Financial Engineering and Risk Management (2022)
Clausius Scientific Press, Canada

DOI: 10.23977/ferm.2022.050315
ISSN 2523-2576 Vol. 5 Num. 3

# *Design of currency portfolio strategy based on ARIMA-SAM-GDBT*

**Meixiang Zhai[1], Keming Chen[1], Yana Wang[1], Xiaojun Men[2,*]**

*[1]School of Chemical Engineering, North China University of Science and Technology, Tangshan, Hebei, 063210, China*
*[2]School of Science, North China University of Science and Technology, Tangshan, Hebei, 063210, China*

***Keywords:*** ARIMA, SAM, GDBT, residual white noise test

***Abstract:*** This paper uses AIRMA model to make short-term price prediction. In order to avoid the error of long-term prediction, the SAM-GDBT model was established, and the trading strategy was obtained by comparing the importance values of characteristics of buying and selling behaviors at different times. Further, residual white noise test was carried out for the above models, and compared with other predictional models such as the limit tree (Extra) model. This paper adopts the way of adjusting commission to reflect the change of transaction cost. The increase of transaction cost will naturally affect the buying tendency of traders and the transaction proportion.

## 1. Introduction

A common means for market traders to maximize total returns is by frequently trading volatile assets such as gold and bitcoin. Typically, there is a commission on each transaction, which is 1% for gold and 2% for bitcoin, and there is no cost to holding assets. And it's important to note that gold can only be traded when the market is open, while bitcoin can be traded every day. In addition, the price of bitcoin fluctuates greatly and the investment risk is high [1], while gold is more stable than bitcoin. Therefore, how to obtain the maximum return with the minimum risk is the topic of concern of investors. It is of great significance to understand the fluctuation law of currency trading price, explore its internal trading mechanism, predict the trend of currency trading accurately and reduce the investment risk of investors.

In this case we need to support an investment strategy for a user with $1000. To do this, we need to do the following: Developing a model that yields the optimal trading strategy based on the currency price of the day and the initial investment of the user's assets on September 10, 2021. It is proved that the model has good accuracy and the optimal strategy can be given [2]. Determining the sensitivity of strategy to transaction cost, and finding out the process and results of the impact of transaction cost on strategy.

To develop the best daily trading strategy, this paper first consider using ARIMA time series to predict the price of currencies and determine whether traders buy or sell by observing price changes. Therefore, we need to determine the transaction amount [3]. In this paper, we consider the use of GBDT to find the best time for gold and bitcoin transactions, moving average analysis with the help

of SMA to obtain the importance of the characteristics of the changes in the amount of gold and bitcoin transactions by observing the ratio. In addition, we consider the use of residual white noise test to analyze the prediction results and compare with the current popular prediction models. Finally, a sensitivity analysis was performed on the considered models to observe the magnitude of changes in total assets。

## 2. ARIMA model

ARIMA model, called autoregressive-integral moving average model, usually uses first-order difference to smooth linear trends and second-order difference to second-order curves [4].

The autoregressive moving average model can be expressed as:

$$X_i = \emptyset_1 X_{i-1} + \cdots + \emptyset_p X_{i-p} + \varepsilon_i - \theta_1 \varepsilon_{i-1} - \cdots - \theta_q \varepsilon_{i-q} \tag{1}$$

Difference d of time series refers to the number of differences needed to obtain stationary series for time series. Rank p and rank q were determined by the autocorrelation coefficient and partial autocorrelation coefficient of stationary series, and the model with the smallest AIC and BIC was the best. AIC is the information of red pool, which can be expressed as:

$$\mathrm{AIC} = -2\ln L + 2k \tag{2}$$

BIC is Bayesian information, which can be expressed as:

$$\mathrm{BIC} = -2\ln L + k\ln n \tag{3}$$

We preprocess the data and observe whether its time series graph conforms to stationary series. Then, we use origin for visualization and get the time series diagram, which is shown in Fig.1.

