A study of gold and bitcoin trading strategies based on quantitative trading decision models

Yu Zhang*, #, Shulin Xu#, Liwei Hou#

College of Science, Northeast Electric Power University, Jilin, China, 132000

*Corresponding author: 18847829055@163.com

#These authors contributed equally.

Abstract. In the computer era, high frequency trading is one of the hottest topics in the financial investment market. In this paper, a quantitative trading decision analysis is conducted for a portfolio consisting of gold and bitcoin, applying gray prediction to get the future asset price, applying time series to predict the risk of purchasing the asset, and combining the two to make a reasonable buying and selling operation. To prove the accuracy of the prediction, the MSE indicator is used to measure the error of the gray prediction model, and the results show that the error is small. Finally this paper also proves that the model is well robust by varying the transaction cost.

Keywords: QTDM, Bitcoin, Gold, gray forecast, time series model.

1. Introduction

Market traders often buy and sell volatile assets with the goal of maximizing total returns. Trading relies on our practical experience and our response to market movements, that is, trading strategies. A good trading strategy can control our investment risks, help us correct loopholes in technical analysis, help us formulate orderly transactions and increase our profits. Quantitative trading originated in the stock market in the 1970s and developed rapidly after that and popularity, especially in the futures market, programmatic has gradually become mainstream. Some data show that programmatic futures trading in mature markets abroad has accounted for 70%-80% of the total trading volume, while it has just started in China. Trading solely on the basis of people's ideas is limited and unreasonable, while the natural accuracy and 100% execution rate of programmatic trading have brought advantages to its profitability [1].

Market traders buy and sell volatile assets frequently, with a goal to maximize their total return. To make the strategy more realistic and representative, we have a commission for each trade.We have selected two of the most popular currency assets in the market to buy and sell: gold and Bitcoin

The data we currently: the closing price of a troy ounce in USD on the specified date and the USD price of a single bitcoin on the specified date. And we set a rule that Bitcoin can be traded on a daily basis, but gold is only traded on days when the market is open. What we have to do is as follows:

• Develop a model that gives the best daily trading strategy based only on price data up to that day. Using this model and strategy, forecasts the value of an initial $1,000 investment from September 11, 2016 to September 10, 2021.

• Present evidence that this model provides the best strategy.

• Determine how sensitive the strategy is to transaction costs. Find out how transaction costs affect strategy and outcomes

2. Model Establishment: Gray Forecast Model

2.1. Model Establishment:Gray Forecast Model

Gray Forecast Model is a forecasting method that establishes a mathematical model and makes forecasts through a small amount of incomplete information is an effective tool to deal with the problem of small sample prediction[2]. By analyzing the data obtained from the survey, we believe that the prediction based on the 5-day sample data is the most reasonable. Part of the information in

the gray system is known, and the other part of the information is unknown. There is an uncertain relationship between the factors.

After the accumulation generation of the original data and the generation of weighted provisional value, we confirmed through data inspection that the generated data could establish the GM(1,1) [3]model for grey prediction.

The grey derivative of $x^{(1)}$ is defined as $d(k)=x^0(k)=x^1(k)-x^1(k\quad 1)$

Let $z^1(k)$ be the adjacent value of the sequence x^1 to generate a sequence

We get the formula $z^1(k)=ax^1(k)+(1-a)x^1(k-1)$

Therefore, the grey differential equation model of GM(1,1) is defined as $d(k)+az^1(k)=b$, where a is called the development coefficient, $z^1(k)$ is called the whitening background value, and b is called the grey action amount. The following equations are obtained:

$$\begin{cases} x^0(2)+az^1(2)=b \\ x^0(3)+az^1(3)=b \\ \cdots\cdots \\ x^0(n)+az^1(n)=b \end{cases} \tag{1}$$

Listed in the form of a matrix:

$$u=\begin{bmatrix} a \\ b \end{bmatrix}, Y=\begin{bmatrix} x^0(2) \\ x^0(3) \\ \cdots \\ x^0(n) \end{bmatrix}, B=\begin{bmatrix} -z^1(2) & 1 \\ -z^1(3) & 1 \\ \cdots & \\ -z^1(n) & 1 \end{bmatrix} \tag{2}$$

Then GM(1,1) can be expressed as Y=Bu, and the next step is to find the values of a and b.

Then we can use linear regression or $(B^TB)^{\ 1}B^TY$ (normal equation) to find the values of a and b according to the principle of least squares .

