Modeling and Forecasting of Gold Prices on Financial Markets

Rebecca Davis
Department of Mathematical Sciences
Pentecost University College
Accra-Ghana.

Vincent Kofi Dedu
Department of Mathematics
Kwame Nkrumah University of Science and Technology
Kumasi-Ghana.

Freda Bonye
Department of Mathematics
University for Development Studies
Navrongo-Ghana.

Abstract

Monthly adjusted close price of gold (112 observed prices) was used for the analysis. An ARMA model was fitted using the first 106 observed prices and the model was used for a 6-step-ahead forecast. The forecast values were then compared to the original corresponding prices. The actual values fell within the forecast limits; and (limits) widened with increasing lead time.

Key Words: ARMA Model, Daniell Spectral Window, Forecast Limit, Gold Price.

1. Introduction

The Gold Coast had long been a name for the region used by Europeans because of the large gold resources found in the area. The Gold Coast achieved independence from the United Kingdom in 1957, becoming the First nation Sub-Saharan African to do so from European Colonialism. The name Ghana was chosen for the new nation to reflect the ancient Empire of Ghana, which once extended throughout much of West Africa. Ghana is endowed with mineral deposits such as gold, diamond, manganese and bauxite. Of all the minerals mined from the earth, none is more useful than gold. Its usefulness is derived from a diversity of special properties. Gold conducts electricity, does not tarnish, malleable and very easy to work with, can be drawn into wire, can be hammered into thin sheets, alloys with many other metals, can be melted and cast into highly detail shapes, has a wonderful color and a brilliant luster. Gold is a memorable metal that occupies a special place in the human mind.

Gold prices spanning a period of 124 months were subjected to Time Series Analysis. The goal was to model these prices, forecast the future prices and compare observed (Actual) prices with forecast values to ascertain the model robustness and predictive strength.

2. Data Analysis

A monthly adjusted close price of gold from January 2003 to April 2012 (112 observed prices) is used for the analysis. The data is freely available for download at Yahoo Finance web site. The time series plot of the data is represented graphically in Figure 1 below.

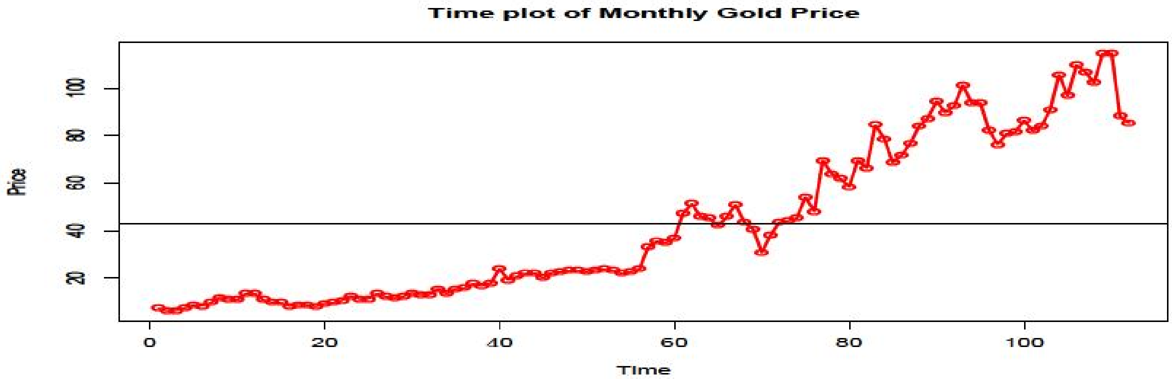

Figure 1: Monthly Price of Gold

Looking at the time series plot, there is an increasing trend in the prices with higher values displaying more variation. By taking logarithm of the data, the variance becomes stable but there is still an increasing trend. As a result, we take the difference of logarithms. The differenced logarithm series (called returns in finance) shown in Figure 4 (see appendix) looks much more stationary when compared with the original time series.

