Forecasting Futures Gold Prices Using Pulse Function Intervention Analysis Approach

Sediono[1*], Putu Eka Andriani[2], Ariadna Sopia Miranda[3], and Idrus Syahzaqi [4]
[1,2,3,4]Department of Mathematics, Universitas Airlangga, Surabaya, Indonesia
*Corresponding author: sediono101@gmail.com

Received: 6 November 2024
Revised: 24 January 2025
Accepted: 30 January 2025

ABSTRACT – **Gold is a precious metal that plays an important role in global trade and is often use as a financial standard in various countries. In 2024, gold prices surged sharply due to global macroeconomic factors, such as economic uncertainty, positioning gold as a safe haven for investors. Accurate predictions of future gold prices are crucial for helping investors make informed decisions and adapt to market changes. In line with Sustainable Development Goal (SDG) 8 on Decent Work and Economic Growth, this study uses the pulse function intervention analysis approach to predict gold prices by identifying patterns of changes in the pre-intervention and post-intervention periods. This study aims to make a significant contribution to the use of comprehensive and relevant predictive tools by considering the effects of interventions, supporting investor decision-making, and contributing to economic growth. The best model was obtained at ARIMA (0,2,1) with intervention parameters $b = 0$, $r = 2$, and $s = 0$. The prediction results show a close alignment with actual data, yielding a MAPE value of 1.289%. Additionally, this model produces the smallest AIC value of 1125.1, an SBC value of 1135.86, and an MSE value of 1403.11, demonstrating excellent predictive capability.**

Keywords– **Futures Gold Price, ARIMA, Pulse Function, Intervention Analysis**

I. INTRODUCTION

Gold is a strategic commodity in the global market with an important role in the world economy. Gold has various benefits, including serving as foreign exchange reserves, a primary payment instrument in several countries, a monetary or economic standard, and a form of investment [1]. Gold also functions as a safe haven, with its demand influenced by extreme loss risks, and this characteristic is reflected in its price fluctuations during times of crisis [2]. In addition to functioning as a safe haven, gold also serves as a hedge, which is a strategy employed by investors to reduce or eliminate potential risks [3]. Gold has proven to be a good asset to protect against inflation and gold returns increase during times of economic recession [4]. Demonstration of this finding can be seen during the COVID-19 pandemic outbreak when investors flocked to buy gold to protect themselves from stock losses [5]. This relationship applies in the US as well as gold prices measured in national currencies.

In 2024, gold prices experienced a significant spike, reaching a nominal record high of $2,331 per troy ounce in April 2024, following a 7 percent increase in the first quarter [6]. This surge was driven by strong demand from central banks in emerging markets and influenced by various global macroeconomic factors, such as global economic uncertainty, geopolitical issues (the Russia-Ukraine crisis, the US-China trade war, and conflicts in the Middle East), and high inflation.

The Russia-Ukraine conflict has caused significant global impacts, including disruptions in the supply and prices of key commodities like oil and gas, as well as creating financial market dynamics that are seen as the largest changes since the 2008 financial crisis [7]. During the ongoing trade conflict between the US-China, gold prices have experienced significant fluctuations due to increasing global economic uncertainty and unstable growth prospects, prompting investors to turn to gold as a safe haven asset, which is reflected in the rising prices of gold in international markets [8]. In this situation, gold is considered a safer asset compared to other financial instruments. When investors lose confidence in the stock market or currencies, they turn to gold as an asset that can maintain its value in the long term [9].

In 2024, many countries made massive investments in gold, both through their central banks and private investors, reflecting increasing global demand. The surge in gold prices in the commodities market, such as the London Bullion Market and NYMEX, was largely driven by high demand from major central banks like China, Russia, and Turkey, which significantly increased their gold reserves as a strategy for asset diversification. Data from the World Gold Council indicates that in the third quarter of 2024, central bank purchases of gold reached 337 tons, with total purchases during the first nine months hitting 800 tons, the highest amount in history. Additionally, reports from research firms like Bloomberg and Reuters confirm this trend, highlighting that these countries are investing in gold as protection against existing risks. Thus, 2024 is expected to be a favorable year for the gold market, driven by strong demand from various sectors [10].

With increasing demand and fluctuating gold prices, the ability to predict futures gold prices becomes very important. Futures gold offers significant profit opportunities for investors, but high volatility risks are also inherent. In this context, methods capable of accurately predicting gold price movements can assist investors and policymakers in making more informed decisions. Predictions for futures gold prices align with the Sustainable Development Goals (SDGs), particularly SDG number 8, which is Decent Work and Economic Growth. The stability of futures gold prices can contribute to sustainable economic growth by helping investors and companies make better decisions and create jobs.

Previous research using pulse function intervention analysis has been conducted by [11] to model and predict global crude oil prices. The data used comes from weekly world crude oil prices during the period from June 8, 2020, to September 19, 2022. The ARIMA (3,2,0) model was chosen as the baseline model before the intervention, while the intervention point was set at the time of conflict occurrence. This model utilizes ARIMA parameters and a pulse intervention function with $b = 0$, $s = 1$, and $r = 2$ to evaluate the direct impact of the intervention event. The prediction results showed a high accuracy value with a MAPE of 2.8982% and an MSE of 10.2687. Another study was conducted by [12] who predicted gold prices using the ARIMA method. The best model produced is ARIMA (1,1,0) with an MAE error value of 94491.42, a MAPE of 1047776409, an MSE of 12281027973, and an RMSE of 110819.7996. However, this research did not use a method that considers intervention variables.