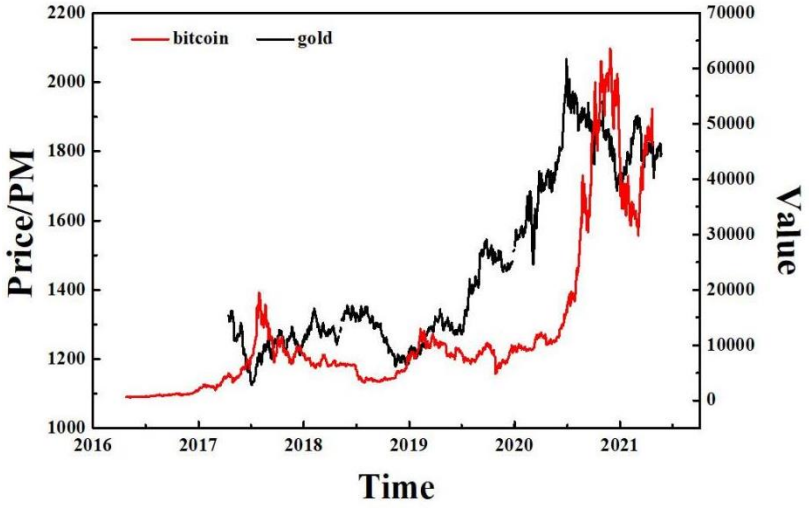


Figure 1 Time series chart of gold and bitcoin price (value)

As can be seen from Fig. 1, the prices of both gold and bitcoin show an upward trend on the whole, and there are significant changes after 2019, which may be random fluctuations. At the same time, we visualized the data in 2019, as shown in Fig. 2, and found that the data showed an overall upward trend without seasonality, so we could judge that the series was non-stationary, and then carried out differential processing.

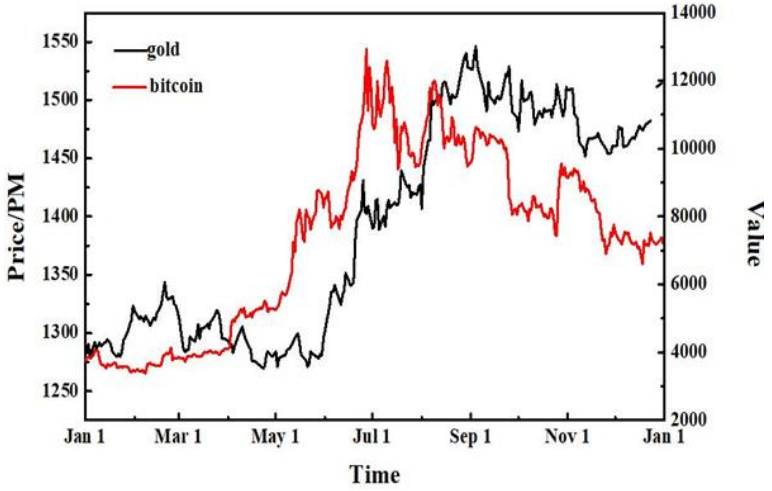


Figure 2 Time series chart of gold and bitcoin price (value) in 2019

The data series are differenced using the function diff () in MATLAB, and by calculating the AIC data after differencing at different orders, we find that the smallest AIC values after first-order differencing are 9948.534 and 29151.816 for gold and bitcoin, respectively. The results is shown in table 1.

Table 1 Gold and bitcoin ACF test table

| The differential order | Variable | t | p | AIC |
|---|---|---|---|---|
| 0 | gold | -0.434 | 0.904 | 9957.828 |
| | bitcoin | -0.238 | 0.934 | 29168.936 |
| 1 | gold | -8.159 | <0.01 | 9948.534 |
| | bitcoin | -8.535 | <0.01 | 29151.816 |
| 2 | gold | -12.877 | <0.01 | 9993.297 |
| | bitcoin | -15.723 | <0.01 | 29187.267 |

Since the smaller the value of AIC and p, the better the effect of the model. Comparing the AIC and p data, we choose the first-order difference analysis, as Fig. 3 shows that the mean and variance of the time series of the first-order difference have been basically smooth, so the number of differences d is set to 1 in this paper.

