APPLICATION OF ARIMA MODEL TO FORECAST GOLD PRICE IN VIETNAM

Thanh Tri Ho – Dao Phan –Van Ninh Nguyen – Juraj Sipko

Abstract

The paper is dealing with the analysis of the gold price in Vietnam. In line with the outbreak of the global financial crisis, the prices of gold have reached very high level. The gold started to be very attractive and safe investment. Therefore, investors are always looking forward to the research analysis related to the future price of gold.

One of the possible methods for the prediction of gold price is ARIMA model historically developed by George Box and Gwilym Jenkins 1976. The paper will use this model for the forecast of gold price in Vietnam. Based on the comprehensive analysis the paper come to the conclusion that how investors could avoid the risks related the price volatility in the international market with gold and as well how this development will influence the prices of gold in Vietnam in the forward market. The forecast Arima (5,1,5) model is excellent in this case with the forecast error is about 3.46%.

Key words: gold price, ARIMA model, international markets for gold, demand, and supply for gold

JEL Code: B23, B26

Introduction

From thousands of year, the gold is viewed as currency to bring global value, while the gold is one of the many investment channels interested. The investors look to gold as a safe investment tool and high profitability in the financial portfolio (Dung 2004). However not the investors would be also lucky and succeed in this wallet. Investing and trading gold is a risky area because the gold prices have fluctuated continuously in recent year. Also, for this reason, the investors are more interested in the role of forecasting gold price in the future. This is a support channel for investors if they know how to use and perform correctly. Therefore, the estimate of gold price in Vietnam is necessary for investment decision making in the present period (Hong 2003).

In this topic, the research team proposed to use the Arima model of Box – Jenkins to forecast the gold's price in the short term. George Box and Gwilym Jenkins (1976) studied the Arima model (Autoregressive integrated moving average), and their names are often used to call generic Arima processes, applied analysis and expected the time series. The Box – Jenkins's method have four steps: identify the test model, estimates, diagnostic tests, and forecast.

1. Theory

1.1 Box – Jenkins model

Arima model is a model used to forecast time series; this model relies on past data to provide forecasting indexes. The Arima model has two parts: AR (Autoregressive) and MA (Moving Average). Arima (p,d,q) with p is the number of autoregressive terms, d is the number of nonseasonal differences, q is the number of moving – average time (Wang, Wang et al. 2014). The general equation of the Arima (p,d,q) is:

$$Y_t = \delta + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \ldots + \varphi_p Y_{t-p} - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q} + e_t$$

With φ is the autoregressive parameter, θ moving average parameter, $\delta = \mu(1 - \varphi_1 - \varphi_2 - \cdots - \varphi_p)$, μ is the average value of time series, e_t is forecast error, $e_t = Y_t - \widehat{Y}_t$. The autoregressive model is Arima (p,d,0) or Ar (p). The general equation of AR(p) is: $Y_t = \delta + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \ldots + \varphi_p Y_{t-p} + e_t$. The moving average model is Arima (0,d,q) or MA(q). The general equation of MA(q) is: $Y_t = \delta - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q} + e_t$. Define the Arima model also means find the times that are related to each other. The $\widehat{Y}_t$ value is related to the $Y_{t-1}, Y_{t-2}, \ldots$ or with Y_{t-p}? The $\widehat{Y}_t$ value is related to the $e_{t-1}, e_{t-2} \ldots$ or e_{t-q}? P,q by how much?

1.2 The basic steps of the Box – Jenkins method

Consider the nature of the time series: According (Wang, Wang et al. 2014), the first thing to note is most of the time series are non-stop, and AR, MA components of the Arima model only related to the stop time series. The random process of Y_t is considered stop if the mean and the variance of the process do not change over time. So to identify the Arima model, we have to do two steps: based on the graph of time series, a graph of the PACF autocorrelation (Ali, Babai et al. 2017).

