Research on Trading Strategy Selection Based on ARIMA and Linear Programming

Yuan Liu[*,†], Shiqin Zhou[†]

Department of Management and Economics, Tianjin University, Tianjin, 300072

[*]Corresponding author: liuyuan220317@163.com

[†]These authors contributed equally.

Keywords: ARIMA; Linear Programming; Trading Strategy; Risk Control.

Abstract: The investment trading market is changing rapidly; how do the traders choose trading strategies in the highly uncertain trading market? Take gold and bitcoin as an example. Based on Bitcoin and Gold price data from September 11, 2016, to September 10, 2021, we used the ARIMA algorithm to predict the closing prices of bitcoin and gold in five years, R^2 is over 0.9 so that the predicted data are similar to the actual data. Then, we built a linear programming model with the maximization of return as the objective function and the variation of cash as the constraint while considering the risk factor. We proposed to split the feasible region to solve the linear programming model where the constraints contain absolute values. Moreover, finally, we obtained the strategy that could be provided for the trader, and the total value of the assets, in the end, is 42,051.42 USD.

1. Introduction

Market traders often buy and sell volatile assets with the expectation of maximizing the total value.[1] Moreover, gold and bitcoin, as the most popular trading varieties, have greater instability in this process.[2] On the one hand, with the deepening of information technology in the industry, a large amount of financial data has been accumulated; on the other hand,[3] with the continuous rise of financial technology, quantitative analysis has become a common analysis tool in market trading.[4] Combining a large amount of financial data with quantitative analysis can consistently achieve long-term and steady returns for investors while bringing more excellent social value and economic benefits.[5]

2. Financial Product Price Prediction Model Based on ARIMA

For the sake of discussion, let us assume we start with a thousand dollars. Our problem is to choose the right investment strategy, which necessarily requires us to predict the price trend of bitcoin and gold. Thus, one must first forecast the price over five years to solve the problem.

We used bitcoin and gold price data from September 11, 2016, to September 10, 2021. After testing, it was found that the given data source has missing data for gold on some days, as opposed to bitcoin data. This missing data is that gold is only traded on trading days, but bitcoin is traded all day.

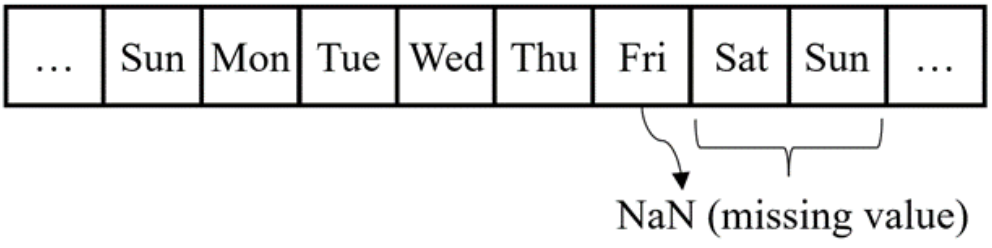


Figure 1. Data Cleaning.

In order to analyze the data, the missing values are processed, filled with the last adjacent value, which is in line with the real-life logic: on non-trading days, the price of gold keeps the same as the closing price of the previous trading day.

