Advances in Social Science, Education and Humanities Research, volume 196
2nd International Conference on Social Science, Public Health and Education (SSPHE 2018)

The Prediction of Gold Price Using ARIMA Model

Xiaohui Yang*
IBSS
Xi'an Jiaotong-Liverpool University
Suzhou, China
Xiaohui.Yang17@student.xjtlu.edu.cn

***Abstract*—Although,2016 and 2017 have risen, the international gold price has been in the doldrums since 2013. The volatility of gold prices will have a profound effect on the investment decisions of individuals, enterprises and countries. This study focuses on the figure of gold prices from July 2013 to June 2018 according to the World Gold Council, and aims to forecast and analyze daily gold price of USD in the first half of the month of July 2018 through the establishment of ARIMA model. This study also uses AC, PAC, AIC, BIC to estimate the accuracy of models. Empirical outcomes demonstrate that ARIMA (3, 1, 2) is the finest model to predict the gold price of USD. The estimate outcomes of ARIMA Model are vital for people to understand the efficiency of gold prices and make great investment choices.**

Keywords—Forecast, gold price, ARIMA model

I. INTRODUCTION

As a precious metal, gold can also be considered as the stock contract products in the commodity market. In the world economic crisis in 2008, a variety of financial products have performed weakly. However, only the gold market had outstanding performance. There are many factors influencing the changes of gold price, for example, the dollar exchange rate, inflation and monetary policy. Scholars at home and abroad generally set up regression models from these factors. Time series analysis method used in this paper is to model and analyze the historical data of the financial time series. This article centers on the figure of gold prices since July 2013 to June 2018 according to the World Gold Council. It is based on the correlation theory of time series to establish the ARIMA (3, 1, 2) Model to forecast the trend of gold price. In addition, it aims to forecast and analyze daily gold price in the first semi-month of July 2018 through the establishment of ARIMA model. This article is composed of five parts, the structure of this article is established as the following: segment 2 describes the literature review, sector 3 defines sample and the methodology that is ARIMA model. Besides, segment 4 positions the consequences, section 5 puts the conclusion.

II. LITERATURE REVIEW

There are a large number of researches and studies in regard to the prediction of gold price, for which remains a popular topic in both literature and industry. Nambier (2012) [1] applied ARIMA model in forecasting gold price. By pointing out the limitation of ARIMA model, it can just forecast immediate future. Pung et al. (2013) [2] used ARIMA model and GARCH model to predict gold price of Malaysia and reached a conclusion that GARCH is more appropriate than ARIMA in simulating changes in the volatility of time series variables. ARIMA is sufficient when the article does not predict the volatility of variables. Baur (2016) [3] used Dynamic Model Averaging (DMA) method to evaluate possible gold price determinants and predict gold price when the forecasting model and the parameter were uncertain. There are also some hybrid models being used to predict gold price. Ye (2014) [4] established ARFIMA-GARCH model to analyze and predict gold price and this results were proved to be highly accurate. Mombeini and Yazdani-Chamzini (2015) [5] adopted ANN (Artificial Neural Network) model and ARIMA model to forecast gold price. As a consequence, ANN model had better performance. Kristjanpoller and Minutolo (2015) [6] used ANN to GARCH model and generated a mixture ANN–GARCH model to foresee the precariousness of spot price and future price of gold. Sharma (2016) [7] stated that ARIMA (3,1,3) is the best appropriate model to forecast gold price of India through applying Box Jenkins Autoregressive Integrated Moving Average method. Hence, it is essential to choose an accurate and appropriate model for the prediction of gold price.

III. DATA AND METHODOLOGY

A. Data Sources

This study predicts the international gold price that is priced in US in the first half of the month of July 2018. The data are collected from the World Gold Council, consisting of 1305 observations of daily gold price from July 1st 2013 to June 29 2018. The sample is divided into two datasets, which are in-sample data and out-of-sample data. In specific, the in-sample data for the construction of model includes 1305 observations from July 1st 2013 to June 29th 2018. In terms of the out-of-sample, there are 10 observations from April 2nd 2018 to April 13rd 2018.