Previous research by [11] used a similar method to model crude oil prices and demonstrated high accuracy in predicting the impact of external events. Meanwhile, employed the ARIMA method to predict gold prices without considering intervention variables, resulting in lower accuracy due to the lack of accounting for the impacts of external events. Based on this description, this study will apply pulse function intervention analysis in forecasting futures gold prices, which has not been implemented in previous research, allowing for a more accurate identification of the impact of events on gold prices. The pulse function intervention analysis method is chosen for its ability to identify and measure the impact of external events on time series data, such as geopolitical events or economic policies, which cannot be captured by classical models like ARIMA. The advantage of this method is its capability to evaluate the direct influence of interventions at specific points in time, resulting in more accurate predictions that are responsive to market changes. Therefore, this method supports more informative decision-making for investors and policymakers. Based on the description, this study uses the intervention modeling that aims to produce accurate world gold price predictions. This study is expected to make a significant contribution to the use of comprehensive and relevant predictive tools by considering the effects of interventions.

II. LITERATURE REVIEW

A. Gold Price

Gold is a precious metal that plays an important role in global trade and is often used as a financial standard in various countries. In addition to its function in trade, gold is also sought after for its stability and relatively constant value, making it a choice for protecting assets against inflation or economic uncertainty. Gold is considered the most liquid and safest investment alternative against stocks in the market [13]. Gold is also used in certain financial regulations, such as for Hajj savings, where value stability is crucial. Furthermore, gold is often chosen as a form of long-term investment because its resale value tends to increase over time [14].

According to Abdullah [15] several factors influence gold prices, including rising inflation that exceeds predictions, which can drive investors to seek safe assets like gold. Financial turmoil, including the monetary crises of 1998 and 2008, can also lead to spikes in gold prices as a response to economic uncertainty. Significant increases in oil prices contribute to rising gold prices, as does the continuously increasing global demand, which will push gold prices up if supply does not keep pace. Additionally, unstable global political conditions can create economic uncertainty, prompting investors to turn to gold. Finally, exchange rate fluctuations, particularly the weakening of the US dollar, can affect the rise in gold prices in international markets.

B. Autoregressive Integrated Moving Average (ARIMA)

The ARIMA (p, d, q) model is a forecasting approach that incorporates differencing to achieve data stationarity. It is a linear model suitable for addressing stochastic series, typically derived from the autoregressive model AR(p), the moving average model MA(q), and their combination known as the ARMA(p, q) model [16]. According to Wei [17], the ARIMA model of order (p, d, q) can be mathematically represented by a specific equation:

$$\phi_p(B)(1-B)^d Z_t = \theta_q(B)a_t \tag{1}$$

Mathematically, the general ARIMA (p, d, q) model with $d = 1$ can be formulated as follows:

$$Z_t = (1+\phi_p)Z_{t-1} + (\phi_2+\phi_1)Z_{t-2} + \cdots + (\phi_p+\phi_{p-1})Z_{t-p} - \phi_p Z_{t-p-1} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}; \quad a_t \sim N(0, \sigma_a^2) \tag{2}$$

In addition, when the data is stationary in terms of mean and variance, then the ARIMA model can be identified using the autocorrelattion plot (ACF) and partial autocorrelation plot (PACF) which are continued by estimating the model using Maximum Likelihood Estimation (MLE). After the parameters are significant, a diagnostic check is carried out using residuals to test whether the white noise assumptions and the normal distribution residuals are met [11].

C. Intervention Analysis

Time series intervention is a statistical technique employed to analyze the effects of an intervention on a variable in time series data. The intervention can be either a deliberate change or an unexpected event that affects the time series pattern. The intervention function aims to determine a model that predicts the effect of an intervention on a time series variable, utilizing the ARIMA model before the intervention and adding an indicator variable as the intervention variable, which takes on values of 1 or 0. Generally, there are two categories of interventions: step and pulse. The step function describes a sudden change in level for a time series variable at a specific time and continues thereafter. Mathematically, it can be expressed as follows [17].

$$I_t^{(T)} = S_t^{(T)} = \begin{cases} 1, & t < T \\ 0, & t \geq T \end{cases} \tag{3}$$

The pulse function is a function that describes a sudden change in the level of a time series variable only at that specific moment, without continuing into subsequent time periods. Mathematically, this can be expressed as follows:

$$I_t^{(T)} = P_t^{(T)} = \begin{cases} 1, & t = T \\ 0, & t \neq T \end{cases} \tag{4}$$

Mathematically, the general form of the intervention model can be written in the following equation:

$$X_t = \frac{\omega_s(B)B^b}{\delta_r(B)} I_t^{(T)} + N_t \tag{5}$$

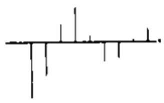
where
X_t : Response variable at time t
I_t : Intervention variable
b : Lag period where intervention I starts affecting X
s : The length of an intervention influences the data following a span of b periods
r : The intervention effect pattern after $b + s$ periods from the intervention at time T
ω_s : $\omega_0 - \omega_1 B - \cdots - \omega_s B^s \delta_r$
δ_s : $1 - \delta_q B - \cdots - \delta_r B^r$
N_t : The best ARIMA model without the influence of intervention

Mathematically, the ARIMA model without the influence of intervention can be expressed in the following equation:

$$N_t = \frac{\theta_q(B)}{\phi_p(B)(1-B)^d} \alpha_t \tag{6}$$

The Cross-Correlation function is utilized to identify the order of the intervention model, which consists of parameters b, s, and r. These values serve as references in determining the transfer function. The transfer function embodies the connection between the input and output within a system. The main purpose of the transfer function is to predict the system's output based on the given input. The illustration and typical impulse response function of order b,r,s from cross-correlation plot is displayed in **Table 1** [18]

Table 1 Illustration and Impulse Response Function of order b, r, s

b, s, r	Impulse Response Function
(2, 2, 0)	