By performing the Augmented Dickey-Fuller unit-root test on the monthly returns, the ADF test statistic is -4.5433 for lag order 4 and a p-value of 0.01 is recorded. With stationary as the alternative hypothesis, we reject the null hypothesis that there is a unit-root in the series. Hence we conclude that the return series is stationary.

3. Model Specification

We find a model that best fits the first 105 observations of the return series. Figure 5 in the appendix shows the estimates of the spectrum using a modified Daniell spectral window convoluted with itself and a span of 5 for both the solid line, [where we have also drawn the 95% confidence limits (dotted lines)] and the dash line [the estimated spectrum using an AR model with the order chosen to minimize the AIC value]. Estimating the spectrum of the return time series using a modified Daniell spectral window convoluted with itself and a span of 5 for both, the frequencies 0.0667 and 0.195 are significant. The estimated spectrum using an AR model with the order chosen to minimize the AIC value agrees with these frequencies (order 9). This suggests one cycle period of 15 months and another cycle period of 5 months.

The estimated autocorrelation and the partial autocorrelation functions (ACF and PACF respectively) for the return series are illustrated in Figure 2 below.

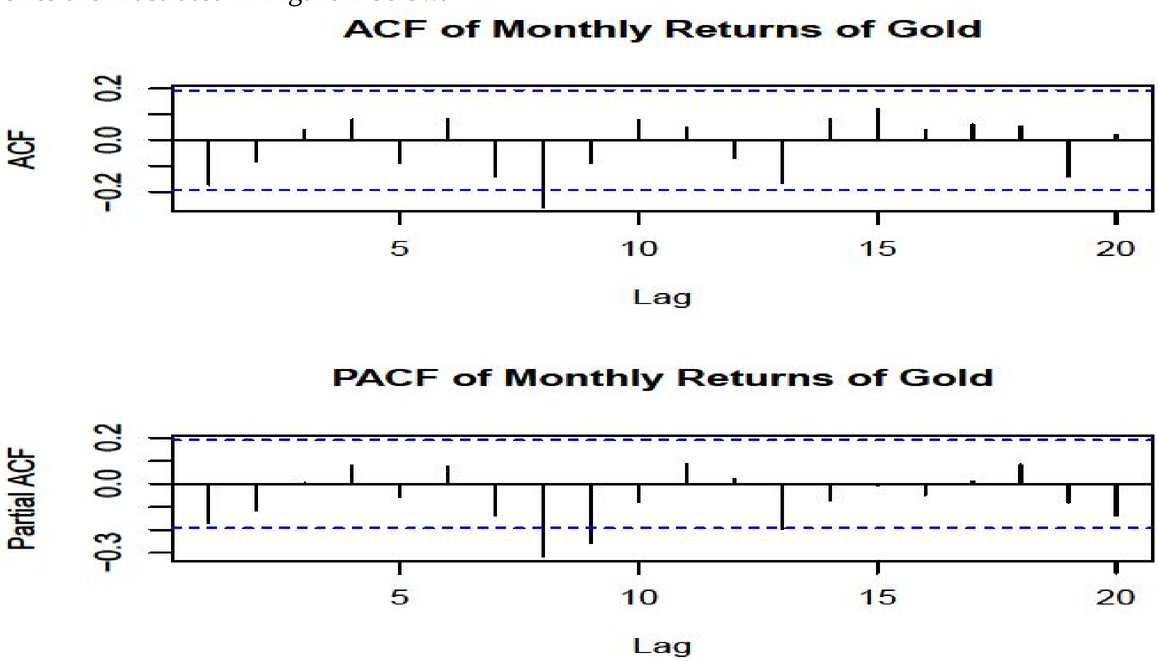

Figure 2: ACF and PACF of monthly gold returns

The ACF suggests an MA(7) model since the autocorrelation is significantly different from zero at lag 7. The PACF also gives a strong evidence to support an AR(9) model. However, none of these plots is very useful in detecting the order of ARMA models. It is very difficult to interpret the plot of extended autocorrelation functions of this series so we use the best subset ARMA approach to specify a model for this data. Results are displayed in Figure 6 (see appendix). From the results, the suggestion is that, the difference in logarithms of gold prices (w_t) should be modeled in terms of w_{t-7}, e_{t-1} and e_{t-8}. We therefore go ahead and investigate further, a subset of ARIMA(7,1,8) model on the logarithms of gold prices(y_t) where y_{t-7}, $e_{t-1}$and e_{t-8} have non-zero coefficients. The estimated coefficient of $y_{t-7}$for this model is not significant so we fit a new model without this coefficient. The new model looks good but from the best subset results displayed in the appendix, we see that this new model is the third best model so we consider the second best model suggested. This model is a subset of ARIMA(7,1,10) in terms of y_{t-7}, e_{t-1}, e_{t-8} and e_{t-10} and it has the minimum AIC value(-154.68) and the maximum log-likelihood value (81.34) among all the fitted models. Also, all the parameters are significant so we choose this model as the best model for the logarithm of gold prices. Using the MLE method of estimation for the parameters, the model for the logarithms of gold prices is given in Equation (1).

$$y_t = -0.1677y_{t-7} + e_t - 0.1276e_{t-1} - 0.2862e_{t-8} + 0.2524e_{t-10} \quad (1)$$

4. Model Diagnostics

To check whether the model assumptions are supported by the data, we take a look at Figure 3. From the plot, the standardized residuals look random, there is no significant autocorrelation in the residuals except at lag 13 and the p-values of the Ljung-Box test up to lag 20 are greater than 0.05. This implies that, the model fits well.