The model identification: The type identification of Arima (p,d,q) is to find the appropriate values of p,d,q. Where p is the number of autoregressive terms, d is the number of

nonseasonal differences, q is the number of moving – average times (Hikichi, Salgado et al. 2017). The determination of p and q will depend on the PACF and ACF graphs. Where ACF is the sample autocorrelation, PACF is the sample partial autocorrelation. Select the value of p if the PACF graph has a high value at lag 1,2,….p and plummeting after p. Choose the value of q if the ACF has a height value at lag 1,2,…,q and falling after q (Wang, Deng et al. 2017). **Parameter estimation of the model:** After finding the (AR-p) regression coefficient and the (MA-q) moving average ratio of the Arima model, we will estimate how many examples from the hight rates coefficients. And then test the model just found (Wang, Wang et al. 2014).

2. Application of Arima model for forecasting gold prices in Vietnam

2.1 The situation of gold price fluctuation in Vietnam

Period 2006 – 2009: According (Tran, Le et al. 2017), although the State does not own license business, the law's provision about the gold market management was still incomplete, so the form of illegal gold trading was created and developed rapidly. In this time, the gold market had about 40-60 tons of gold was imported officially and about 50-60 tons of smuggled gold. Characteristic of this stage was the big difference between domestic gold and world gold price. At this time, the local gold market was unstable, the severe effect on foreign exchange market and exchange rate.

Period 2009 – 2012: At this time, the difference between domestic gold price and world gold price had tended to decrease over 2006 – 2009; instability of the market and its negative impact on the exchange rate, price index, macroeconomic stability still happening but lower than the previous period (Tran, Le et al. 2017).

Period 2012-2015: This is the stage had a new legal framework and the effective enforcement. During this time, the gold exchange and the gold trading on account were not allowed to operate. The state Bank of Vietnam had exclusive rights to import gold and manufacture of gold bars. Characteristics of this stage: The difference between the domestic gold price and the world gold price had a low level. The gold market was stable; the illegal gold trading had stopped working (Tran, Le et al. 2017).

2.2 The application of Arima model for forecasting gold price in Vietnam

Description of data: The data used in the study were taken monthly from 01/2006 to 01/2017. A total of 133 observers corresponding to 133 months. The data each month is average data of all dates. Research data is taken from the website of World gold council.

Explain the variables used in the model:

The original data: Gold

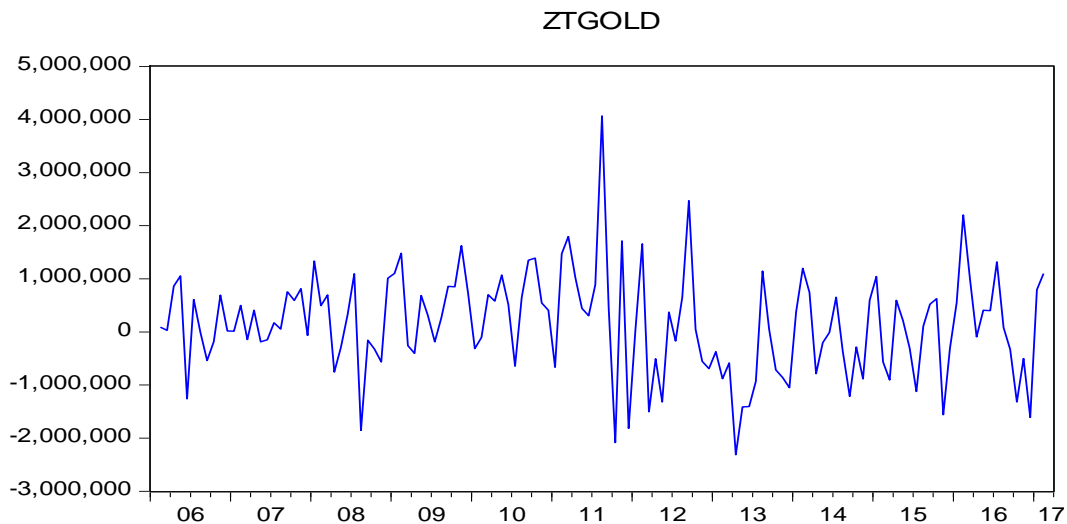The gold data after getting the difference in level 1: Z_tGold

We had a graph of the Vietnam gold price in the initial data string (Figure 2.1) and the gold data after getting the difference in level 1. (Figure 2.2).