B. Data Description

Fig. 1. Daily gold price from 20130701 to 20180629

Fig. 1 displays the daily gold price from July 1st 2013 to June 29th 2018. It shows that gold price fluctuates during this period between 1,000 USD and 1,400 USD.

Copyright © 2019, the Authors. Published by Atlantis Press.
This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

There are many factors that influence the sustained weakness of international gold price in 2014, including the strong dollar rally, the fulfillment of QE in the US and the rise of the fed's interest rate. Gold prices did not experience a slump during the whole year, and even saw a slight increase at the end of the year. Apparently, it can be seen that the most dramatic change happens during 2015 after that the Federal Reserve announced interest rate increase four times. The overall structure of gold prices in 2015 shows a declining trend, with occasional strong rises. In 2016, the global macroeconomic environment has been improving generally, and the inflation expectations of major economies have rebounded. The investment preference of global market for gold has improved significantly. In 2017, the demand for hedging caused by political uncertainty can be the most important factor driving the periodic rise in gold prices. Overall, from 2014 to 2017, despite significant increases in 2016 and 2017, gold price has failed to exceed 1,400 dollar for four consecutive years.

C. Methodology

Box and Jenkins' (1976) [8] state that ARIMA model appears to be one of the broadly applied models in forecasting gold price. This model attempts that the future value of time series data ensures a practical connection with current value and historical value. Analysts, exporters, importers, international firms, speculators and dealers in the foreign exchange market all believe that patterns in the past can afford an implication of future measure in the short term.

The Growth of ARIMA model includes mainly three stages for instance preprocessing of time serie`s, identification, estimation, diagnostic checking and forecasting. In the single-variable model, the first step is to check whether gold prices are stationary through Augmented Dickey-Fuller (ADF) unit root test. The assumptions are:

H_0: There is a unit root

H_1: There is not a unit root

The ADF assessment is performed on the original data, and the test result is shown as Table I. It can be seen from Table I that null assumption cannot be rejected at the 5 percentage significance level. The gold price sequence is a non-stationary process. There are unit roots and the data is not stable. So the data needs to be processed differentially to remove the tendency and seasonality of the data.

TABLE I. ADF TEST RESULTS

Number of obs: 1304		**t-Statistic**	**Prob.**
Augmented Dickey-Fuller test statistic		-2.738	0.0677
Critical values	1%	-3.430	
	5%	-2.860	
	10%	-2.570	

The first-order differential was performed on the original data. Then the ADF test was applied to inspect the smoothness of the data after the first-order difference. The test consequences are shown in Table II. It can be gotten from the test consequences that all the significance levels were rejected from the original hypothesis. There was no unit root and the data was stable.

TABLE II. ADF TEST RESULTS

Number of obs: 1303		**t-Statistic**	**Prob.**
Augmented Dickey-Fuller test statistic		-36.693	0.0000
Critical values	1%	-3.430	
	5%	-2.860	
	10%	-2.570	

After passing the stationary test, the d-value in the ARIMA (p, d, q) model has been decided. Due to that the stationary requirement has been achieved after the first-order differential, the d value of the model is 1. The following stage is to detect the model. The p and q values of the ARIMA (p, 1, q) model need to be identified. The value recognition of p and q can be gained through the observation of auto correlation function (ACF) and partial auto correlation function (PACF) of the sample.

Both of ACF and PACF are applied to figure out decisive evidence of smoothness state. If the AC slowly decreases, PAC shows a sudden break after one lag, then this model is AR. Alternatively, if AC tends to be zero after one lag, PAC shows a zigzag decline, then this model is MA.