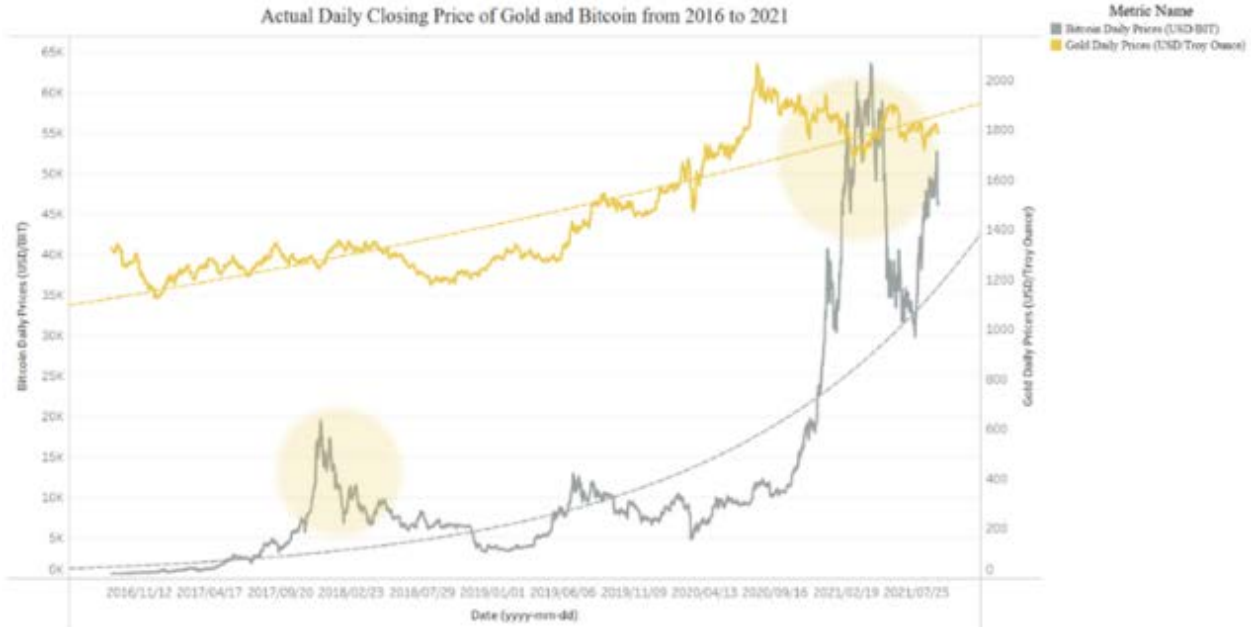


Figure 2. Actual Daily Closing Price of Gold and Bitcoin.

As you can visually see from the chart, the prices of both bitcoin and gold are generally trending upwards. However, what cannot be ignored is that, on the one hand, bitcoin has risen much more than gold. On the other hand, bitcoin is far more volatile than gold. Since 2016, Bitcoin has seen two particularly large swings, the exact places identified in the chart. Gold is more of a preservation asset than bitcoin, and in times of market turmoil, it would be wise to swap bitcoin for gold properly.

A more refined and accurate algorithm for analysis and forecasting of time series data is the Box-Jenkins method, whose common models include: Autoregressive model (AR model),[6] Moving average model (MA model),[7] Autoregressive moving average model (ARMA model),[8] and Autoregressive integrated moving average model (ARIMA model).[9] Among them, the ARIMA model requires only endogenous variables without the help of other exogenous variables, which exactly fits the characteristics of the data source given in the topic. Moreover, the ARIMA model combines the advantages of the AR model, MA model, and ARMA model, and the forecasting performance is relatively better. Therefore, it is chosen as the model used for forecasting.

The ARIMA, the Autoregressive integrated moving average model, separates signal and noise by taking past observations and considering the difference between autoregressive and moving average components. The optimal model parameters can be found by minimizing the AIC (Akaike Information Criterion).