D. Criteria For the Best Model

1. Mean Squared Error (MSE)
   MSE quantifies the mean of the squared discrepancies between the predicted values and the actual values of the dependent variable. It can be systematically written in the following equation [19]:

$$MSE = \frac{\sum_{t=1}^{n}(Z_t - \hat{Z}_t)^2}{n} \tag{7}$$

where
Z_t : Actual value at time t

$\hat{Z}_t$: Predicted value at time t
n : Total data

2. Schwarz Bayesian Criterion (SBC)
   SBC is a criterion based on the Bayesian method introduced by Schwartz. It can be systematically written in the following equation [19]:

$$SBC(M) = n \ln \hat{\sigma}_a^2 + M \ln n \tag{8}$$

   where
   $\hat{\sigma}_a^2$: Estimation of residual variance
   M : Number of parameters in the model
   n : Total residual

3. Akaike Information Croterion (AIC)
   AIC is another criterion used for selecting the optimal model. It can be systematically written in the following equation [19]:

$$AIC = n \ln \hat{\sigma}_a^2 + 2M \tag{9}$$

   where
   $\hat{\sigma}_a^2$: Estimation of residual variance
   M : Number of parameters in the model
   n : Total residual

4. Mean Absolute Percentage Error (MAPE)
   Model validation using MAPE, assesses the accuracy of the model's predictions by calculating the average absolute percentage of the residuals. The MAPE equation is as follows:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{Z_t - \hat{Z}_t}{Z_t}\right| \times 100\% \tag{10}$$

   where
   Z_t : Actual value at time t
   $\hat{Z}_t$: Predicted value at time t
   n : Total data
   To describe the model's performance effectively based on Trimono [20], the MAPE value is interpreted in the **Table 2** below.

Table 2 MAPE Value Interpretation

MAPE Value	Interpretation
MAPE < 10%	The predictive model's ability is very good
10% < MAPE ≤ 20%	The predictive model's ability is good
20% < MAPE ≤ 50%	The predictive model's ability is adequate
MAPE > 50%	The predictive model's ability is poor

III. METHODOLOGY

A. Sources of Data and Research Variables

The data utilized in this study is gold futures price data with a weekly period from July 2022 to July 2024 obtained from the official investing.com website [21]. The dataset is divided into two parts: one for training and the other for testing. The training data is utilized to develop the model, consisting of pre-intervention and post-intervention testing data. The training data includes data from the second week of July 2022 to the second week of June 2024, totaling 104 data points. Meanwhile, the testing data is used to compare the predicted results with actual data, specifically from the third week of June 2024 to the last week of July 2024, totaling 7 data points.

The variables in this study are found in **Table 3**.

Table 3 Research Variables

Variable	Variable Name	Unit
y_t	Futures Gold Prices	USD/t.oz
x	Time period	Weekly period

B. Step of Analysis

The analysis steps used in the research are as follows:

1. Modeling the futures gold price data with ARIMA on pre-intervention data and an intervention analysis approach.
   a. Data Exploration
      i. Creating time series plots and estimating extreme variables that may lead to intervention
      ii. Splitting the dataset into training and testing subsets
      iii. Dividing the training data into two segments: the pre-intervention segment for data collected prior to the occurrence of extreme events and the post-intervention segment for data gathered after such events have taken place
   b. Estimating the Best ARIMA Model with Pre-Intervention Data
      i. Determining the stationarity of the data through the analysis of time series plots, ACF plots, PACF plots, and the ADF test
      ii. If the data exhibits non-stationarity in variance, it is transformed using the Box-Cox transformation. Subsequently, if the data obtained from the Box-Cox transformation remains non-stationary in mean, differencing is applied
      iii. Identifying the model by creating and comparing ACF and PACF plots of the pre-intervention data that have met the stationarity assumptions
      iv. Estimating parameters in the model using the least squares method or Ordinary Least Squares (OLS)
      v. Conducting diagnostic tests by examining models whose parameters have significance values $< \boldsymbol{\alpha}$, checking for residual white noise, and performing normality tests on the residual data
      vi. Choosing the optimal model based on minimum AIC, SBC and MSE
      vii. Generating forecasts for post-intervention data utilizing the best ARIMA model obtained from the pre-intervention dataset
   c. Modeling with an Intervention Approach
      i. Determining the parameters of the intervention model (b, s, r) by analyzing the Cross Correlation Function (CCF) plot that compares the post-intervention data with the predicted values from the optimal ARIMA model
      ii. Estimating the parameters of the intervention model using the Ordinary Least Squares (OLS) approach
      iii. Conducting diagnostic tests by examining models whose parameters have significance values $< \boldsymbol{\alpha}$, checking for residual white noise, and performing normality tests on the residual data
2. Predicting and Analyzing Futures Gold Prices with an Intervention Analysis Approach
   a. Writing down the best model obtained
   b. Checking the accuracy of the model in making predictions using testing data
   c. Predicting data based on the best model
3. Interpreting the Obtained Model

IV. RESULTS AND DISCUSSIONS

A. Descriptive Statistics

Descriptive analysis of gold prices is conducted to describe data as the research variable in order. The gold price data is divided into training data for pre-intervention and post-intervention periods.

Table 4 Descriptive Analysis of Futures Gold Prices

Variables	Mean	Variance	Minimum	Median	Maximum
Full Data	2002.9	211.4	1655.6	1989.1	2469.8
Pre Intervensi	1927.5	133.2	1655.6	1948.9	2206.3
Post Intervensi	2392.3	39.5	2330.6	2384.1	2469.8

Based on **Table 4**, it shows that the futures gold prices before the intervention, from the second week of July 2022 to the third week of March 2024, had an average of 1927.5 USD/t. oz and a variance of 133.2 USD/t. oz. Meanwhile, after the intervention, the average futures gold price was 2392.3 USD/t. oz with a variance of 39.5 USD/t. oz. The post-intervention data started from the fourth week of March 2024 to the fourth week of July 2024.