(5)Prediction

The corresponding whitening model is $\frac{dx^1(t)}{dt}+ax^1(t)=b$, and the solution for $x^1(t)$ is $x^1(t)=\left(x^0(1)-\frac{b}{a}\right)e^{-a(t-1)}+\frac{b}{a}$, Let $t+1=t$ have $x^1(t+1)=\left(x^0(1)-\frac{b}{a}\right)e^{-a}+\frac{b}{a}, k=1,2,3,...,n-1$

2.2. Model Establishment:Time series forecasting model

2.2.1 Time series forecasting model

Since in order to make the prediction model have good prediction effect even when the known data sequence is short, [N/2] is chosen as the maximum number of columns of the mean-generation function matrix in this paper[4]. Through the analysis of the element generation in the mean generating function matrix, it can be seen that the mean generating function matrix contains all the possibilities that there may be at least two cycles in the known sequence at the same time, and the elements in the same position in each cycle are averaged to eliminate the chance of individual data. This is the meaning of averaging in the mean generating function. After obtaining the mean generating function matrix, in order to predict the known sequence, it is necessary to generate the cycle extrapolation matrix F from the mean generating function matrix. Assuming that the length of the known sequence is N and the number of time points required for prediction is q, F is expressed as follows:

$$F = \begin{bmatrix} F_{[N/2]} \\ F_{\mathrm{q}} \end{bmatrix} \tag{3}$$

$$F_{[N/2]} = \begin{bmatrix} \bar{x}_1(1) & \bar{x}_2(1) & \ldots & \bar{x}_{\left[\frac{N}{2}\right]}(1) \\ \bar{x}_1(1) & \bar{x}_2(2) & \ldots & \bar{x}_{\left[\frac{N}{2}\right]}(2) \\ \bar{x}_1(1) & \bar{x}_2(1) & \ldots & \bar{x}_{\left[\frac{N}{2}\right]}(3) \\ \bar{x}_1(1) & \bar{x}_2(2) & \ldots & \bar{x}_{\left[\frac{N}{2}\right]}(4) \\ \ldots & \ldots & \ldots & \ldots \\ \bar{x}_1(1) & \overline{x}_2(i_l) & \ldots & \bar{x}_{\left[\frac{N}{2}\right]}\left(\left[\frac{N}{2}\right]\right) \end{bmatrix} \tag{4}$$

$$F_{\mathrm{q}} = \begin{bmatrix} \bar{x}_1(1) & \bar{x}_2(i_2) & \ldots & \bar{x}_{\left[\frac{N}{2}\right]}(1) \\ \bar{x}_1(1) & \bar{x}_2(i_1) & \ldots & \bar{x}_{\left[\frac{N}{2}\right]}(2) \\ \ldots & \ldots & \ldots & \ldots \\ \bar{x}_1(1) & \overline{x}_2(i) & \ldots & \bar{x}_{\left[\frac{N}{2}\right]}(q) \end{bmatrix} \tag{5}$$

$\bar{x}_2(i_1)$ means to take one of the $\bar{x}_2(1)$ and $\bar{x}_2(2)$, and $\bar{x}_2(i_2)$ mean the other one

It can be seen that the periodic extension matrix is essentially the averaging function matrix according to the length of each column vector period, which is periodically expanded to fill in zero elements. There are two reasons for this: On the one hand, the upper limit of the number of rows of the average generating function matrix is extended, so that the matrix F has the function of fitting and prediction; on the other hand, the sparse matrix H is transformed into a full matrix F, and then the sequence is simulated. Each periodic law (that is, the matrix contains at least two periods to no periodic law) can be reflected in the combined forecasting. Periodic extension matrix F $F_{[N/2]}$ and F_{q} are used for fitting and prediction respectively.

Before using the periodic extension matrix F for fitting and prediction, there is a further preparatory work, that is, the Gram-Schmidt orthogonalization of the F matrix. The main use of the prediction model is that due to the arbitrariness of the known data sequence, there may be correlations between the column vectors after generating the mean-generating function matrix and cannot be formed. Orthogonal basis, which then violates the original intention of the mean-generating function matrix, is bound to have a certain degree of influence on the prediction results.