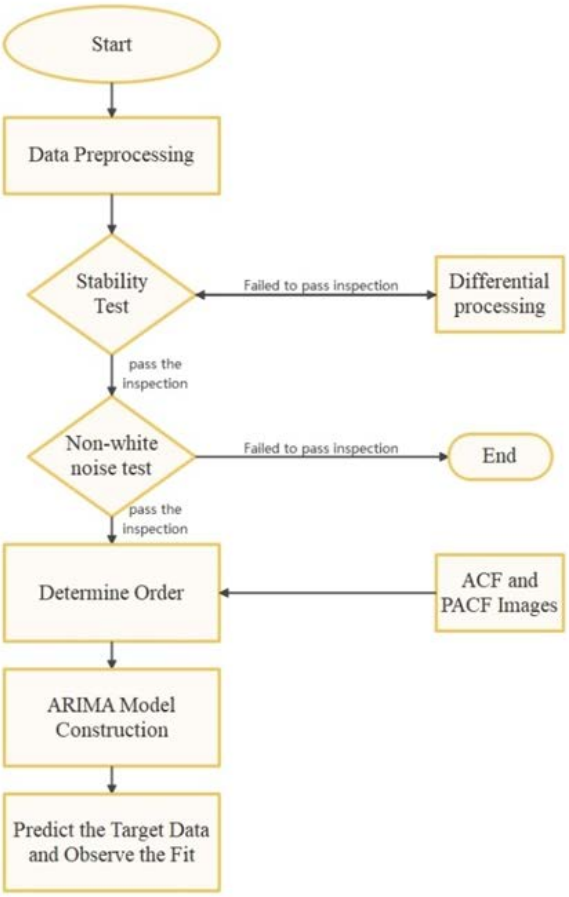


Figure 3. Modeling Steps of ARIMA Model.

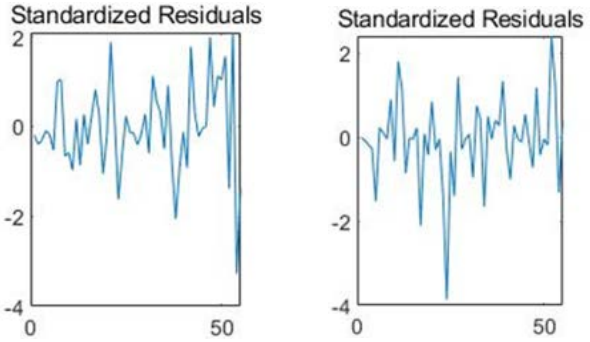
Using the ARIMA algorithm for prediction, the following relationship equation holds:

$$y_t = \propto_1 y_{t-1} + \propto_2 y_{t-2} + \cdots + \propto_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + c + \varepsilon_t \quad (1)$$

Differentiate the sample time series, and if the mean and variance of the series do not change significantly after processing, the fitted curve meets the requirement of stability. The difference is carried out according to the following relation:

$$f_t = f_t - f_{t-1} \quad (2)$$

Performing one differencing on the bitcoin data yields the following bitcoin data curve, as shown in Figure 4(a):

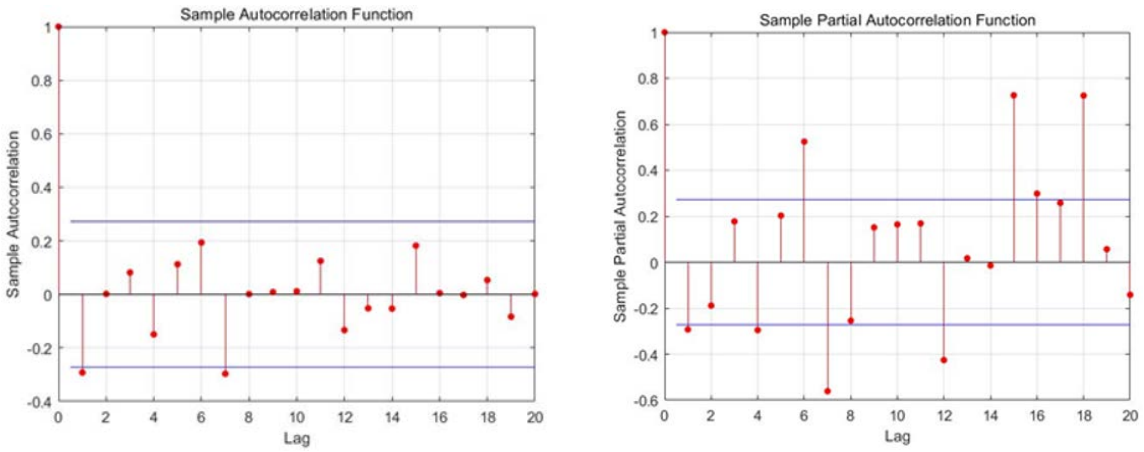


(a) Bitcoin (b) Gold

Figure 4. Stability Test.