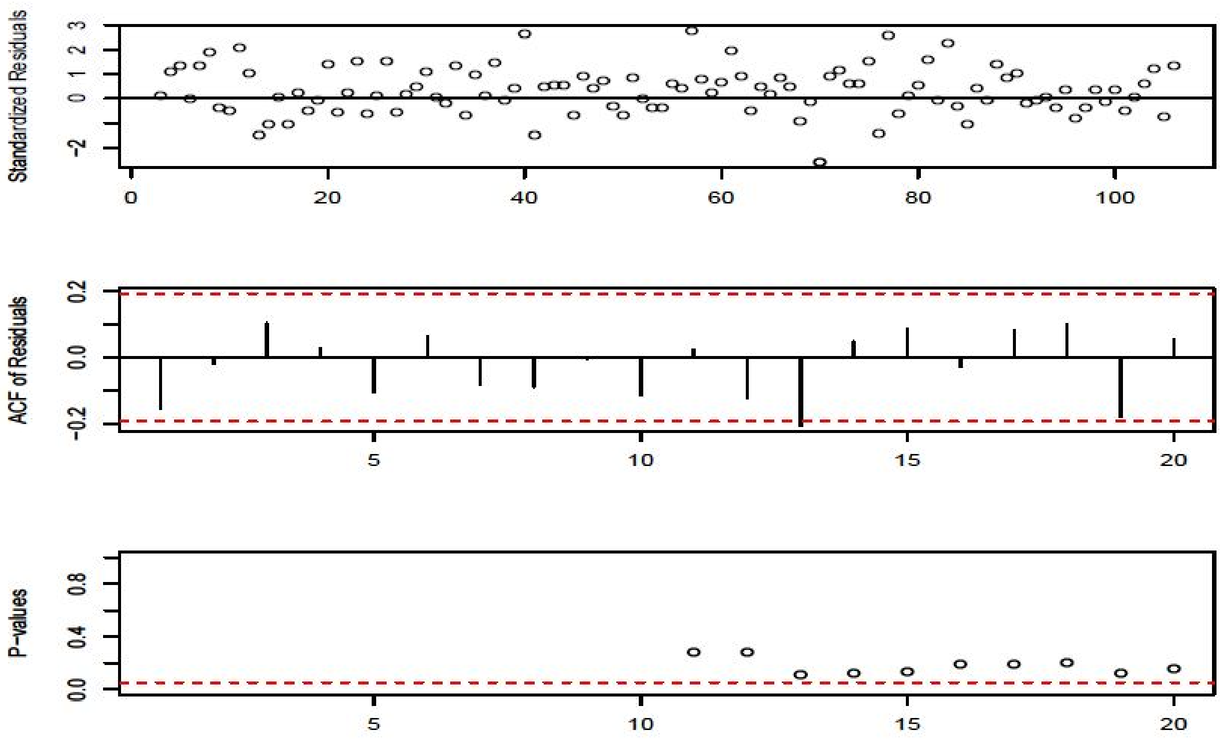

Figure 3: Diagnostic Display for the fitted ARIMA Model of Monthly Gold Prices

The Q-Q plot of the residuals in Figure 7 (see appendix) indicates that the tails are lighter than a normal distribution. We therefore do a Shapiro-Wilk test. The Shapiro-Wilk test records a p-value of 0.682, which does not reject normality of the error terms at any of the usual significance levels. We continue the model diagnostics by including y_{t-8} in the model to test for over fitting. Results from the software shows that, it is not only the coefficient of y_{t-8} that is not significant but also, other coefficients which were previously significant become insignificant and the AIC value increases. We can therefore say that, the chosen model in the previous section is the best and we proceed to use the model for forecasting.

5. Forecasting

The fitted ARIMA model is used to do a 6-step-ahead forecast. The forecast values are then compared to the original corresponding prices. Table 1 provides details of the forecast values. All the actual values fall within the forecast limits and the limit becomes wide as the lead time increases. Figure 8 (see appendix) displays the last two years of the monthly gold prices together with 95% forecast limits for six additional months and the actual gold prices in these months.

Table 1: Forecasting Results

Time	Actual	Forecast	Lower limit	Upper limit	Forecast Error
107	106.91	105.74207	85.12956	131.3455	1.167932
108	102.10	106.35153	79.75878	141.8107	-4.251535
109	114.41	107.15382	75.93778	151.2020	7.256179
110	114.73	106.68603	72.02437	158.0286	8.043975
111	87.98	100.68859	65.10369	155.7238	-12.708587
112	85.29	98.57158	61.28103	158.5541	-13.281580

6. Discussion of Results and Conclusion