The process of orthogonalization is as follows:

(1)Divide the periodic extension matrix F into [N/2] column vectors: $F^* = [f_1, f_2, \ldots f_{[N/2]}]$

(2)Take the column vector fi as the basic vector to orthogonalize the rest of the vectors to get [N/2] column vectors $f^*_1, f^*_2, \ldots f^*_{[N/2]}$

(3)The orthogonalized vectors form a new periodic extension matrix: $F^* = [f_1, f_2, \ldots f_{[N/2]}]$

$$\left.\begin{aligned} f^*_1 &= f_1 \\ f^*_2 &= f_2 - \frac{(f^*_1, f_2)}{(f^*_1, f^*_1)} f^*_1 \\ &\cdots \\ f^*_{[N/2]} &= f_{[N/2]} - \frac{(f^*_1, f_{[N/2]})}{(f^*_1, f^*_1)} f^*_1 - \frac{(f^*_2, f_{[N/2]})}{(f^*_2, f^*_2)} f^*_2 - \ldots - \frac{(f^*_{[N/2]-1}, f_{[N/2]})}{(f^*_{[N/2]-1}, f^*_{[N/2]-1})} f^*_{[N/2]-1} \end{aligned}\right\} \quad (6)$$

At this time, the orthogonal periodic extension matrix F^* can be used for fitting and prediction. The fitting and prediction formulas are as follows:

$$x(t) = \sum_{i=2}^{[N/2]} \varphi_i f_i(t) \qquad t \le N \quad (7)$$

$$x(N+q) = \sum_{i=2}^{[N/2]} \varphi_i f_i(N+q) \quad (8)$$

According to equation (6), the vector $\varphi_i (i = 2,...[N/2])$ is obtained by the least square method:

$$\phi = (F^{*T} F^*)^{-1} F^{*T} X' \quad (9)$$

Among them X' is the new sequence X obtained by standardizing the known data sequence:

$$x'(t) = \frac{x(t) - \bar{x}}{\sigma} \quad (10)$$

In the formula, $\bar{x}$ and σ are the average and standard deviation of the known sequence X, respectively. The reason for standardizing the known data sequence here is to make the known number sequence fluctuate around the zero value and be at the same value level.

Finally, let me explain the essential principle of using periodic extension matrix fitting and prediction. It can be seen from equation (10) and (11) that the use of periodic extension matrix fitting and prediction is actually to separate any kind of periodicity that may exist in the known sequence, and then give corresponding weights (such as periodic extension). The first column vector in the matrix indicates that each element in the known data sequence has an independent period. For the predicted value of the unknown sequence, the contribution of this period is the average value of all elements in the known data sequence and multiplied by this period. The weight of sex; and the last column of vector period is the largest, and the vector contains only two periods, then the contribution of the vector when predicting the nth value is the average value of each element in the previous periods with the same period position of the predicted value The meaning of the remaining column vectors can be deduced by analogy. Therefore, it can be seen that the method of predicting with the mean-generating function model contains the possibility of multiple periodic changes in the known sequence, and then the coefficient vector $\varphi = [\varphi_1, \varphi_2, ..., \varphi_l]$ is used as the size of the possibility of different periods. In balance, this method is reasonable and feasible.

2.2.2 ARMA time series model

Recording the development process of random events in the order of time constitutes a time series[5]. To observe and study the time series, to find the law of its change and development, and to predict its future trend is time series analysis.

(1) Model establishment

ARMA(p,q) model:

$$\begin{cases} x_t = \phi_0 + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \\ \phi_p \neq 0, \quad \theta_q \neq 0 \\ E(\varepsilon_t) = 0, \quad Var(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t \\ Ex_s \varepsilon_t = 0, \forall s < t \end{cases} \tag{11}$$

The model with the above structure is called p-order autoregressive model, denoted as ARMA(p,q) Stationary series modeling[6].

(2) Calculate the sample correlation coefficient:

Sample autocorrelation coefficient:

$$\hat{\rho}_k = \frac{\sum_{t=1}^{n-k}(x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2} \tag{12}$$

Sample partial autocorrelation coefficient:

$$\hat{\phi}_{kk} = \frac{\hat{D}_k}{\hat{D}} \tag{13}$$

(3) Model recognition:

Table 1 gives the basic principles of model identification, from which we can get the suitable models to use.

Table 1. The basic principle

$\hat{\rho}_k$	$\hat{\phi}_{kk}$	Choose a model
Tailing	P-order tailing	AR(p)
q-order tail	Tailing	MA(q)
Tailing	Tailing	ARMA(p,q)

(4)Approximate distribution of sample correlation coefficient:

Barlett: $\hat{\rho}_k \sim N(0, \frac{1}{n}) \ , n \to \infty$

Quenouille: $\hat{\phi}_{kk} \sim N(0, \frac{1}{n}) \ , n \to \infty$

(5)Sequence prediction:

Linear prediction function

$$x_t = \sum_{i=0}^{\infty} C_i x_{t-1-i} \tag{14}$$

Principle of Minimal Forecasting Variance

$$Var_{\hat{x}_t(l)}\left[e_t(l)\right] = \min\left\{Var\left[e_t(l)\right]\right\} \tag{15}$$

2.2.3 Analysis and Evaluation of results

The time series is used to predict the rise of the two currencies, score the bitcoin buy as shown in the Figure 1 and the bitcoin score comparison value according to Figure 2, and set the scoring criteria. The reasonableness of the scoring criteria is judged.

