time Series, 5th ed., Chapman & Hall, New York, NY

[6].Box, G.E.P., and G. M. Jenkins. 1970. Time series analysis: forecasting and control. Holden Day, San Francisco, CA.

[7].Nelson, Daniel B. 1991. "Conditional Heteroscedasticity in Asset Returns: A New Approach." Econometrica.59:2, pp. 347–70.

[8].Rabemananjara, R. and J. M. Zakoian. 1993. "Threshold Arch Models and Asymmetries in Volatility." Journal of Applied Econometrics January/March, 8:1, pp. 31–49.

[9].Engle, Robert F. and Victor Ng. 1993. "Measuring and Testing the Impact of News on Robert Engle 167 Volatility." Journal of Finance. December, 48:5, pp. 1749–78.

[10].Bollerslev, T. 1986. Generalized autoregressive conditional heteroscedasticity, Journal of Econometrics, Vol 31 pp.307-326

[11]. Box, George; Jenkins, Gwilym (1970) Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day.