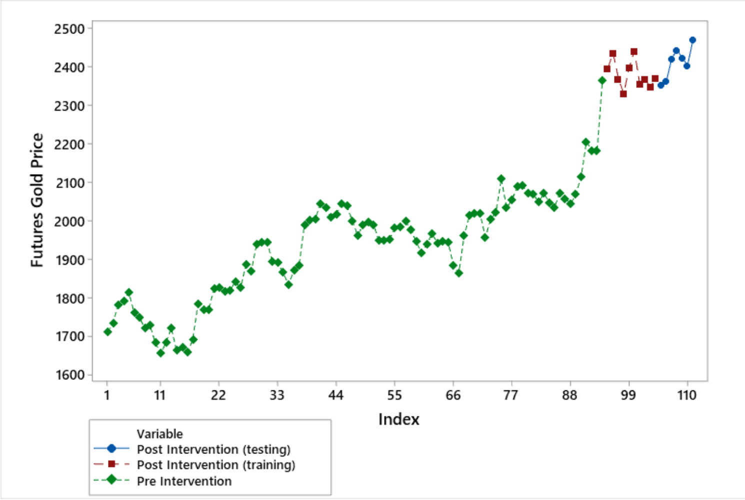


Figure 1 Plot Time Series of Futures Gold Prices

Based on **Figure 1**, the plot of futures gold prices shows that the data fluctuates and experiences a sharp increase at point $t = 94$ on March 24, 2024, with a price of 2366.60 USD/t. oz. The surge in gold prices in 2024 was driven by a significant increase in demand from central banks of various countries and institutional investors as a safe haven asset in the face of economic and geopolitical instability. As a result, the global gold price hit its peak for the first time, which was later recognized as the intervention point in the pulse function intervention analysis. This analysis examines the distribution of data both prior to and following the intervention, namely at $t = 1$ (July 10, 2022) to $t = 93$ (March 28, 2024) as pre-intervention data, while $t = 94$ (March 31, 2024) to $t = 111$ (July 28, 2024) as post-intervention data.

B. *ARIMA Modeling of Pre-Intervention Futures Gold Price Data*

Pre-intervention data were used to determine the best ARIMA model. To determine the model, the data first needed to be identified whether they had stationarity in mean (no trend) and stationarity in variance (no large fluctuations). Then, the stationarity of the transformed data in variance is checked using the box-cox transformation that shown in **Figure 2**.

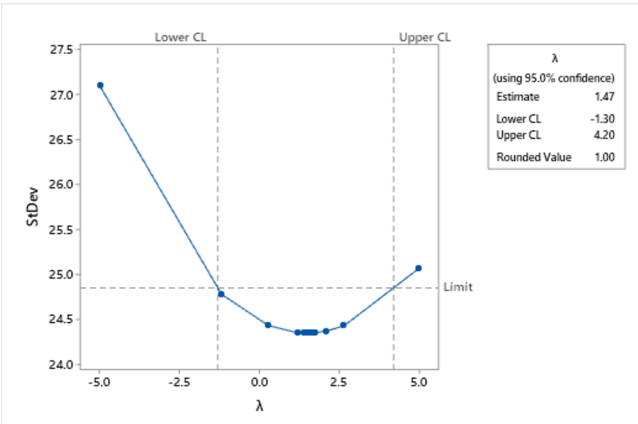
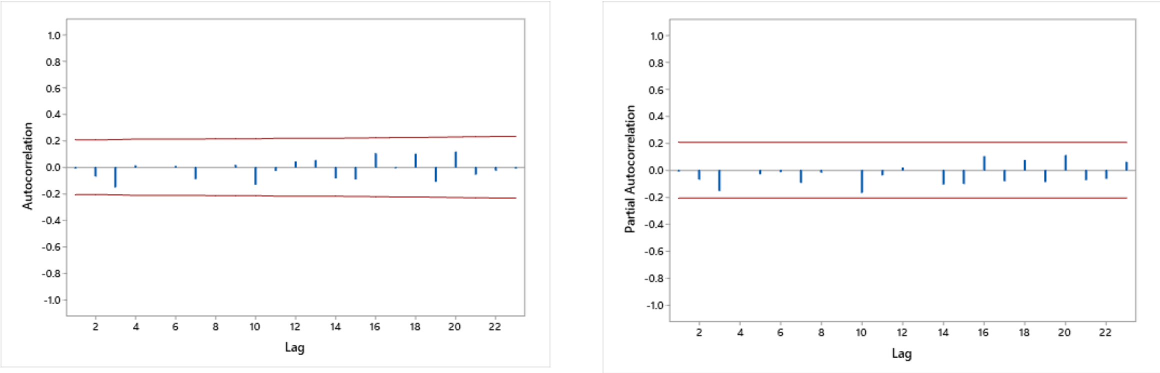


Figure 2 Box-cox Plot of Pre-Intervention Futures Gold Prices Data

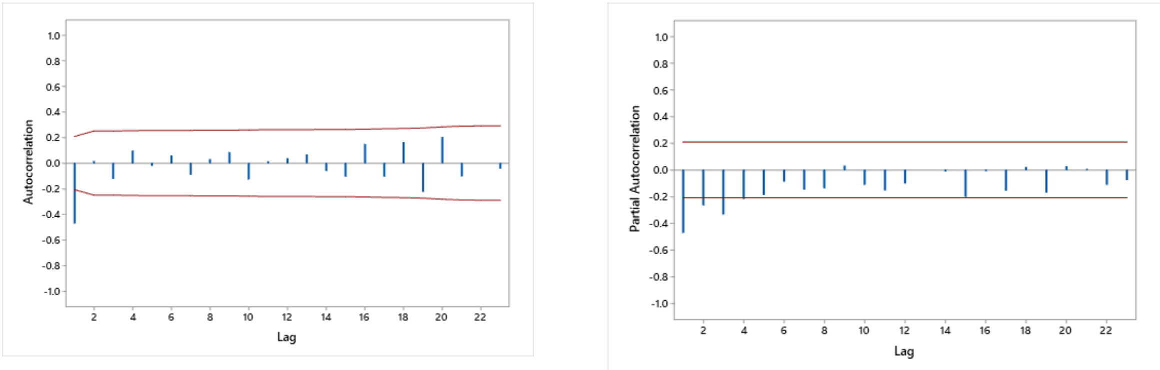
Based on **Figure 2**, The Box-Cox transformation shows a rounded value (λ) of 1.00 that indicating the data is stationary in variance. Then, the stationarity of the transformed data in terms of mean was checked using differencing at lag 1 by examining the trend analysis graph. The ACF and PACF plots following the first differencing are presented in **Figure 3**.