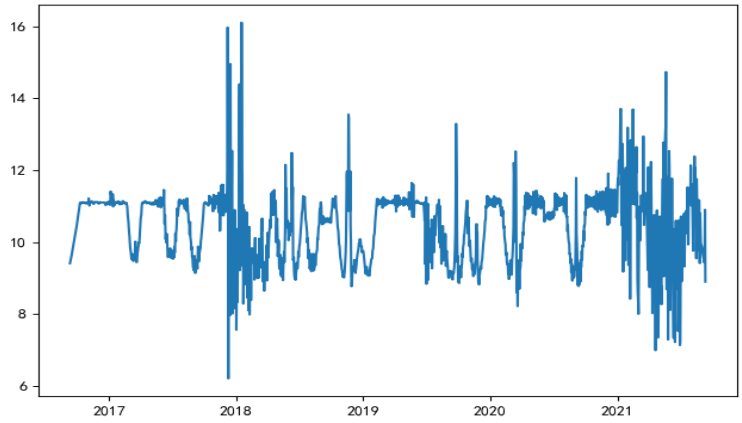


Figure 1. Bitcoin Buy Score

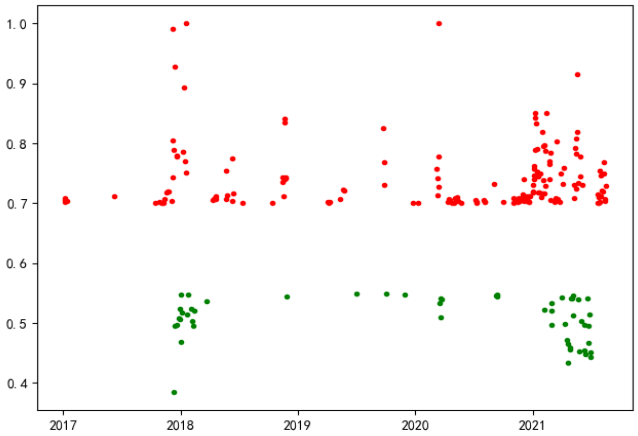


Figure 2. Bitcoin Score Comparison Value

Predict the increase of the two currencies through time series, score the buy, and set the scoring standard:

$$A = \frac{g \times 10 + B \times 5 - C + 1}{F} \tag{16}$$

Judging the rationality of the scoring criteria:

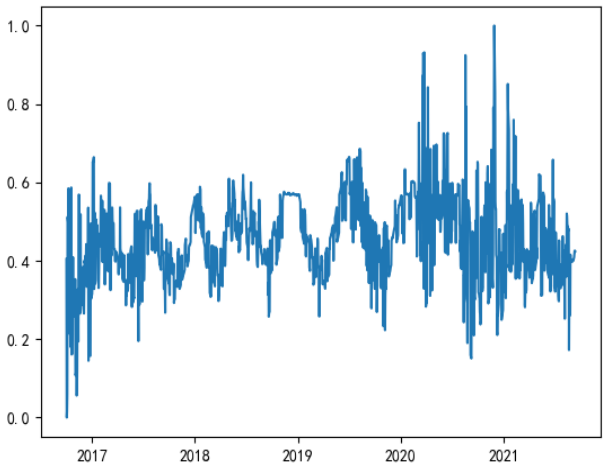


Figure 3. Gold Buy Score

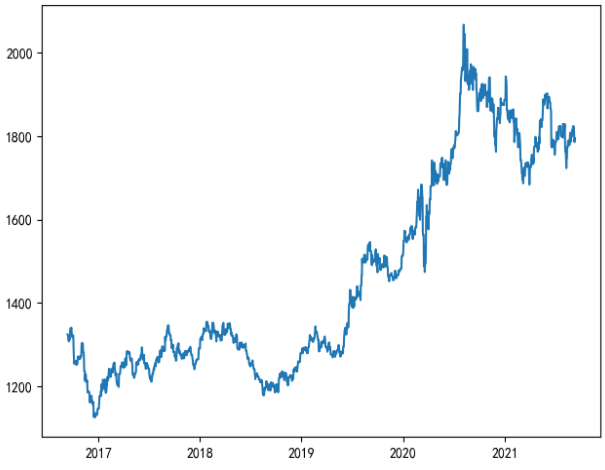


Figure 4. Gold Price Trend

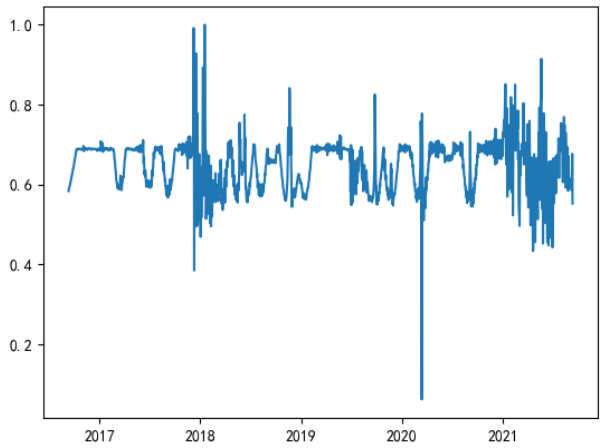


Figure 5. Bitcoin Buy Score

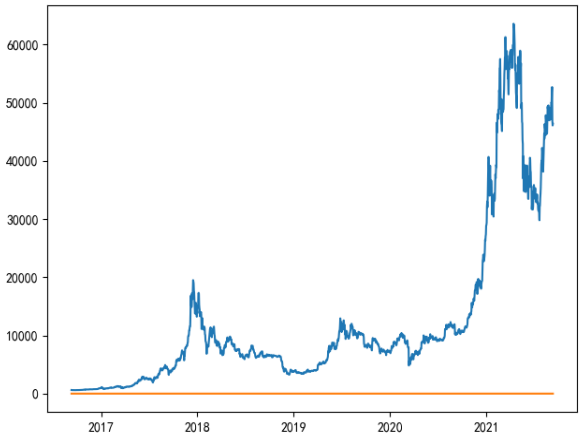


Figure 6. Bitcoin Price Trend

Comparing the two figures 4 and 6, when the price trend is declining in the same time period, the buy score is increasing, which shows that the scoring standard is set reasonably. (During data processing, if a certain value is much larger than other values, you should choose to discard the value or assign the value to data that is not much different from other values)[7].

The buying and selling rules are formulated according to the buying score chart of gold and the buying score chart of Bitcoin. As can be seen from the figure 3and 5, the score of gold accounts for the majority in the period from 0.3 to 0.6, and it is in the middle of the whole score, which is not suitable for buying. If you want to buy or sell, choose to buy gold when the score is greater than 0.6, and sell gold when the score is less than 0.3; when Bitcoin is in the majority between 0.7 and 0.55, choose to buy Bitcoin when the score is greater than 0.7 Sell bitcoins , buy bitcoins when the score is less than 0.55.

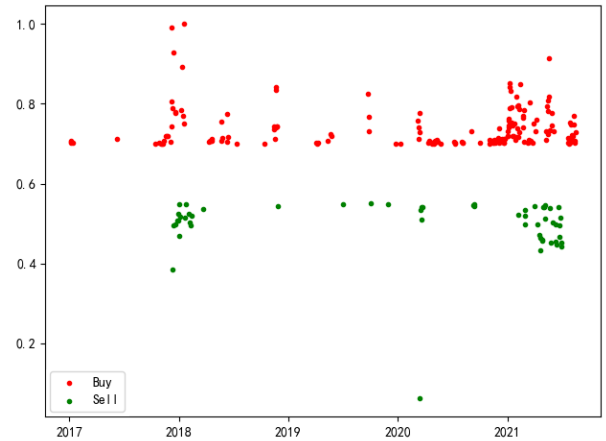


Figure 7. Bitcoin Buy/Sell

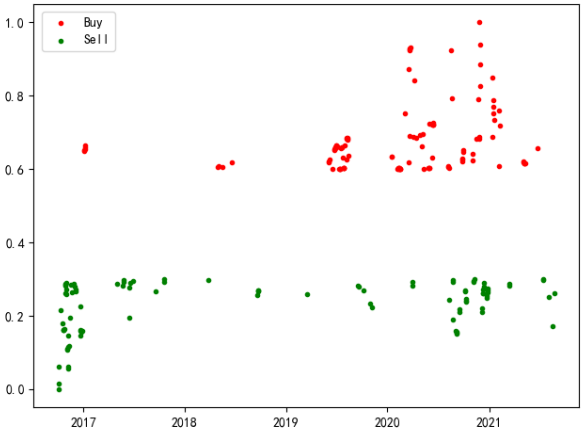


Figure 8. Gold Buy/Sell

Figure 7, 8 is the distribution of bitcoin and gold purchase and sale scores, red dots are purchase scores, green is sale scores, through the figure 7, 8 can be derived from the following purchase rules.