The forecast limits provide us with a clear measure of the uncertainty in the forecasts. It is easier to see that the forecast limits spread out and the forecast error increases in absolute value as we get further into the future. Updating the forecast once the observations at previous times have been observed, did not cause much change. (Example, updating $\widehat{Y}_t(l)$ to $\hat{Y}_{t+l-1}(1)$ for $l > 1$ when $Y_{t+1}, Y_{t+2}, \dots, Y_{t+l-1}$have been observed). Table 2 shows the actual values, forecast values and updated forecast values. The mean of the forecast errors before updating is -2.2956 and the mean after updating is -2.2716.

Table 2: Forecasting Results

Time	Actual	Forecast	Forecast Update
107	106.91	105.74207	105.74207
108	102.10	106.35153	106.20250
109	114.41	107.15382	107.67730
110	114.73	106.68603	105.82694
111	87.98	100.68859	99.55256
112	85.29	98.57158	100.04824

Talking about the prediction of direction of price movements, (whether the price of gold will increase or decrease at a particular time in the future) the fitted model was 66.67% correct. To get a more accurate forecast results, it is advisable to start from the scratch and fit a new model (if necessary) to a new data which contains currently observed gold prices.

References

Ngai Hang Chan, “Time Series Applications to Finance," John Wiley & Sons, Inc., 2008.

Jonathan D. Cryer and King-Sik Chan, “Time Series Analysis with applications in R," Second Edition Springer Science+Business media, LLC, 2008.

E. Kosciełny-Bunde, H. E. Roman, A. Bunde, S. Havlin and H. J. Schellnhuber, “Long-range power-law correlations in daily temperature fluctuations," Phil. Mag. B 77, 1998, 1331-1339.

Y. H. Liu, P. Cizeau, M. Meyer, C. K. Peng, H. E. Stanley, “Quantification of Correlations in Economic Time Series," Physica A 245, 1997, 437-440.