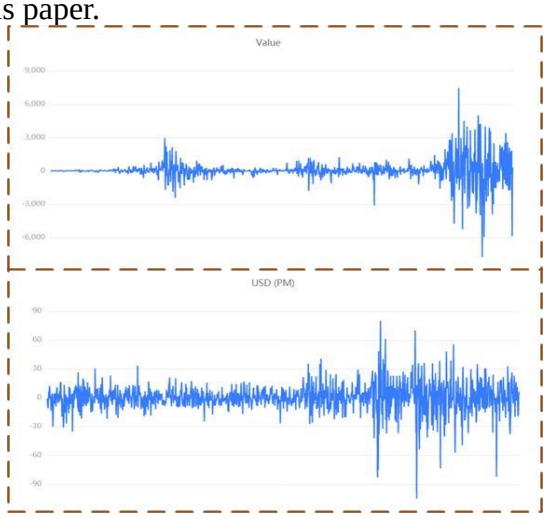


Figure 3 Bitcoin, gold first order difference sequence diagram

The ACF calculation formula:

$$\mathrm{ACF} = \sum_{t=k+1}^{n} \frac{(Z_t-\bar{Z})(Z_{t-k}-\bar{Z})}{\sum_{t=1}^{n}(Z_t-\bar{Z})^2} \tag{4}$$

We use MATLAB to calculate the autocorrelation coefficient (ACF) of the data sequence, and the data results are shown in Figs. 4(a) and 4(b).

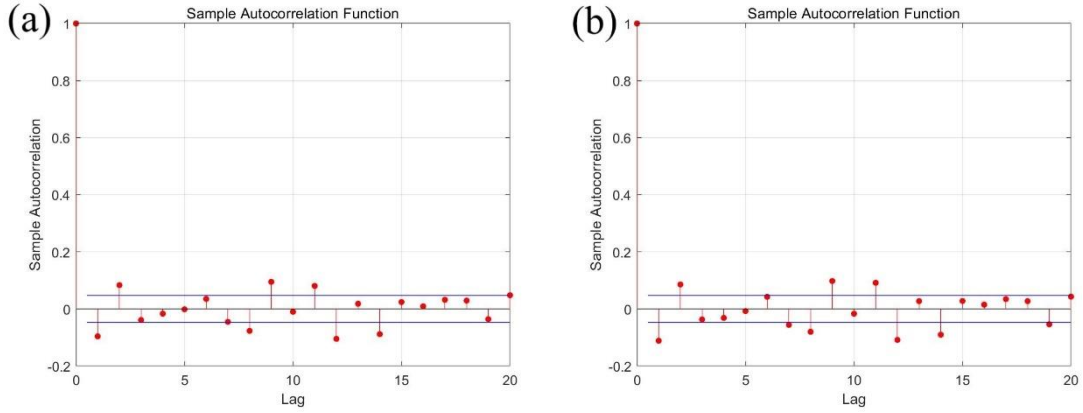


Figure 4 ACF diagram of bitcoin(a) and gold(b) after first-order difference sequence

As can be seen from the Figs. 4, the autocorrelation plots of both gold and bitcoin are untruncated and suitable for an ARIMA model. We finally settled on a first-order difference ARIMA model for forecasting. After comprehensive comparison of AIC and AFC data, we finally decided to adopt ARIMA (0, 1, 0) model for prediction.

Through the above analysis, ARIMA (0, 1, 0) model is finally selected for prediction. Here, we predict the time area of the given data and compare it with the actual data, as shown in Fig.5(a) and Fig.5(b). The results show that the $R^2$ of the predicted data gold and bitcoin is 0.998 and 0.999 respectively, which indicates that our model has good reliability in short-term prediction.