Buying and selling rules: 1. Determine whether it is a gold trading day 2. When bitcoin and gold can be bought at the same time, if the gold buy score -0.6> (bitcoin buy score -0.7) X2 (gold score buy standard than bitcoin) The currency is low, and there is more room for growth) 3. Buying quota = current cash quota X buying score X (1-handling fee) / current price 4. Selling quota = holding share * (1-score + selling standard)

Note: Pay attention to handling fees when calculating cash.

This calculates our predicted value of 226648.4205.

2.3. Model Establishment:Measure model error with MAE indicator[8]

Suppose the predicted value: $\hat{y} = \{\hat{y}_1, \hat{y}_{2,\ldots,}\hat{y}_n\}$, the true value: $y = \{y_1, y_{2,\ldots,}y_n\}$

Mean Absolute Error, $MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}i - yi|$, range $[0, +\infty)$, is equal to 0 when the predicted value exactly matches the true value, that is, a perfect model; the larger the error, the larger the value.

2.4. Model Establishment:Durbin-Watson test[9]

Durbin-Watson test, also known as DW test, is used to test the first-order autocorrelation of residuals in regression analysis (especially for time series data).

Assuming that the residual is e_t, the correlation equation of each residual is $e_{t} = \rho e_{t-1} + v_t$, and the null hypothesis of the test is $\rho = 0$: the alternate hypothesis is $\rho \neq 0$: the test statistic: $d = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}$ is approximately equal to 2(1-ρ) (this proof process requires the molecular The formula is expanded, and then the autocorrelation coefficient between e_t and e_{t-1} is used), so the closer the statistic value is to 2, the better. Generally, between 1 and 3, there is no problem. If it is less than 1, it means that the residual has autocorrelation.

3. Analysis and Evaluation of result

The MAE index test is carried out on the predicted value obtained by the gray prediction, and the obtained data is 0.3428, and the accuracy of the model can be determined by the definition of the MAE index.

Based on the original data, the predicted value for the next day is obtained after prediction, and the comparison between the predicted and actual values can be drawn from the same table, as in Figures 9 and 10.