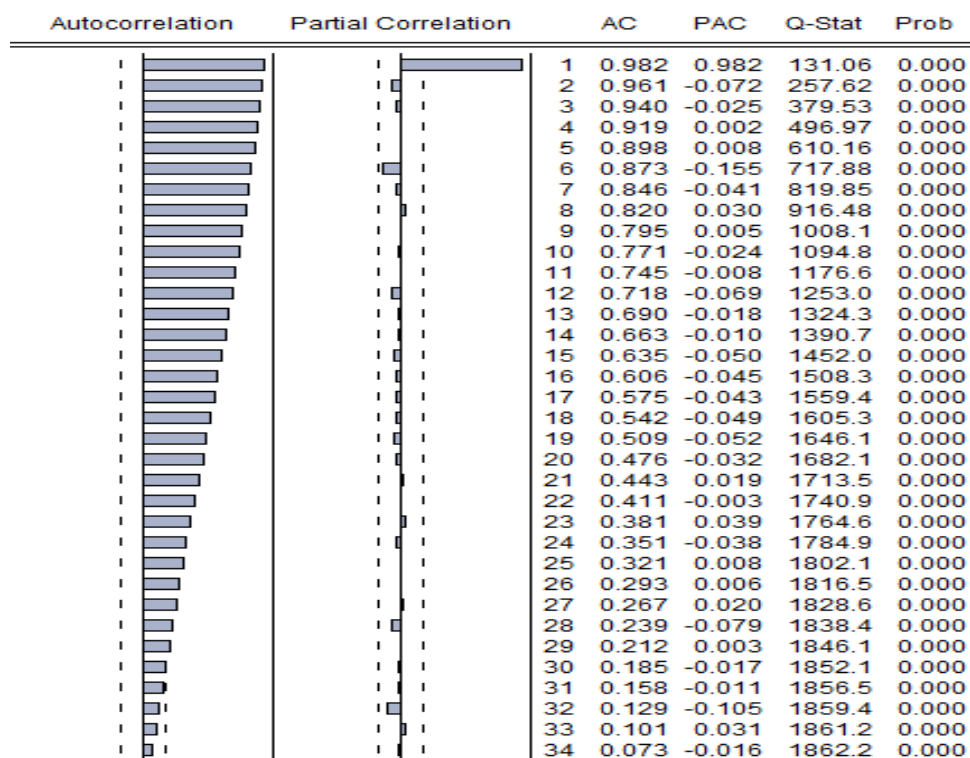**Figure 2.1: The original data string of gold**
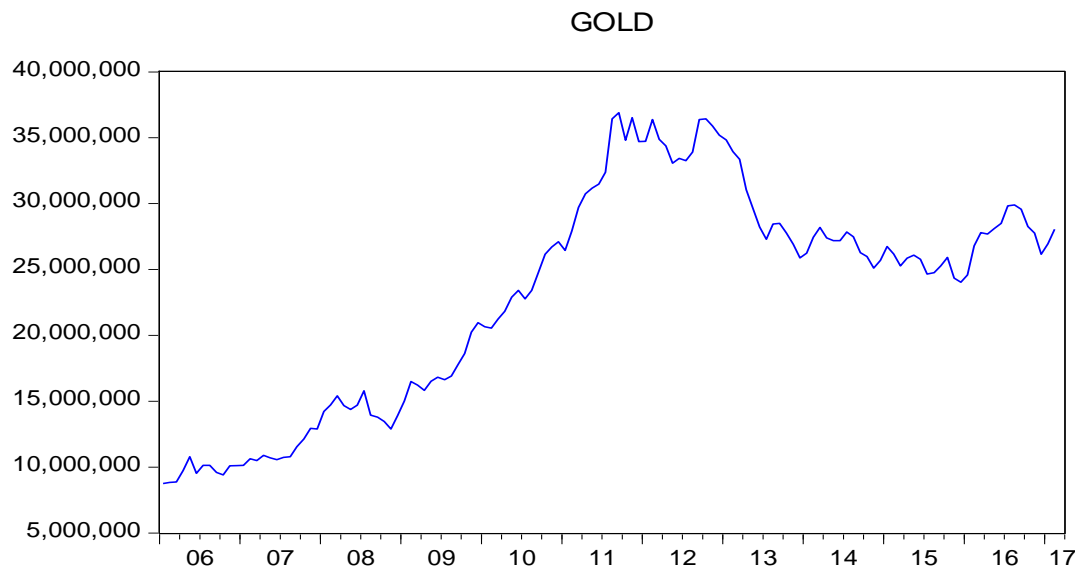
Source: Own calculation