Figure 5 ARIMA test data predicts evaluation results

## 3. SMA mean line analysis and GBDT dynamic model

The simple moving average (SMA) is one of the core technical indicators used by traders and investors for technical analysis of stocks, indexes or securities. Through observation, we find that the trading of bitcoin and gold is similar to the trading of stocks, so we choose SMA for long-term ema analysis to determine bull and bear markets [5], which can be expressed as:

$$\mathrm{MA} = (C_1 + C_2 + \cdots + C_n)/n \tag{5}$$

where C is the closing price of the day (gold and bitcoin prices), and N is the moving average

period.

Here we use 100 days and 300 days as moving average cycles. Based on this one can derive the exact time point for trading, as well as the buy-sell strategy, as shown in Fig. 6, where the red positive triangle indicates the best buy point and the green inverted triangle indicates the best sell point, and the long-term strategy can ignore the time when gold and bitcoin trading do not overlap and can exclude the case when gold is not in the trading day.

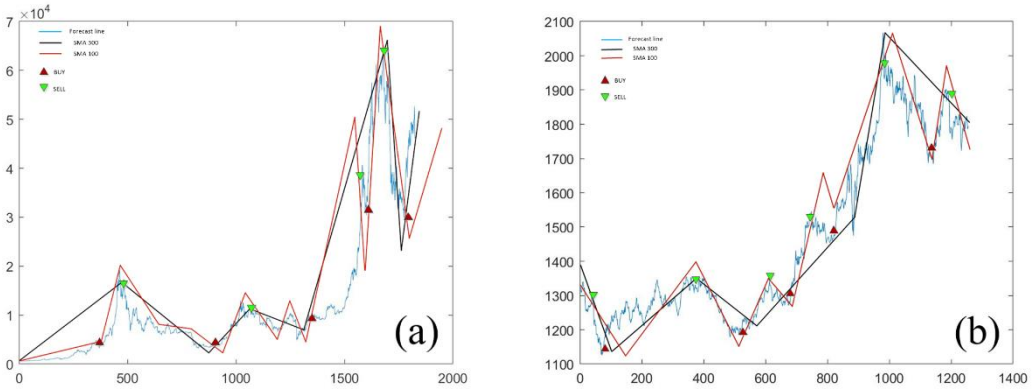


Figure 6 Bitcoin(a) and Gold(b) SMA analysis diagram

GBDT model is an additive model, which trains a group of CART regression trees serally and finally adds the predicted results of all regression trees to obtain a strong learner, and each new tree fits the direction of the negative gradient of the current loss function [6]. Here we use GBDT for dynamic planning, combined with SMA average line, to help customers find the most appropriate trading points and trading strategies.

Through the GBDT model of machine learning, we came up with the optimal trading ratio at the trading point. In the final decision of the day's trading ratio and whether to buy or sell, this paper presuppose that all US dollars will be put into the transaction, and in the proportion of gold and bitcoin trading, we consider to make the purchase ratio according to the ratio of their characteristic importance [7]. The results is shown in Fig.7.

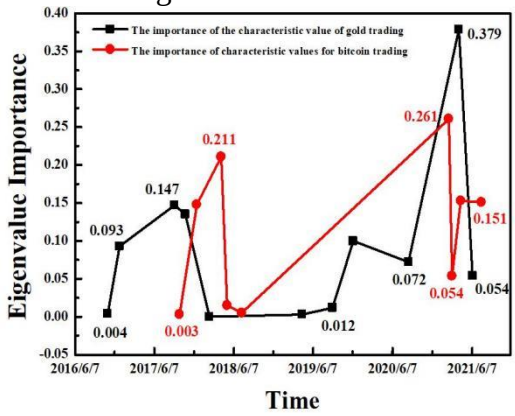


Figure 7 The importance of the characteristic value for bitcoin/gold trading

We calculate that on September 10, 2021, the user's initial investment value of $1000 is $36,269.8, and the profit is more than $35,000, indicating that our model has good guidance for user profit.