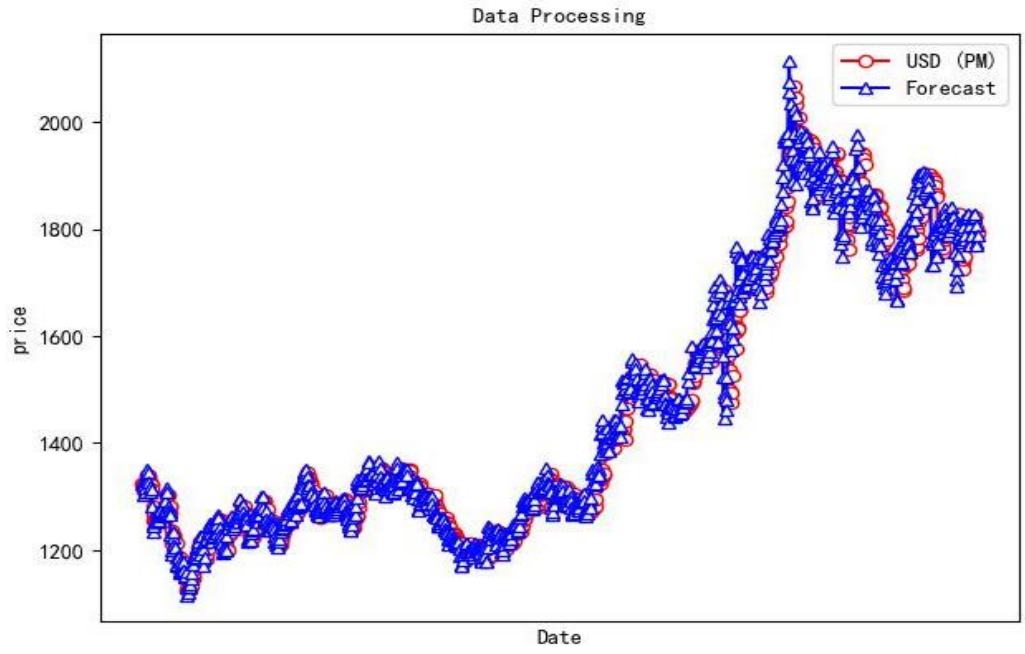


Figure 9. Gold Predictions vs. Real Data

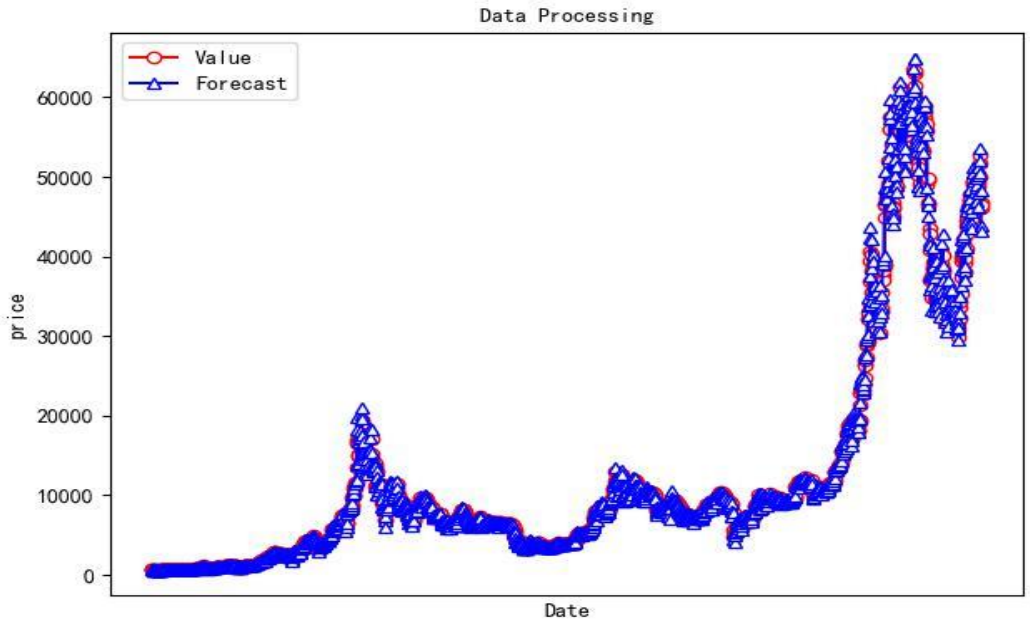


Figure 10. Bitcoin Predictions vs. Real Data

It can be seen from the figure9 and figure10 that the predicted value and the real data are basically in the same position, and it can be seen that the predicted value obtained by the gray prediction is more real.

Residual Analysis: residual is the difference between the fitted value of the model and the actual observed value[10]. The residuals should look like white noise. If it is a normal distribution, it will be a straight line, that is, the red line. The result is roughly in line with white noise. We do the residual analysis graph as shown in the figure 11.

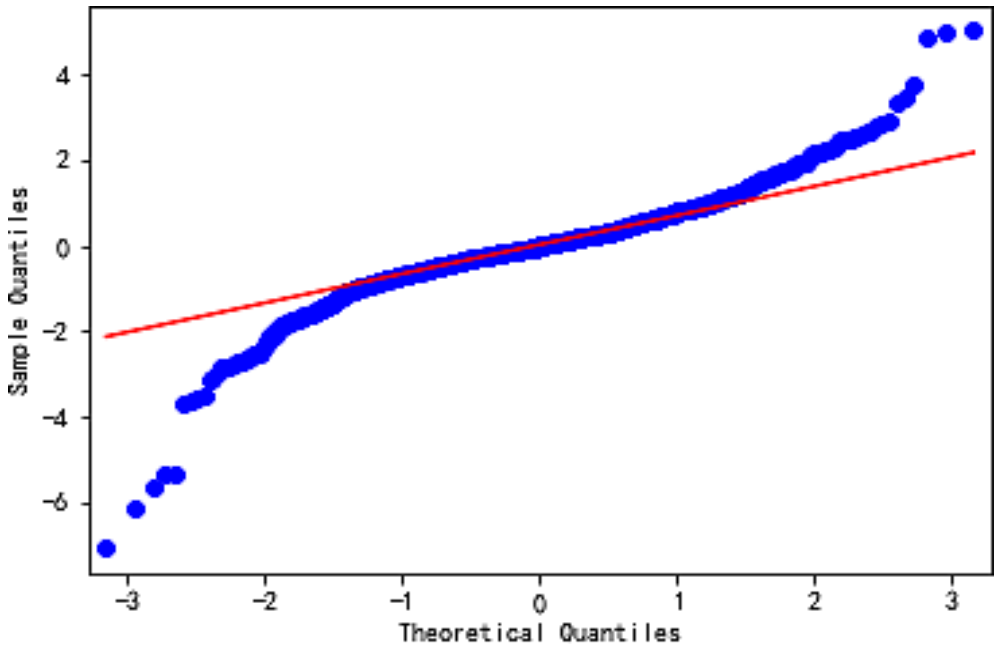


Figure 11. Residual Analysis Chart

After the Durbin-Watson test, the data is 2.0179203733750732, indicating that there is no problem.