R. N. Mantegna, H. E. Stanley, “Scaling Behaviour in the Dynamics of an Economic Index," Nature 376, 1995, 46-49.

Robert A. Yaffee with Monnie McGee, “An Application to Time Series Analysis and Forcasting with Applications of SAS and SPSS," Academic Press, Inc., San Diego London, 1999.

Appendix

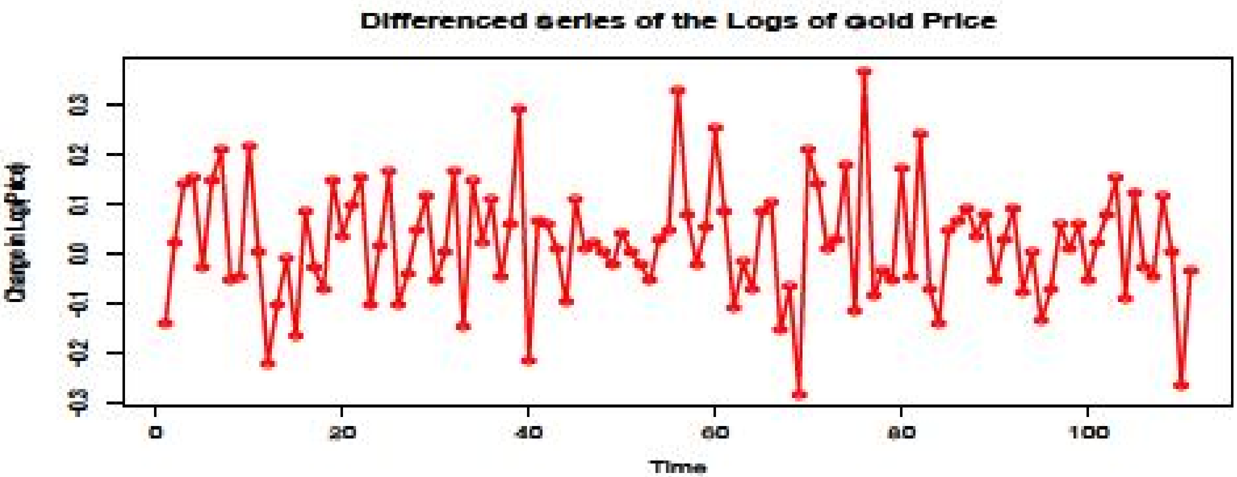

Figure 4: Differenced Series of the Logs of Gold Price

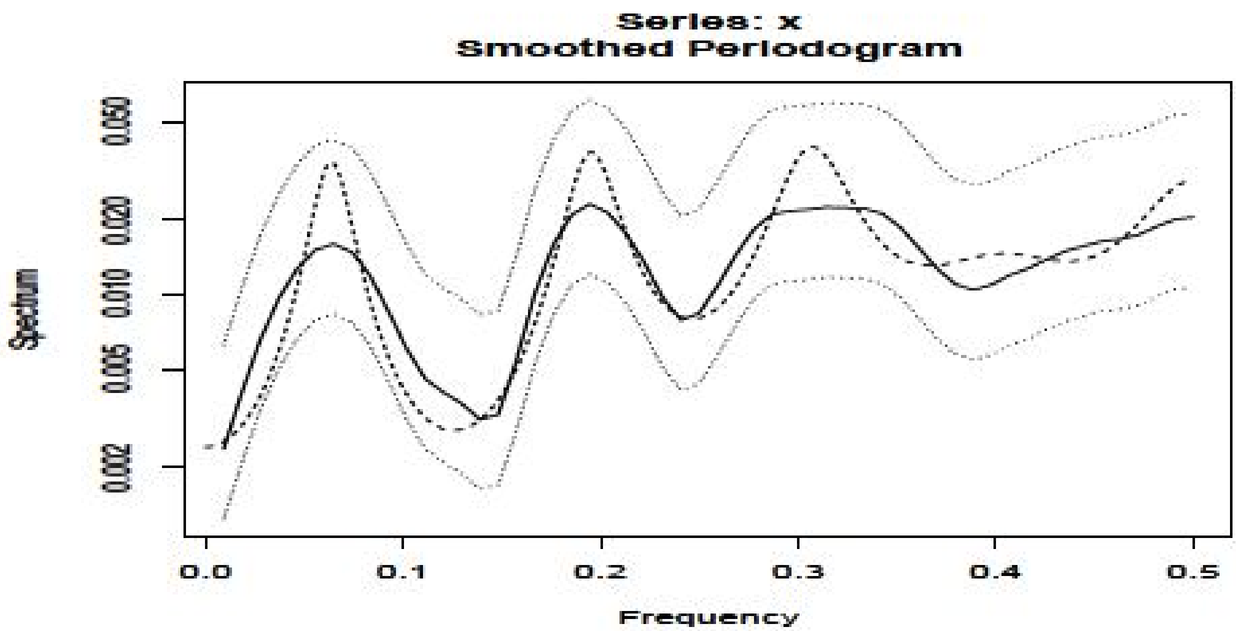

Figure 5: Estimated Spectrum for monthly gold returns

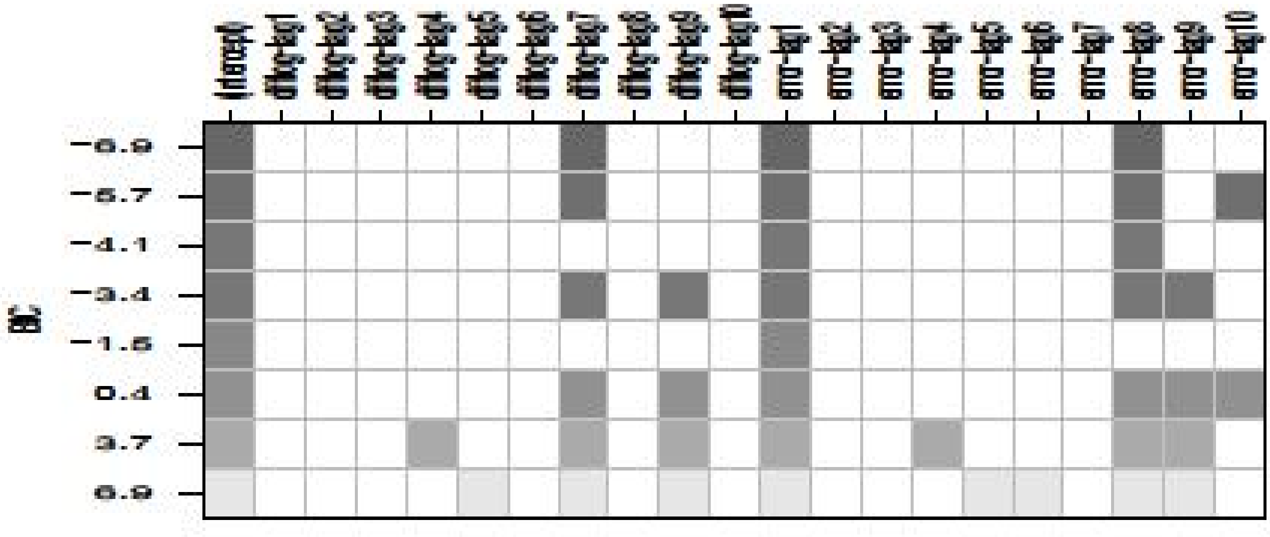

Figure 6: Best Subset ARMA model for Differenced of Log(Gold)