(a) (b)

Figure 3 (a) ACF Plot of Pre-Intervention Futures Gole Price Data with $d = 1$ (b) PACF Plot of Pre-Intervention Futures Gole Price Data with $d = 1$

The pattern in the ACF and PACF plots shows that no lags fall outside the boundary lines. As a result, the time series analysis model to be tested cannot yet be determined. Therefore, a second differencing should be applied to the data to create a dataset that better aligns with stationary properties and to ascertain the suitable time series analysis model.



(a) (b)

Figure 4 a) ACF Plot of Pre-Intervention Futures Gole Price Data with $d = 2$ (b) PACF Plot of Pre-Intervention Futures Gole Price Data with $d = 2$

Based on **Figure 4**, there is a significant 1st lag in the ACF plot and significant at 1st, 2nd and 3rd lags in PACF plot. Based on the ACF and PACF plots, a preliminary model of MA(1) with order q equal to 1 and AR(3) with order p equal to 1, 2, or 3 is suggested. The stationarity test indicates that the data is stationary after applying differencing twice, thus the order d is 2. The preliminary model estimates based on the analysis above are ARIMA (0,2,1), ARIMA (1,2,0), ARIMA (1,2,1), ARIMA (2,2,0), ARIMA (2,2,1), ARIMA (3,2,0), and ARIMA (3,2,1). To confirm that the data is stationary in mean and variance, an ADF test was conducted, as presented in the **Table 5**.

Table 5 Results of the Augmented Dickey-Fuller Test Pre-Intervention Futures Gole Price Data with $d = 2$

t-statistics	P-value
-3.7378	0.025

In **Table 5**, it is known that the $p - value = 0.025$ is less than the alpha of 5%. As a result, it can be inferred that the data is stationary with respect to both mean and variance. Next, a preliminary model estimation will be conducted through identification on the ACF and PACF plots as shown in the figure.

Following the determination of the tentative models, then perform parameter estimation and significance testing of parameters from those models, as presented in the following tables.

Table 6 Pre-Intervention ARIMA Model

Model		Parameter	Estimate	P-Value	Description
ARIMA (0,2,1)	Probabilistic	MA(1)	0.9769	0.000	Significant
	Deterministic	MA(1)	0.9759	0.000	Not Significant
		Constant	0.029	0.928	
ARIMA (1,2,0)	Probabilistic	AR(1)	-0.4754	0.000	Significant

Model		Parameter	Estimate	P-Value	Description
	Deterministic	AR(1)	-0.4756	0.000	Not Significant
		Constant	-0.064	0.894	
ARIMA (1,2,1)	Probabilistic	AR(1)	-0.014	0.896	Not Significant
		MA(1)	0.9784	0.000	
	Deterministic	AR(1)	-0.013	0.905	Not Significant
		MA(1)	0.9767	0.000	
		Constant	0.010	0.975	
ARIMA (2,2,0)	Probabilistic	AR(1)	-0.604	0.000	Significant
		AR(2)	-0.278	0.000	
	Deterministic	AR(1)	-0.604	0.000	Not Significant
		AR(2)	-0.278	0.010	
		Constant	-0.50	0.915	
ARIMA (2,2,1)	Probabilistic	AR(1)	-1.3681	0.000	Significant
		AR(2)	-0.3712	0.000	
		MA(1)	-0.9742	0.000	
	Deterministic	AR(1)	-1.3684	0.000	Not Significant
		AR(2)	-0.3715	0.000	
		MA(1)	-0.9740	0.000	
		Constant	-1.45	0.879	
ARIMA (3,2,0)	Probabilistic	AR(1)	-0.698	0.000	Significant
		AR(2)	-0.491	0.000	
		AR(3)	-0.358	0.001	
	Deterministic	AR(1)	-0.698	0.000	Not Significant
		AR(2)	-0.491	0.000	
		AR(3)	-0.357	0,001	
		Constant	-0.29	0.947	

Based on **Table 6**, it is known that ARIMA (0,2,1) probabilistic, ARIMA (1,2,0) probabilistic, ARIMA (2,2,0) probabilistic, ARIMA (2,2,1) probabilistic, and ARIMA (3,2,0) probabilistic have p-values less than 0.05, indicating that the parameters are considered significant.

Next, residual assumptions for the tentative models were conducted, including white noise to demonstrate the absence of correlation among residuals using the Ljung-Box test, and assessing the normality of residuals to confirm that the residual data follows a normal distribution through the Kolmogorov-Smirnov test.