We believe that the commission is increased by 0.01 to 0.05, and the calculation of the final total assets is shown in Table 2 below. It can be concluded that when the commission increases, the total assets show a decreasing trend. (It should be noted that the total assets at this time are calculated by the actual value, excluding the influence of the inaccurate predicted value on other topics.)

Table 2. Total assets

Gold commission	Bitcoin commission	Total assets
0.01	0.01	270048.0513
0.01	0.02	252424.6982
0.01	0.03	235802.0964
0.01	0.04	220131.9419
0.01	0.05	205367.8523
0.02	0.01	205701.0632
0.02	0.02	192278.4709
0.02	0.03	179618.0091
0.02	0.04	167682.8948
0.02	0.05	156437.8085
0.03	0.01	156346.5104
0.03	0.02	146145.606
0.03	0.03	136523.8438
0.03	0.04	127453.2752
0.03	0.05	118907.0635
0.04	0.01	118571.5683
0.04	0.02	110836.2201
0.04	0.03	103539.9868
0.04	0.04	96661.67971
0.04	0.05	90180.95259
0.05	0.01	89721.76169
0.05	0.02	83869.22618
0.05	0.03	78348.88522
0.05	0.04	73144.71079
0.05	0.05	68241.31292

4. Conclusions

The essence is to build a quantitative trading decision model (referred to as QTDM). Here we can divide it into two steps to build QTDM. The first is to solve the decisions that traders need to make on the day (it should be noted that the decisions of the day are essentially for the profit and loss of the next day). We use the data of the previous days to predict the future data, apply the gray forecast to calculate the predicted value, time series calculate the risk, reasonably score the steps of buying or selling, choose the best strategy, and combine the two parts to determine the stock price. The profit and loss situation is predicted, so that the model has good credibility for short-term stock price prediction.

It is proved that the model proposes the best strategy, which can carry out the accuracy test and analysis of the predicted results, that is, give the indicators to demonstrate the rationality of the decision. We use the MSE indicator to measure the model error detection of the gray forecast. If the error of the constructed model is larger, the result of the decision will be worse, because there is a huge difference between the actual profit and the simulated profit. Applying the time series, the ARMA model was used to predict the increase in the comparison of the three models.

It is necessary to demonstrate the impact of transaction costs on the results. We will change the commission, choose different percentages, use non-predictive data (the original data given to us) for calculation, present the data visually, and list the percentages under different percentages. result image.

References

[1] Song Chengguo; Research on the history and development of China's futures market [D]; Soochow University; 2010

[2] Zhang Peng. Research on key techniques of grey forecasting for small sample time series [D]. University of Electronic Science and Technology,2020.

[3] Wu, Chuanxu. Application of gray model GM(1,1) in price index forecasting [J]. Journal of Shaanxi University of Technology, 1998 03 44-47.

[4] Wei Shengsheng,Cao Hongxu,Wei Fengying. Test of homogeneous function model forecasts [J]. Meteorology,1998 05 43-47.

[5] Jiang Le. Research and application of stock price analysis based on time series [D]. Dalian University of Technology,2015.

[6] Zhang Tingting. Time series mining based on ARMA model [D]. Hefei University of Technology,2013.

[7] Hao Shuang,Li Guoliang,Feng Jianhua,Wang Ning. A review of structured data cleaning techniques [J]. Journal of Tsinghua University (Natural Science Edition),2018 58 12 3-16.

[8] Jia Mingzhu. Research and application of time series analysis method based on machine learning [D]. Xi'an University of Science and Technology,2020.

[9] Liu Ming,Wang Yongyu. Discussion on the application of Durbin-Watson autocorrelation test [J]. Quantitative Economic and Technical Economics Research,2014 31 06 154-161.

[10] Li Xin. Research on portfolio countermeasures to eliminate systematic risk based on the problem of stock price volatility [D]. Tianjin University of Technology, 2012.