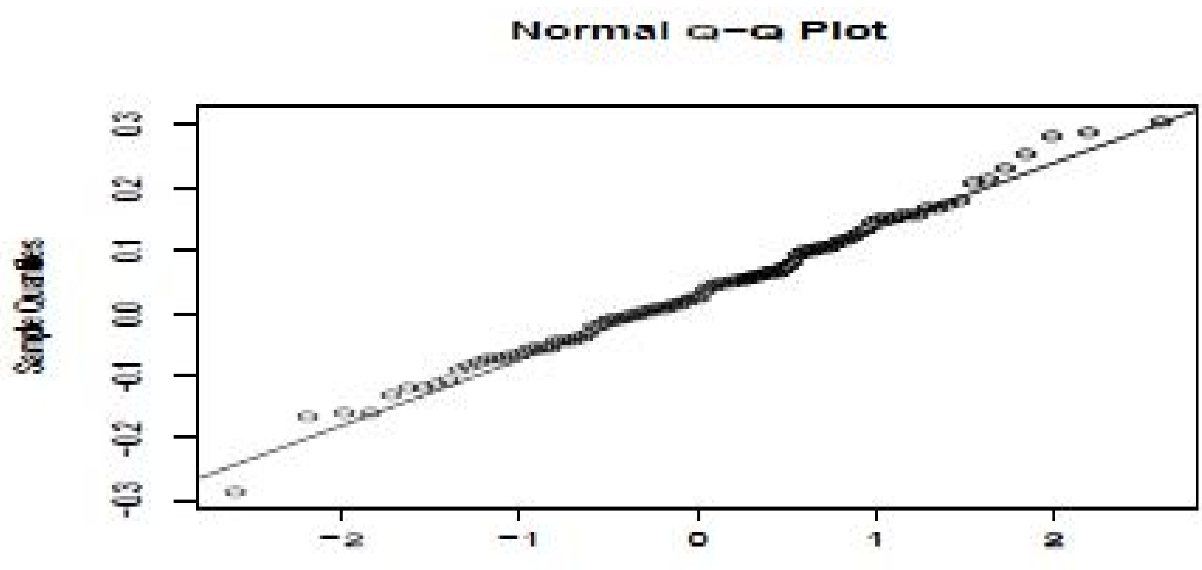

Figure 7: Residuals from the fitted ARIMA Model for Monthly Gold Prices

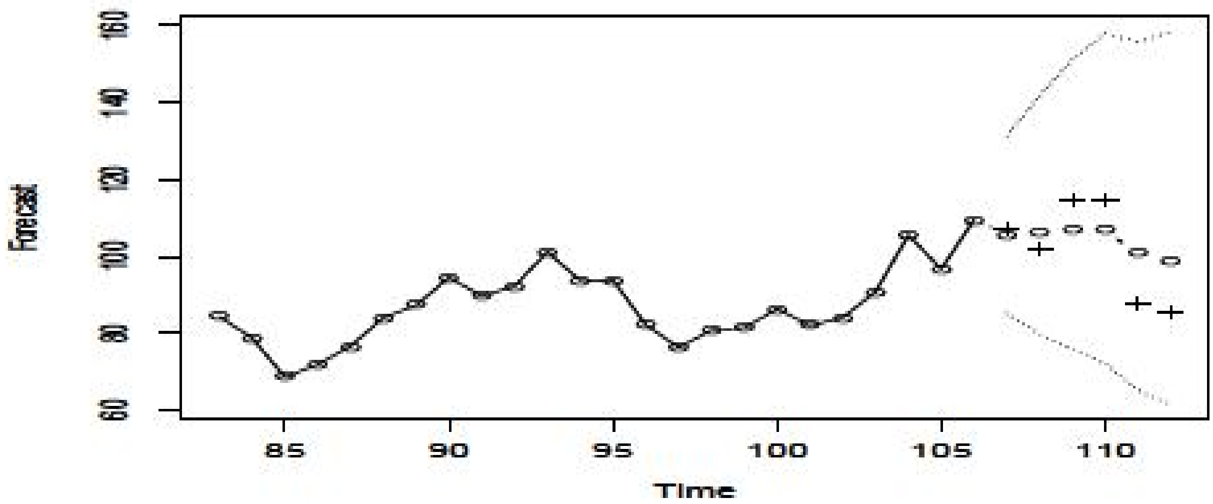

Figure 8: Forecasts, Forecast Limits and Actual Monthly Gold Prices