Figure 2.2: The gold data after getting the difference in level 1

ZTGOLD

Source: Own calculation

From figure 2.2, we can see that observation data has stopped, and this is also clearly shown in figure 2.3 and figure 2.4.

Figure 2.3: ACF & PACF of the original data string

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.982	0.982	131.06	0.000
		2	0.961	-0.072	257.62	0.000
		3	0.940	-0.025	379.53	0.000
		4	0.919	0.002	496.97	0.000
		5	0.898	0.008	610.16	0.000
		6	0.873	-0.155	717.88	0.000
		7	0.846	-0.041	819.85	0.000
		8	0.820	0.030	916.48	0.000
		9	0.795	0.005	1008.1	0.000
		10	0.771	-0.024	1094.8	0.000
		11	0.745	-0.008	1176.6	0.000
		12	0.718	-0.069	1253.0	0.000
		13	0.690	-0.018	1324.3	0.000
		14	0.663	-0.010	1390.7	0.000
		15	0.635	-0.050	1452.0	0.000
		16	0.606	-0.045	1508.3	0.000
		17	0.575	-0.043	1559.4	0.000
		18	0.542	-0.049	1605.3	0.000
		19	0.509	-0.052	1646.1	0.000
		20	0.476	-0.032	1682.1	0.000
		21	0.443	0.019	1713.5	0.000
		22	0.411	-0.003	1740.9	0.000
		23	0.381	0.039	1764.6	0.000
		24	0.351	-0.038	1784.9	0.000
		25	0.321	0.008	1802.1	0.000
		26	0.293	0.006	1816.5	0.000
		27	0.267	0.020	1828.6	0.000
		28	0.239	-0.079	1838.4	0.000
		29	0.212	0.003	1846.1	0.000
		30	0.185	-0.017	1852.1	0.000
		31	0.158	-0.011	1856.5	0.000
		32	0.129	-0.105	1859.4	0.000
		33	0.101	0.031	1861.2	0.000
		34	0.073	-0.016	1862.2	0.000

Source: Own calculation

Figure 2.4: ACF & PACF of the gold data after getting the difference in level 1

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.192	0.192	4.9605	0.026
		2	-0.019	-0.057	5.0075	0.082
		3	0.034	0.051	5.1689	0.160
		4	0.005	-0.013	5.1727	0.270
		5	0.212	0.227	11.436	0.043
		6	0.171	0.088	15.526	0.017
		7	-0.076	-0.115	16.336	0.022
		8	-0.142	-0.127	19.221	0.014
		9	-0.037	0.001	19.420	0.022
		10	0.091	0.073	20.622	0.024
		11	0.199	0.143	26.405	0.006
		12	0.056	0.016	26.870	0.008
		13	0.024	0.096	26.957	0.013
		14	0.040	0.044	27.194	0.018
		15	0.058	0.004	27.704	0.023
		16	0.104	-0.019	29.364	0.022
		17	0.082	0.018	30.397	0.024
		18	0.071	0.088	31.187	0.027
		19	-0.008	-0.012	31.196	0.038
		20	-0.029	-0.033	31.329	0.051
		21	0.010	-0.011	31.346	0.068
		22	-0.056	-0.101	31.847	0.080
		23	-0.017	-0.027	31.895	0.102
		24	0.043	0.032	32.201	0.122
		25	-0.023	-0.003	32.286	0.150
		26	-0.045	-0.031	32.630	0.173
		27	-0.030	-0.030	32.778	0.204
		28	-0.092	-0.112	34.215	0.194
		29	-0.021	-0.055	34.291	0.229
		30	0.054	0.031	34.795	0.250
		31	-0.049	-0.043	35.208	0.276
		32	-0.120	-0.090	37.770	0.222
		33	-0.093	-0.021	39.329	0.207
		34	-0.095	-0.072	40.966	0.191