For the gold data to be differenced twice, the gold data curve obtained is shown in Figure 4(b).

After applying stability test to the bitcoin data, we can obtain its autocorrelation and partial autocorrelation plots using SPSS software, as shown in Figure 5:

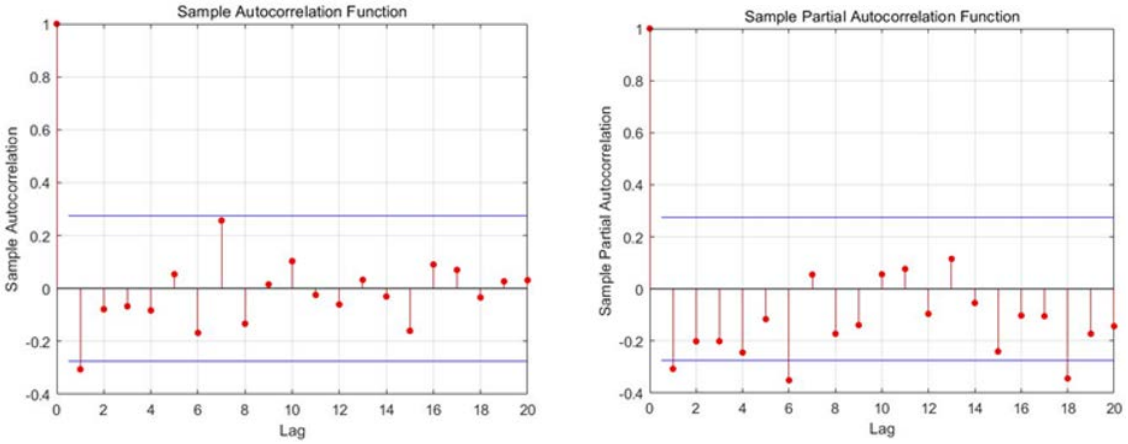


(a) Autocorrelation. (b) Partial Autocorrelation

Figure 5. Determine the Order of the Bitcoin Data.

Based on the truncated, it can be determined that the bitcoin data corresponds to p, q of 5, 5 respectively.

After smoothing the gold data, we can use SPSS software to get its autocorrelation diagram and partial autocorrelation diagram, as shown in Figure 6:



(a) Autocorrelation. (b) Partial Autocorrelation.

Figure 6. Determine the Order of the Gold Data.

The gold data can be determined to correspond to p, q of 1, 1 respectively, based on the truncate.

From this, we have determined the respective price forecasts for bitcoin and gold over five years based on the ARIMA model.

The results are shown in Figure 7 and Figure 8:

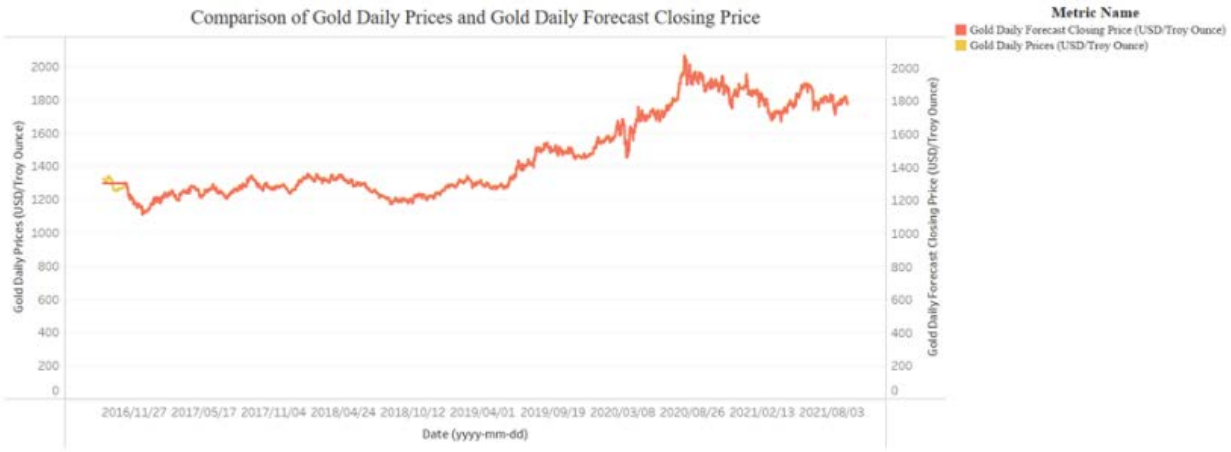


Figure 7. Gold Daily Forecast Closing Price.