Table 7 Result of White Noise Test and Normality of Residuals Test

Model	P-Value Ljng-Box				MSE	Normality P-value	Description
	Lag 12	Lag 24	Lag 36	Lag 48			
ARIMA (0,2,1) Probabilistic	0.851	0.866	0.966	0.950	1360.32	0.092	White Noise and Normally Distributed
ARIMA (1,2,0) Probabilistic	0.051	0.109	0.273	0.264	2085.21	>0.150	Not White Noise and Normally Distributed
ARIMA (2,2,0) Probabilistic	0.065	0.113	0.352	0.286	1956.28	>0.150	Not White Noise and Normally Distributed

Model	P-Value Ljng-Box				MSE	Normality P-value	Description
	Lag 12	Lag 24	Lag 36	Lag 48			
ARIMA (3,2,0) Probabilistic	0.323	0.590	0.911	0.954	1742.05	>0.150	White Noise and Normally Distributed
ARIMA (2,2,1) Probabilistic	0.016	0.052	0.178	0.130	2079.24	>0.150	Not White Noise and Normally Distributed

Based on **Table 7**, the ARIMA (0,2,1) probabilistic model is the proposed model that meets the assumptions of white noise and residual normality, and this model also has the smallest Mean Squared Error (MSE) value of 1360.32, mathematically expressed as follows:

$$\phi_p(B)(1 - B)^d Z_t = \theta_q(B)a_t \quad (11)$$

with a value of $p = 0,\ d = 2,\ q = 1$ then:

$$\begin{aligned}\phi_0(B)(1 - B)^2 Z_t &= \theta_1(B)a_t\\ (1 - B)^2 Z_t &= (1 - \theta_1 B)a_t\\ (1 - 2B - B^2)Z_t &= (1 - \theta_1 B)a_t\\ Z_t &= \frac{(1 - \theta_1 B)a_t}{(1 - 2B - B^2)}\\ Z_t &= (1 + (2 - \theta_1)B + (3 - 2\theta_1)B^2 + \cdots)a_t \quad (12)\end{aligned}$$

Substitute the value $\theta_1 = 0.9769$ obtained from the Table

$$\begin{aligned}Z_t &= (1 + (2 - 0.9769)B + (3 - 2(0.9769))B^2 + \cdots)a_t\\ Z_t &= (1 + 1.0231B + 1.0462\,B^2 + \cdots)a_t\\ Z_t &= a_t + 1.0231a_{t-1} + 1.0462\,a_{t-2} + \cdots \quad (13)\end{aligned}$$

The equation Z_t represents the best ARIMA model without the influence of intervention, used as noise or Z_t.

C. *Pulse Function Intervention Analysis of Futures Gold Price Data*

Once the optimal ARIMA model has been identified using the pre-intervention data, it is necessary to establish the intervention order *b*, *r*, and *s*, by analyzing the cross-correlation graph that compares the post-intervention training data with the forecasted values generated by the pre-intervention ARIMA model. The results are presented in **Figure 5**.

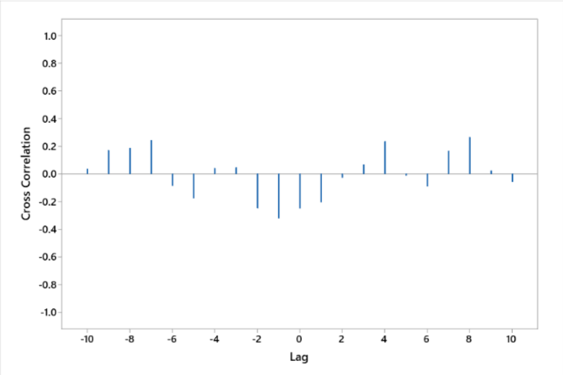


Figure 5 Cross-Correlation Plot

Based on **Figure 5**, the order for the intervention model can be identified, indicating that the futures gold price experiences an increase at the point of intervention. This shows an intervention order of $b = 0$ because there is no delay in the start of the intervention, $r = 2$ is due to the presence of a wave pattern in the intervention, and $s = 0$ because the duration of an intervention does not affect the data after a significant lag.

The findings from the estimation of the pulse function intervention model are displayed in **Table 8**.

Table 8 Results of Significant Test for Intervention Model

Intervention Model	Parameter	Estimate	P-Value	Description
ARIMA (0,2,1)	MA(1,1)	0.95918	0.000	Significant

Intervention Model	Parameter	Estimate	P-Value	Description
with $b = 0$, $r = 2$ and $s = 0$	ω_0	71.8761	0.011	Significant
	δ_1	-1.2392	0.000	Significant
	δ_2	-0.7995	0.000	Significant

Based on the **Table 8**, it was found that the ARIMA model parameters (0,2,1) with orders $b = 0$, $r = 2$, and $s = 0$ have passed the model significance test. Then, check the assumptions of residual white noise and normality.

Table 9 Results of the Assumption Test for the Intervention Model Residuals

P-Value Ljung-Box				Normality P-Value	AIC	SBC	Description
Lag 6	Lag 12	Lag 24	Lag 36				
0.1418	0.1098	0.1304	0.309	$>$ 0.150	1125.1	1135.86	White Noise and Normally Distributed

Based on the **Table 9**, it was found that the intervention model has met the white noise assumption because the p-values for all lags are greater than 0.05. Additionally, the residuals of the intervention model are normally distributed, as indicated by the normality p-value being greater than 0.05. Based on the output, the model has the smallest AIC and SBC values, namely 1125.1 and 1135.86.

D. *Forecasting and Intervention Analysis of Gold Futures Price Data*

Based on the discussion that has been conducted, the intervention model used to predict gold futures prices is as follows. The ARIMA (0,2,1) model with intervention parameters $b = 0$, $r = 2$, and $s = 0$ is written as follows:

$$Z_t = \frac{\omega_s(B)B^b}{\delta_r(B)} I_t^T + N_t \text{ with } N_t = (1 + 1.0231B + 1.0462\, B^2 + \cdots)a_t \quad (14)$$

with the value of $b = 0,\ r = 2,\ s = 0$ then:

$$Z_t = \frac{\omega_0(B)B^0}{\delta_2(B)} I_t^{(94)} + N_t$$
$$Z_t = \frac{\omega_0(B)}{(1 - \delta_1 B - \delta_2 B^2)} P_t^{(94)} + N_t$$
$$Z_t = (\omega_0 + \omega_0\delta_1 B + \omega_0(\delta_1^2 + \delta_2)B^2 + \omega_0(\delta_1^3 + 2\delta_1\delta_2)B^3 \ldots)P_t^{(94)} + N_t \quad (15)$$