Source: Own calculation

We used the autocorrelation and partial autocorrelation graph of Z_tGold data string to get the Arima model. Look at the autocorrelation and partial autocorrelation graph of figure 2.4; we can see that p can be values 1 and 5; q can be values 1 and 5. We can estimate four models: Arima (1,1,1); Arima (1,1,5); Arima (5,1,1); Arima (5,1,5). We continue to test the residual of four models to see if it has white noise or not by test the p-value of the residual. All four models have no white noise. Next step, we compare the RMSE (Root mean squared error), MAE (Mean absolute error) of each model.
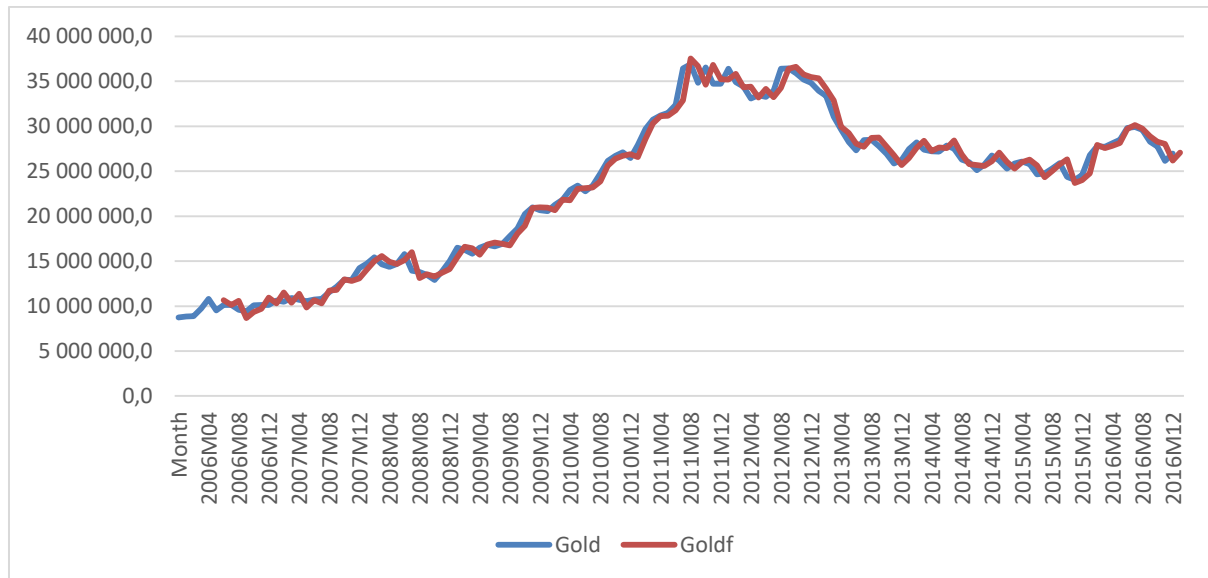
Table 2.1: RMSE & MAE of four models

Mô hình Arima	RMSE	MAE
Arima (1,1,1)	928932.1	716887.0
Arima (1,1,5)	918322.3	712972.9
Arima (5,1,1)	915193.6	706869.6
Arima (5,1,5)	874510.6	671072.6

Source: Own research

Look at table 2.1 we can see the Arima (5,1,5) model has the smallest RMSE, MAE. So we choose the Arima (5,1,5) model to forecast.

Chart 2.1: The current gold price and forecast gold price

Source: Own research

Look at the chart we can see the current gold price and forecast gold price always had the same trend. This proves that the forecast Arima (5,1,5) model is excellent.