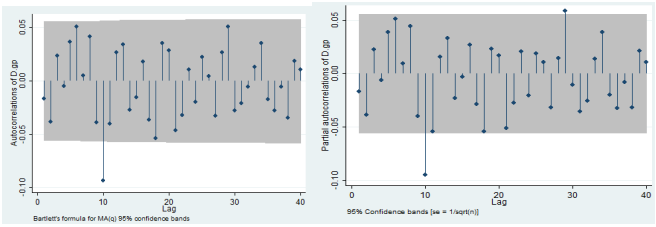

Fig. 2. Autocorrelation function and partial auto correlation function

Fig. 2 displays that both ACF and PACF does not show significant tailing and truncation, so they were judged as ARIMA models.

Table III shows the summary results of ARIMA models. The value of AIC and BIC in Table III shows that gold price besides the model ARIMA (3, 1, 2) when regressed on historical gold price because the AIC is 9947.061 and the BIC value is 9983.273 which are minimum based on comparison of models. It indicates that the order of auto-regressive progression is 3, the order of the moving average progression is 2 and the order of the difference is 1.

It can be found in Table IV that except for the parameter of y*L^3, all other parameters are quite significant. the overall

parameter significance test is also proved to be significant, which implies that the variables in the first column variables exactly shows their influence.

TABLE III. COMPARISON OF DIFFERENT MODEL INDICATORS

Model	LL	AIC	BIC
ARIMA(1,1,1)	-4973.204	9954.409	9975.101
ARIMA(1,1,2)	-4972.194	9954.388	9980.254
ARIMA(1,1,3)	-4971.429	9954.858	9985.898
ARIMA(1,1,4)	-4972.152	9958.305	9994.517
ARIMA(3,1,3)	-4966.262	9948.524	9989.909
ARIMA(2,1,2)	-4971.57	9955.14	9986.179
ARIMA(3,1,2)	-4966.53	9947.061	9983.273

TABLE IV. MODELING RESULTS OF ARIMA (3, 1, 2)

D.goldprice		Coef	SE	z	p>[z]	[95% Conf. Interval]	
_cons		0.0035	0.3069	0.01	0.991	-0.598	0.605
ar	L1	1.4325	0.0584	24.54	0.000	1.318	1.547
	L2	-0.910	0.0598	-15.23	0.000	-1.027	-0.793
	L3	0.029	0.0294	0.98	0.325	-0.029	0.087
ma	L1	-1.457	0.0479	-30.40	0.000	-1.551	-1.363
	L2	0.906	0.0485	18.68	0.000	0.011	1.002
sigma		10.91	0.1369	79.69	0.000	10.64	11.18

Based on Table III and Table IV above, we can determine that choosing the ARIMA (3,1,2) model as the best appropriate gold price forecasting model, its corresponding expression is:

$$Y_t = 0.0035 + 1.4325Y_{t-1} - 0.910Y_{t-2} + 0.029Y_{t-3} + \varepsilon_t - 1.457\varepsilon_{t-1} + 0.906\varepsilon_{t-2}$$

IV. EMPIRICAL RESULTS AND DISCUSSION

The ARIMA(3,1,2) model is used to predict gold price. The results of in-sample estimation and out-of-sample estimation are as follows.

The in-sample prediction results of international gold price on 25th March 2018 to 29th June 2018 are listed in Table V. It can be obtained from Fig. 3 the relative error between the predicted values and the actual values are all relatively small, less than 1.2%. It indicates that the model has good prediction results.

From the in-sample chart, as can be seen that the actual gold price trend rises first, then declines, and finally stabilizes. The predicted goldprice_hat line is below the actual gold price trend line, indicating that the in-sample forecasting data is lower than the actual data, and the forecasting data is on average about $1.25 lower than the actual data. At the same time, within the time period predicted in the sample, the predicted trend maintains a relatively small range of stability.