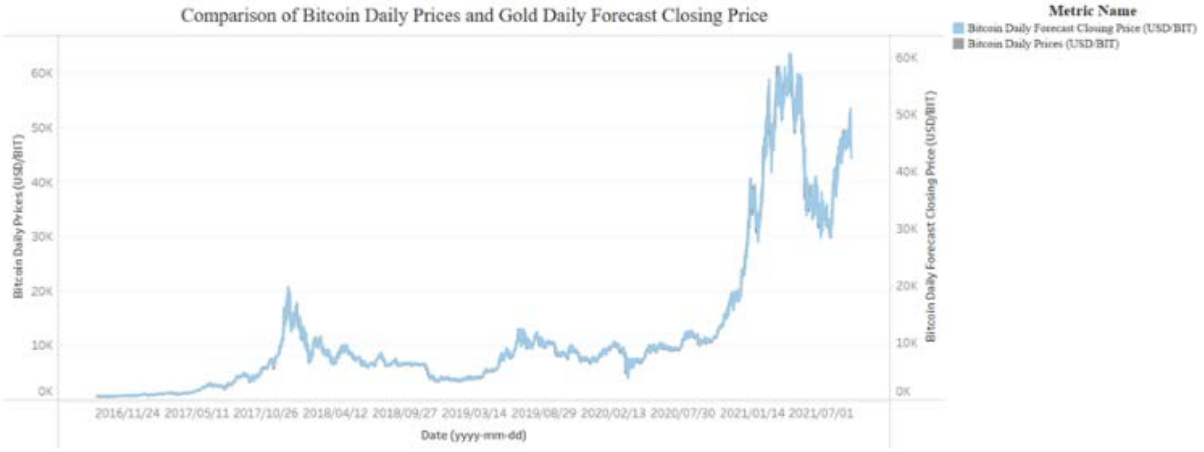


Figure 8. Bitcoin Daily Forecast Closing Price.

Among them, the predicted price of gold and the historical data satisfy the following equation:

$$y_{Gt} = 0.32687y_{G(t-1)} - \varepsilon_{G(t-1)} - 8.3586e^{-5} + \varepsilon_{Gt} \tag{3}$$

Likewise, for Bitcoin, it satisfies the following equation:

$$\begin{aligned} y_{Bt} = {} & 0.32902y_{B(t-1)} - 0.58875y_{B(t-2)} - 0.40092y_{B(t-3)} - 0.92569y_{B(t-4)} \\ & +0.92569y_{B(t-5)} - 0.73078\varepsilon_{B(t-1)} + 0.79933\varepsilon_{B(t-2)} - 0.75667\varepsilon_{B(t-3)} \\ & +0.024274\varepsilon_{B(t-4)} - 0.33615\varepsilon_{B(t-5)} + 0.678678 + \varepsilon_{Bt} \end{aligned} \tag{4}$$

3. Trading Strategy Based on Linear Programming

Having obtained daily forecast data, traders need to make investment decisions based on the past data and forecast data of the day. As mentioned in Assumption 1, traders only need to consider two factors: return and risk when making investment decisions. Therefore, we try to maximize the benefits and minimize the risks. So, we use a linear programming model to make decisions.

Linear programming studies the extremum problem of a linear objective function under linear constraints [10].

It is generally solved using the Simplex algorithm.