Substitute the value of $\omega_0 = 71.87618,\ \delta_1 = -1.23923$ and $\delta_2 = -0.79957$ obtained from Table

$$Z_t = (71.87618 + 71.87618(-1.23923)B + 71.87618\big((-1.23923)^2 + (-0.79957)\big)B^2 + 71.87618((-1.23923)^3 + 2(-1.23923)(-0.79957))B^3 \ldots)P_t^{(94)} + N_t$$
$$Z_t = (71.87618 + (-89.07111)B + (-167.84963)B^2 + (-279.22290)B^3 \ldots)P_t^{(94)} + (1 + 1.0231B + 1.0462\, B^2 + \cdots)a_t$$
$$Z_t = 71.87618P_t^{(94)} + (-89.07111)P_{t-1}^{(94)} + (-167.84963)P_{t-3}^{(94)} + (-279.22290)P_{t-4}^{(94)} \ldots + a_t + 1.0231a_{t-1} + 1.0462\, a_{t-2} + \cdots \quad (16)$$

From the model above, predictions will be made for the next 7 weeks. The prediction results can be seen in the **Table 10**.

Table 10 Results of Gold Futures Price Prediction

Date	Prediction Results	Actual Data	APE (%)	Squared Error
16/06/2024	2405.183	2353.9	2.1786	2629.95
23/06/2024	2362.582	2362.5	0.0034	0.00672
30/06/2024	2393.578	2421.1	1.1367	757.460
07/07/2024	2443.539	2444.5	0.0393	0.92355
14/07/2024	2461.166	2423.0	1.5751	1456.64
21/07/2024	2453.116	2404.5	2.0218	2363.52
28/07/2024	2418.68	2469.8	2.0697	2613.25
			MAPE = 1.289	MSE = 1403.11

Based on the **Table 10**, the prediction results have a MAPE value of 1.289% and a small MSE value of 1403.11. This shows that the model has very good predictive ability and it can be concluded that the prediction for the next 7 weeks after the 104th data has no significant difference with the actual data. The comparison graph shows that the overall prediction results tend to be close to the actual data. However, the fluctuation pattern of the predicted data (See **Figure 6**)shows a decreasing trend while the actual data shows an increasing trend. The difference between the prediction results and the actual data can be caused by several other factors such as inflation, interest rates and other economic problems.

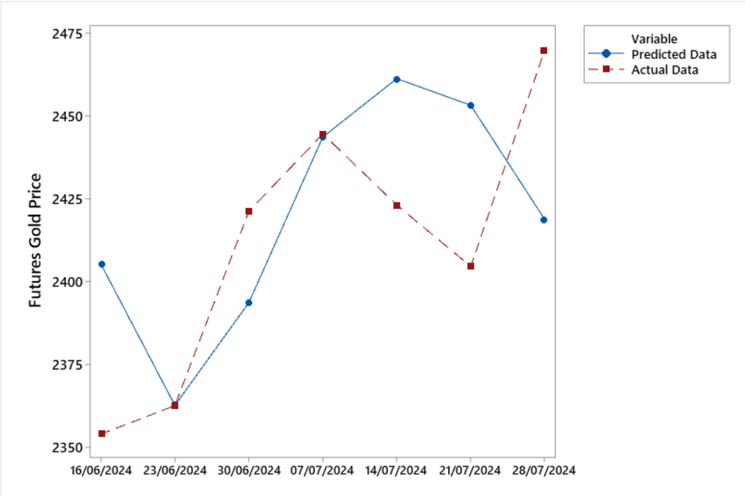


Figure 6 Comparison Plot of Actual Data and Predictions

V. CONCLUSIONS AND SUGGESTIONS

A. *Conclusions*

Based on the analysis and discussion results, the conclusions drawn from this study are as follows:

1. The futures gold price before the intervention averaged approximately 1927.5 USD/t. oz with a standard deviation of 133.2 USD/t. oz. After the intervention. the average futures gold price was 2392.3 USD/t. oz with a standard deviation of 39.5 USD/t. oz.
2. The probabilistic ARIMA (0,2,1) model is the best model derived from pre-intervention data of futures gold prices. The probabilistic ARIMA (0,2,1) model can be mathematically expressed as follows:

$$\dot{Z}_t = a_t + 1.0231a_{t-1} + 1.0462\, a_{t-2} + \cdots \text{ with } \dot{Z}_t = Z_t$$

3. The best intervention model for futures gold prices is the ARIMA (0.2.1) model with pulse function intervention parameters $\boldsymbol{b} = \mathbf{0}.\ \boldsymbol{r} = \mathbf{2}$ and $\boldsymbol{s} = \mathbf{0}$. with the model equation as follows:

$$\dot{Z}_t = 71.87618P_t^{(94)} + (-89.07111)P_{t-1}^{(94)} + (-167.84963)P_{t-3}^{(94)} + (-279.22290)P_{t-3}^{(94)} \ldots + a_t + 1.0231a_{t-1} + 1.0462\, a_{t-2} + \cdots$$

4. The prediction results from the best model using the intervention analysis approach with pulse function show that the overall prediction tends to be close to the actual data. However, the fluctuation pattern of the predicted data shows a decreasing trend while the actual data shows an increasing trend. The difference between the prediction results and the actual data can be caused by several other factors such as inflation, interest rates and other economic problems. The model has a MAPE value of 1.289% and a small MSE value of 1403.11. The model also has the smallest AIC and SBC values, namely 1125.1 and 1135.86 this indicates that the model has very good predictive capability.