TABLE V. IN-SAMPLE PREDICTION RESULTS OF GOLD PRICE

Date	Gold price	Gold price_hat	Relative error
2018/6/25	1268.70	1283.11	1.136%
2018/6/26	1260.30	1246.02	-1.133%
2018/6/27	1254.60	1241.81	-1.020%
2018/6/28	1251.55	1254.95	0.272%
2018/6/29	1250.45	1253.44	0.239%

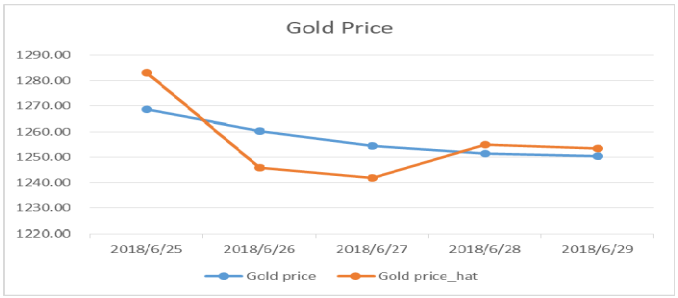

Fig. 3. The comparison of predicted results and actual figure of the in-sample data

Next, using the model ARIMA (3,1,2) to forecast gold price from July 2 to July 13, 2018. The results of the out-of-sample prediction are shown in Table VI.

According to the out-of-sample chart, it can be found out that the results of the out-of-sample prediction are approximately the same as the results predicted in the above-mentioned sample. Out-of-sample forecasting data is on average about $0.46 higher than actual data. The actual gold price out of the sample rose first and then fell in the short term and remained stable in four trading days, and then rose rapidly again. The out-of-sample forecast trend line is lower than the actual gold price trend line. Meanwhile, the forecast price of gold price remains basically stable within the predicted time range.

Moreover, it can be seen from Fig. 4 that there was a drop on July 6 and this forecast result is in line with the downward trend of actual international gold prices. However, at the same time, it can be seen that as the forecast period increases, the relative error of the model prediction also becomes larger, which also indicates that the ARIMA model is more suitable for short-term prediction.

In addition, it is necessarily admitted that there is a gap between the prediction results and the actual data of this model.

Because the ARIMA model only predicts the price of gold based on its own time series of gold prices, the prerequisite for this model to be established is that other influencing factors do not change much and the period is short. However, there are various factors affecting analysis and forecast of international gold prices, such as product operating cost, gold supply and demand, oil prices, dollar exchange rates, inflation, stocks prices, and international political situation. Additionally, these cause factors actually interact with each other and then have a significant impact on gold price. Therefore, it is hard to forecast gold price as its generation process involves significant elements, which can be regarded as an extremely complex system and a non-linear problem.

The characteristic of the time series itself is to extract useful information from historical data to predict future trends and other factors affecting prices. The other factors affecting the settlement price are only reflected in a random manner. This is also a shortcoming of the time series model. This research cannot control other influencing factors. When the data is selected, it has chosen to avoid policies and other factors which may cause significant fluctuations.

TABLE VI. OUT-OF-SAMPLE PREDICTION RESULTS OF GOLD PRICE

Date	Forecast	Reality	Relative error
2018/07/02	1249.09	1247.80	0.103%
2018/07/03	1247.91	1251.75	-0.307%
2018/07/04	1249.81	1255.65	-0.465%
2018/07/05	1250.66	1255.50	-0.386%
2018/07/06	1247.36	1255.35	-0.636%
2018/07/09	1271.44	1262.05	0.744%
2018/07/10	1270.56	1254.00	1.320%
2018/07/11	1261.73	1251.40	0.826%
2018/07/12	1260.61	1245.90	1.181%
2018/07/13	1216.54	1241.70	-2.027%