The data obtained from the forecasts are not yet ready to build a linear programming model. We need to introduce new variables:

Table 1. Variables

Symbol	Description
y_{Bt}	The actual price of bitcoin on day t
y_{Gt}	The actual price of gold on day t
f_{Bt}	Forecast price of bitcoin on day t
f_{Gt}	Forecast price of gold on day t
$MA5_{Bt}$	The bitcoin value of the five-day moving average on day t
$MA5_{Gt}$	The gold value of the five-day moving average on day t
k_t	Bitcoin range of change on day t
(C_t, B_t, G_t)	Cash, bitcoin, and gold holdings on day t when the stock market closes

Then, we construct the Linear Programming Model:

$$\max \quad C_t + B_t f_{Bt} + C_t f_{Gt} - C_{t-1} - B_{t-1}y_{B(t-1)} - G_{t-1}f_{G(t-1)} \tag{5}$$

$$s.t.\begin{cases} C_t = C_{t-1} - |B_t - B_{t-1}| \times y_{B(t-1)} \times 0.02 - |G_t - G_{t-1}| \times f_{G(t-1)} \times 0.01 \\ -(B_t - B_{t-1}) \times y_{B(t-1)} - (G_t - G_{t-1}) \times f_{G(t-1)} \\ n_t(G_t - G_{t-1}) = 0 \\ m_t(B_t - B_{t-1}) = 0 \\ k_t(G_t - G_{t-1}) = 0 \\ C_t, G_t, B_t \geq 0 \end{cases} \tag{6}$$

Ct, Bt, and Gt are the only decision variables, and other parameters are known variables.

The absolute value condition exists in the linear programming constraint, and we handle it by dividing the solution space into four parts based on the relationship of $G_t, G_{t-1}, B_t,$ and B_{t-1} dividing the solution space into four parts.

$$G_t \geq G_{t-1}, B_t < B_{t-1} \tag{7}$$

$$G_t < G_{t-1}, B_t \geq B_{t-1} \tag{8}$$

$$G_t < G_{t-1}, B_t < B_{t-1} \tag{9}$$

$$G_t \geq G_{t-1}, B_t < B_{t-1} \tag{10}$$

The optimum solution $G_t, G_{t-1}, B_t,$ and B_{t-1} is found in four different solution spaces with the optimum value corresponding to the objective function, where the optimal solution with the largest value is the optimum solution corresponding to the original linear programming model.

The model is imported into MATLAB to solve it cyclically, and the following data results are obtained.