In order to prove that our model has good accuracy, we first conducted residual white noise test on the model in this paper [8]. A horizontal comparison was made between the model in this paper and prediction models such as Extra. Taking Extra as an example, as shown in Fig. 8, the fitting

degree of gold and bitcoin was only 0.971 and 0.95. In addition, the values of MSE, RMAE, MAE and MAPE of Extra model were (1006.819, 31.73, 24.144, 1.339), which were significantly higher than those of this model (5.5643, 2.35889, 2.1296, 0.012). The results show that the model has good prediction effect.

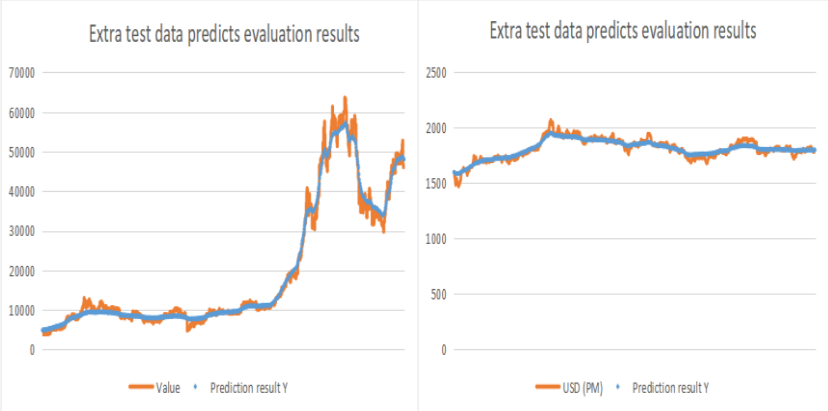


Figure 8 Prediction renderings of Extra

## 4. Sensitivity analysis

In order to verify the sensitivity of the model established in this paper to transaction costs, we consider observe the change of assets held by end users by fixing the commission of one currency, adjusting the level of the commission of another, and simultaneously increasing the commission, which is shown in Fig.9. By analyzing the changes in trading ratios, we can also intuitively see that there are large deviations in returns even after fine-tuning the trading ratios, so we not only prove again that our trading strategy is optimal, but also verify that the model has good sensitivity.


Figure 9 Adjusted return curve after trading ratio

## 5. Conclusion

In this paper, we use ARIMA model to predict the price of gold and bitcoin in the short term. Due to the limitations of short-term prediction, long-term price changes cannot be seen. Therefore, in order to improve the investment efficiency, we established a SAM-GBDT decision model, used the moving average analysis to find the analysis cycle, and then used the GBDT gradient decision tree to find the best trading time point.

## References

*[1] Xin Jianxuan. Analysis on the influencing factors of bitcoin price fluctuation [D]. Harbin Institute of Technology, 2020.*
*[2] Guilherme Carmona. On the optimality of monetary trading [J]. Economic Theory, 2019,71(prepublish).*
*[3] Zhu Tong, Liao Chenyang, Zhou Ziyang, Li Xinyu, Zhang Qingfu. Fast popularity value calculation of virtual cryptocurrency trading stage based on machine learning [J]. Frontiers in Physics,2022.*
*[4] Zhou Liwen. Application of ARIMA model on prediction of China's corn market [J]. Journal of Physics: Conference Series, 2021, 1941(1).*
*[5] Ye Lu, Wang Zhizheng. Identification and prediction of bull and bear markets in China's stock market [J]. Statistics and Decision, 2021(20):161-165.*
*[6] JinShan Yang, ChenYue Zhao, HaoTong Yu, HeYang Chen. Use GBDT to predict the stock market [J]. Procedia Computer Science, 2020, 174.*
*[7] Yu Liwei, Ning Ling. Quantitative Investment strategy and empirical test of stock market [J]. Statistics and Decision, 2016(06):145-149.*
*[8] Jonathan B. Hill, Kaiji Motegi. A max-correlation white noise test for weakly dependent time series [J]. Econometric Theory, 2020.*