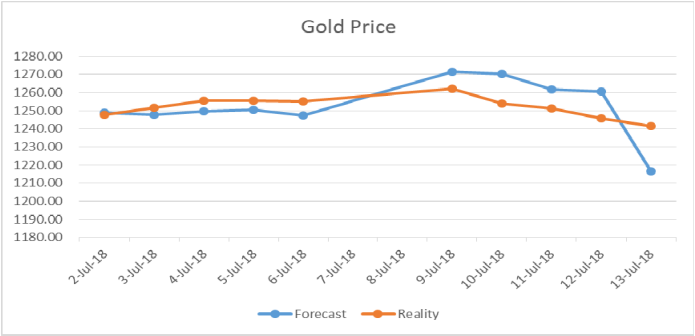

Fig. 4. The comparison of predicted results and actual figure of the out-of-sample data

V. CONCLUSIONS

This research paper mainly aims to investigate and carry out the prediction of the future international gold price to determine the most accurate and appropriate model for forecasting. By using ARIMA model with the daily time series data, this study briefly predicts the future international gold price. Based on the study results shown, ARIMA (3, 1, 2) appears to be the best model for predicting gold price. This research indicates that ARIMA model can be applied to predict the international gold price. Furthermore, this article selected the international gold closing prices from July 1, 2013 to July 29, 2018. After multiple screenings, the ARIMA (3, 1, 2) model was chosen to analyze and forecast the gold price. The results show that the model used can reflect the true gold price trend to a certain extent and can provide consumers with a certain guiding role in gold investment. However, due to the limitations of the model itself and the complexity of the real world, there is still a certain gap between this model and the real data, which cannot be fully accurately reflected and forecasted. Therefore, in the future when we conduct related research on gold prices to more comprehensively consider various influencing factors and improve the prediction model. With respect to the empirical analysis results of this study, several significant facts can be found. Initially, other uncertain factors especially policies do have significant effect on the international gold price, which is quite difficult to forecast. Additionally, the tendencies of the international gold price of the in-sample data and out-of-sample data shows different directions, which indicates different cause factors. Thus, in the future study, it might be helpful to establish a model by adding dummy variables to explain and investigate the impact of non-numeric factors on the international gold price.

REFERENCES

[1] G, D.M., G. Nambiar, and R. M, "Forecasting Price And Analysing Factors Influencing The Price Of Gold Using Arima Model And Multiple Regression Analysis," International Journal of Research in Management, Economics and Commerce, vol. 2, no. 11, pp. 548-563, 2012.

[2] Pung, Y.P., Nor, H.M., Maizah, H.A, "Forecasting Malaysian gold using GARCH model," Applied Mathematical Sciences, vol. 7, no. 58, pp. 2879-2884, 2013.

[3] Baur, D. G., Beckmann, J., & Czudaj, R, "A melting pot - gold price forecasts under model and parameter uncertainty," International Review of Financial Analysis, vol. 48, pp. 282-291, 2016.

[4] YE. J, ZHAO. K, WANG. C, & LIU. W, "Analysis and Forecasts of Gold Price Based on the ARFIMA-GARCH Model," Journal of Qingdao University(Natural Science Edition), 4, p. 10, 2014. Available from: 10.3969/j.issn.1006-1037.2014.11.03 (Accessed: 15 Apr 2018).

[5] Mombeini, H. and A. Yazdani-Chamzini, "Modeling Gold Price via Artificial Neural Network Journal of Economics," Business and Management, vol. 3, no. 7, pp. 699-703, 2015.

[6] Kristjanpoller, W, & Minutolo, M, "Gold price volatility: A forecasting approach using the Artificial Neural Network–GARCH model," Expert Systems With Applications, vol. 42, pp. 7245-7251, 2015.

[7] Sharma, RK, "Forecasting Gold price with Box Jenkins Autoregressive Integrated Moving Average Method," Journal of International Economics, vol .7, no. 1, Jan-Jun, pp. 32-61, 2016. The World Gold Counci.

[8] Box, G.E.P., Jenkins, G.M., "Time Series Analysis: Forecasting and Control," San Francisco: Holden-Day, 1976.