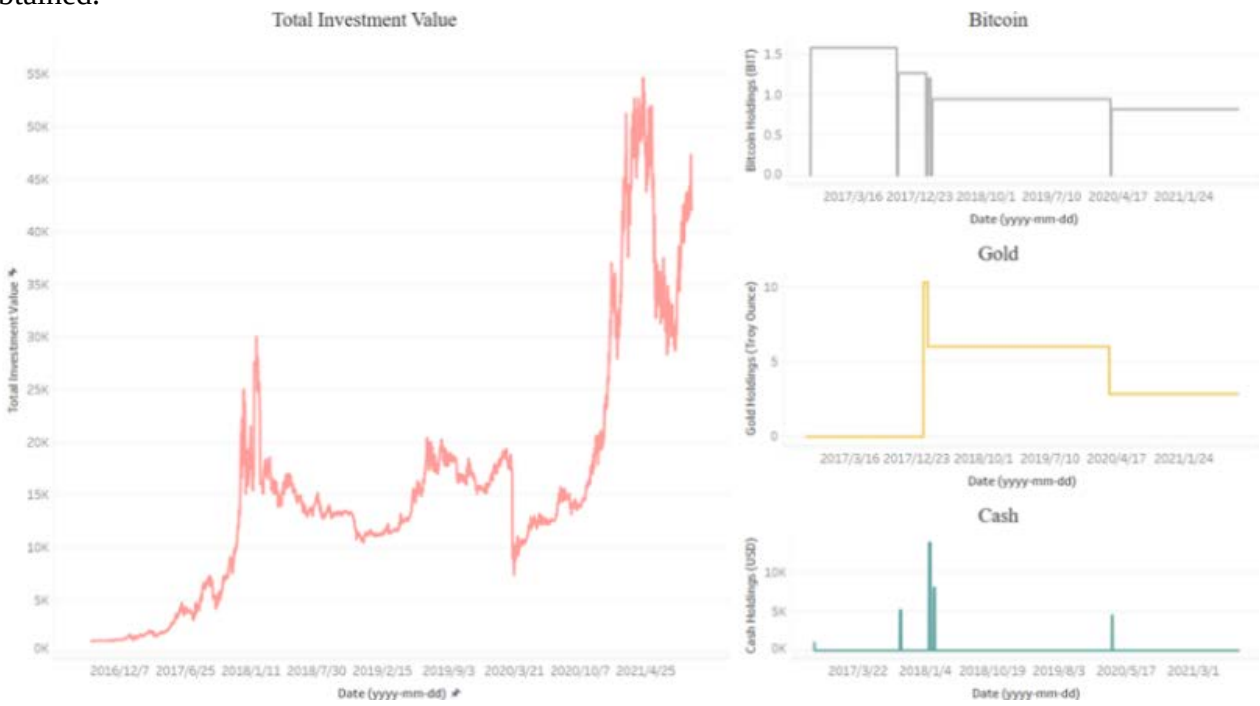


Figure 9. The Strategy of Buying Assets.

4. Conclusion

In this paper, we used an ARIMA model to predict the five-year closing prices of bitcoin and gold. Secondly, we evaluated the daily trading risk and built a linear programming model to find the optimal trading strategy for each day, taking the trading risk and the estimated return into account. When choosing a prediction model, we did not choose a complex model such as neural networks or a series of other complex models. We believe that the model is used to solve practical problems. As long as the chosen model solves the problem well enough, it is a suitable model, rather than saying that the

more complex the model is, the better. So, we chose the ARIMA model for prediction. And R^2 is very close to 1, and the actual prediction result is rather good. We believe that traders will be forced to liquidate when the bitcoin price falls too much. It is a means of managing risk. However, maybe we can find a better way to deal with it in practice. In this article, we mainly focus on investing in bitcoin and gold; however, our approach can be used in other investment situations, as it takes both risk and reward into consideration.

References

[1] Gan Ke. Analysis of factors affecting bitcoin price [D]. Central China Normal University, 2021.

[2] Chen Zheshi. Research on the Internet and Financial Futures Influencing Factors and Prediction Methods of Bitcoin Price [D]. Harbin Institute of Technology, 2021.

[3] Ye Wuyi, Sun Liping, Miao Boqi. Research on Dynamic Cointegration of Gold and Bitcoin: Based on Semiparametric MIDAS Quantile Regression Model [J]. Systems Science and Mathematics, 2020, 40(07): 1270 -1285.

[4] Liu Xiaofeng. An Empirical Study on the Influencing Factors of Digital Currency Prices [D]. Shandong University, 2020.

[5] Hu Shuyue. Research on the asset attributes of Bitcoin [D]. Southwestern University of Finance and Economics, 2019.

[6] Wen Pan, Ji Kun, Wen Ruizhi, Ren Yefei. Research on the applicability of multi-dimensional autoregressive (AR) model to calculate spatial coherence function of ground motion [J]. Vibration and Shock, 2022, 41(02): 140-150.

[7] Wang Huanting. Bayesian analysis of moving average model MA under time series [J]. Journal of Huaihai Institute of Technology (Natural Science Edition), 2019, 28(02): 1-4.

[8] Su Rong. Thoughts on bond investment in the post-epidemic era—Analysis based on ARMA model [J]. China Business Review, 2021(23):83-86.

[9] Ren Min, Wang Chenming, Yan Hongyan. The impact of the new crown pneumonia epidemic on the stock price of my country's pharmaceutical industry: an empirical analysis based on the ARIMA model [J]. Northern Economic and Trade, 2022(02): 93-96.

[10] Xia Jing. Stock portfolio based on linear programming [J]. China Management Information, 2019, 22(21): 142-143.