B. Suggestions

Based on the results and discussion in the previous chapter, the author offers the following recommendations:

1. The findings of this research can be utilized by parties involved in gold import and export to understand gold price predictions. which can provide considerations for buying and selling gold.
2. For future researchers interested in the same topic. it is suggested to investigate the factors influencing futures gold prices using different methods such as machine learning techniques

ACKNOWLEDGEMET

The authors would express the gratitude to Universitas Airlangga, Faculty of Science a and Technology, especially Statistics Study Program for upportunity given to implement the knowledge to this research and publication. Thanks to the accompanying lectures, namely Drs. H. Sediono, M.Si. and Idrus Syahzaqi, S.Stat,. M.Stat. who have provided guidance during this research.

REFERENCES

[1] E. A. Florentina, J. Gunawan, and N. S. Hakim, “Analisis Daya Saing Produk Emas Indonesia di Pasar Global,” *Jurnal Sains dan Seni ITS*, vol. 11, no. 5, pp. I47–I52, 2023.

[2] M. B. Triki and A. Ben Maatoug, “The GOLD market as a safe haven against the stock market uncertainty: Evidence from geopolitical risk,” *Resources Policy*, vol. 70, p. 101872, 2021.

[3] I. F. Puspitasari, N. Andriyani, and N. Hidayah, “Emas sebagai safe haven dan inflation hedging di tengah ketidakpastian perekonomian global selama pandemi Covid-19,” *JURNAL PENDIDIKAN EKONOMI: Jurnal Ilmiah Ilmu Pendidikan, Ilmu Ekonomi Dan Ilmu Sosial*, vol. 16, no. 2, pp. 250–258, 2022.

[4] T. C. Chiang, “The effects of economic uncertainty, geopolitical risk and pandemic upheaval on gold prices,” *Resources Policy*, vol. 76, p. 102546, 2022, doi: https://doi.org/10.1016/j.resourpol.2021.102546.

[5] S. Baker, N. Bloom, S. Davis, and S. Terry, “COVID-induced economic uncertainty and its consequences,” *VoxEU. org*, vol. 13, 2020.

[6] John Baffes, Jeetendra Khadan, and Kaltrina Temaj, “Gold Shines Amid Geopolitical Uncertainties ,” World Bank . Accessed: Nov. 04, 2024. [Online]. Available: https://blogs.worldbank.org/en/opendata/gold-shines-amid-geopolitical-uncertainties

[7] M. K. Alam, M. I. Tabash, M. Billah, S. Kumar, and S. Anagreh, “The impacts of the Russia–Ukraine invasion on global markets and commodities: a dynamic connectedness among G7 and BRIC markets,” *Journal of Risk and Financial Management*, vol. 15, no. 8, p. 352, 2022.

[8] C. Andreas, I. A. Rahmayanti, and S. M. Ulyah, “The impact of US-China trade war in forecasting the gold price using ARIMAX model,” in *AIP Conference Proceedings*, AIP Publishing, 2021.

[9] A. F. Yuliana and R. Robiyanto, “Peran Emas Sebagai Safe Haven Bagi Saham Pertambangan Di Indonesia Pada Periode Pandemi Covid-19,” *Jurnal Ilmiah Bisnis dan Ekonomi Asia*, vol. 15, no. 1, pp. 1–11, 2021.

[10] World Gold Council, “2024 Central Bank Gold Reserves Survey.” [Online]. Available: https://www.gold.org/goldhub/data/2024-central-bank-gold-reserves-survey

[11] N. Aliffia, S. Sediono, S. Suliyanto, M. F. F. Mardianto, and D. Amelia, “MODELING OF WORLD CRUDE OIL PRICE BASED ON PULSE FUNCTION INTERVENTION ANALYSIS APPROACH,” *MEDIA STATISTIKA*, vol. 16, no. 2, pp. 136–147, Feb. 2024, doi: 10.14710/medstat.16.2.136-147.

[12] R. Elizabeth and S. Sitorus, “Gold Price Forecasting Using Autoregressive Integrated Moving Average (ARIMA) Method,” *Journal of Mathematics Technology and Education*, vol. 1, no. 1, pp. 11–18, 2021.

[13] M. B. Triki and A. Ben Maatoug, “The GOLD market as a safe haven against the stock market uncertainty: Evidence from geopolitical risk,” *Resources Policy*, vol. 70, p. 101872, 2021.

[14] D. Nudia, “Emas Sebagai Instrumen Investasi Jangka Panjang,” *Shar-E: Jurnal Kajian Ekonomi Hukum Syariah*, vol. 8, no. 1, pp. 177–187, 2022.

[15] I. Kurniawan, “Analisis keuntungan investasi emas dengan IHSG,” *Jurnal manajemen bisnis dan kewirausahaan*, vol. 3, no. 2, 2019.

[16] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[17] W. W. S. Wei, “Time series analysis: univariate and multivariate,” *Methods. Boston, MA: Pearson Addison Wesley*, 2006.

[18] D. C. Montgomery, C. L. Jennings, and M. Kulahci, “Introduction to Time Series Analysis and Forecasting (Wiley Series in Probability and Statistics).”

[19] H. N. Tantika, N. Supriadi, and D. Anggraini, "Metode seasonal ARIMA untuk meramalkan produksi kopi dengan indikator curah hujan menggunakan aplikasi R di kabupaten Lampung Barat," *Matematika: Jurnal Teori dan Terapan Matematika*, vol. 17, no. 2, 2018.

[20] T. Trimono, A. Sonhaji, and U. Mukhaiyar, "Forecasting Farmer Exchange Rate in Central Java Province Using Vector Integrated Moving Average," *Media Statistika*, vol. 13, no. 2, pp. 182–193, 2020.

[21] Investing, "Emas Berjangka GCZ4." Accessed: Oct. 05, 2024. [Online]. Available: https://id.investing.com/commodities/gold-historical-data