2.3 Expansion of the gold price forecast

The equation of prediction: $Z_t Goldf = 130740.9 + 0.741463.Z_t Gold_{t-1} - 0.678517.Z_t Gold_{t-2} - 0.518136.Z_t Gold_{t-3} + 0.416901.Z_t Gold_{t-4} - 0.530531.Z_t Gold_{t-5} - (-0.523059).e_{t-1} - 0.518119.e_{t-2} - 0.755595.e_{t-3} - (-0.450518).e_{t-4} - 0.760356.e_{t-5}$

The forecast gold price = (The Gold price)$_{t01-2017}$ + $Z_t Goldf$ = 26948229.9 + 127824.8 = 27076054.7

Compared to the actual gold price in February of 2017 is 28047662.7 VND and the forecast gold price in February of 2017 is 27076054.7 VND. The difference between the actual price and the forecast price is 971608 VND. The forecast error is about 3.46%.

Conclusion

The forecast result shows the forecast value is not much different from the actual value. This means the reliability of the model is very good. In fact, the gold price is also affected by

another factor such as the world gold price, oil price, exchange rate of VND/USD will make the forecast error higher. So the combination of the prediction results and the current economic situation to have a correct view about the gold price fluctuations.

However, it is necessary to reaffirm that current Arima model is very good in applying it to the gold price forecast. It helps the investors make decisions and investment plan avoid the risk of gold trading.

References

Box-Jenkins Model. 2013.

Vietnam: Gold Trade Shifts Direction: Local Gold Price Dropped by 2.7 Percent in the First Six Months of 1999. 1999. *Vietnam Economic News, ,* pp. 1.

ALI, M.M., BABAI, M.Z., BOYLAN, J.E. and SYNTETOS, A.A., 2017. Supply chain forecasting when information is not shared. *European Journal of Operational Research,* **260**(3), pp. 984-994.

DRITSAKI, C., 2015. BoxJenkins Modeling of Greek Stock Prices Data. *International Journal of Economics and Financial Issues,* **5**(3), pp. 740-747.

DUNG, T., 2004. Global Strife Pushes Local Gold Price to New Heights. *Vietnam Investment Review,* (650), pp. 15.

GANGOPADHYAY, K., JANGIR, A. and SENSARMA, R., 2016. Forecasting the price of gold: An error correction approach. *Iimb Management Review,* **28**(1), pp. 6-12.

HIKICHI, S.E., SALGADO, E.G. and BEIJO, L.A., 2017. Forecasting number of ISO 14001 certifications in the Americas using ARIMA models. *Journal of Cleaner Production,* **147**, pp. 242-253.

HONG, N., 2003. Wedding Boom Raises Gold Price. *Vietnam Investment Review,* (630), pp. 15.

TAM, W. and REINSEL, G.C., 1997. Tests for Seasonal Moving Average Unit Root in ARIMA Models. *Journal of the American Statistical Association,* **92**(438), pp. 725-738.

TRAN, T.N., LE, C.D. and HOANG, T.T.P., 2017. Does the State Bank Widen the Gap Between International and Domestic Gold Prices? Evidence from Vietnam. *Global Business Review,* **18**(1), pp. 45-56.

WANG, H.R., WANG, C., LIN, X. and KANG, J., 2014. An Improved ARIMA Model for Hydrological Simulations. *Nonlinear Processes in Geophysics Discussions,* **1**(1), pp. 841-876.

WANG, K.W., DENG, C., LI, J.P., ZHANG, Y.Y., LI, X.Y. and WU, M.C., 2017. Hybrid methodology for tuberculosis incidence time-series forecasting based on ARIMA and a NAR neural network. *Epidemiology and infection,* **145**(6), pp. 1118-1129.

Contact

Ho Thanh Tri

Tomas Bata University

T.G. Masaryka 5555 760 01 Zlín, Czech Republic

tho@fame.utb.cz

Phan Đao

Ton Duc Thang University

19 Nguyen Huu Tho street, Dist. 7, HCMC, Viet Nam

phandao@tdt.edu.vn

Nguyen Van Ninh

Tomas Bata University

T.G. Masaryka 5555 760 01 Zlín, Czech Republic

n1nguyen@fame.utb.cz

Juraj Sipko

Institute of Economic Research SAS

Sancova 56 Bratislava 811 05

jurajsipko@gmail.com