Problem Chosen	2022 MCM/ICM Summary Sheet	Team Control Number
C		2204128

Sailing through the Thunderstorm: Trading Strategy Model based on ARIMA and Linear Programming

Summary

Bitcoin, the new favorite of many investors, has taken the international trading market by storm recently with its amazingly high returns; meanwhile, Gold sits firmly at the top of the precious metals trading pile.However, Bitcoin and Gold both have high risk rates.The purpose of this report is to build a price prediction and investment planning model to achieve: short-term price trend prediction within a certain error range; investment planning; and investment risk reduction. We have built three models: Model 1: Gold and Bitcoin Value Forecasting Model; Model 2: Daily Trading Strategy Planning Model; and Model 3: Investment Risk Assessment Model.

For model 1, the time series of the provided data were first analyzed for change trends and verified for smoothness. We performed first-order difference processing on the data to use ARIMA forecasting model. Next, due to the trailing phenomenon in both ACF and PACF images of the data (Figure 5), the ARIMA parameters could not be determined directly, and we applied AIC and BIC to finally determine the models for Bitcoin and Gold forecasting as ARIMA (2,1,2) and ARIMA (0,1,0). Finally, we trained the prediction models separately and predicted the prices for the next 15 days for each trading day with 95% prediction accuracy. Due to the low amount of data in the early stage, we introduced a Grey-forecasting model for the initial forecasts to reduce the ARIMA forecast bias due to insufficient data.

For Model 2, with prediction data from Model 1. In order to simulate the real market situation of Bitcoin and Gold buyers, we build a linear programming model. The linear programming model uses the sum of the possible trade assets acquired within 15 days including and after that day as the objective function for planning, and the first of the 15 strategy solutions is taken as the strategy for that day

For model 3, it is considered that there is a large trade risk in the trade market when Bitcoin and Gold prices are volatile, which may also affect investors' decisions. Therefore, we established 3 risk assessment indicators and used hierarchical analysis to evaluate the importance of these 3 indicators and finally derived their weights separately. After that, the trade risks were calculated for the dates with trade transactions (purchase or sale) based on the derived weights, respectively, and then the risks were ranked. The trade dates with higher trade risk were singled out and the trading strategy for that day was adjusted, and the method we used was to reduce the number of trades for that day to 80% of its size.

After predicting and simulating the original training set, an investor with an initial capital of $1,000, with 5 years of buying and selling, ended up holding more than $160,000 in Bitcoin, Gold & Cash (market capitalization at the time), achieving a 16,000% return, while fully following the investment strategy proposed in this model.

Finally, the model was tested for sensitivity, and when the initial distributions of Bitcoin and Gold prices were randomly generated by a uniform random distribution, the final convergence distribution of the model did not differ much as shown in Figure 13.

Keywords: ARIMA; Time Series Analysis; Grey Prediction; Linear Programming;Analytic Hierarchy Process.

Contents

1 Introduction

1.1 Problem Background

Stocks are playing a more and more important role in the financial field. Not only does the stock market establish a close bridge between financiers and investors, but also plays an important role in allocating resources. The trading volume and price of stocks often fluctuate, showing the characteristics of instability. Market traders can flexibly buy and sell stocks according to their own needs and the actual changes of the market.

The daily prices of gold and bitcoin are constantly changing. Experienced market traders will be accompanied by a return commission every time they buy and sell. In order to maximize returns, market traders often need to make careful decisions about their plans to buy and sell gold and bitcoin.

1.2 Restatement of the Problem

For this problem, we have learned the value of gold and bitcoin every day from November 9, 2016 to October 9, 2021, and the cost of each purchase and sale of bitcoin and gold is 1% and 2% of the trade amount respectively. Traders asked us to establish a reasonable mathematical model, using only the prices of gold and bitcoin in the past five years, to help traders make which decision to buy, hold or sell stocks.

- Establish a mathematical model to make the best strategy for traders every day based on the price data up to that day. In addition, calculate how much the initial $1000 investment will be worth on October 9, 2021 using this model.
- Proved that the established mathematical model can provide the best trade strategy.
- Calculate the sensitivity of transaction price to trade strategy, and analyze how transaction price affects trade strategy and results.
- Write to traders, to introduce the established model, provide trade strategies and summarize the analysis results.

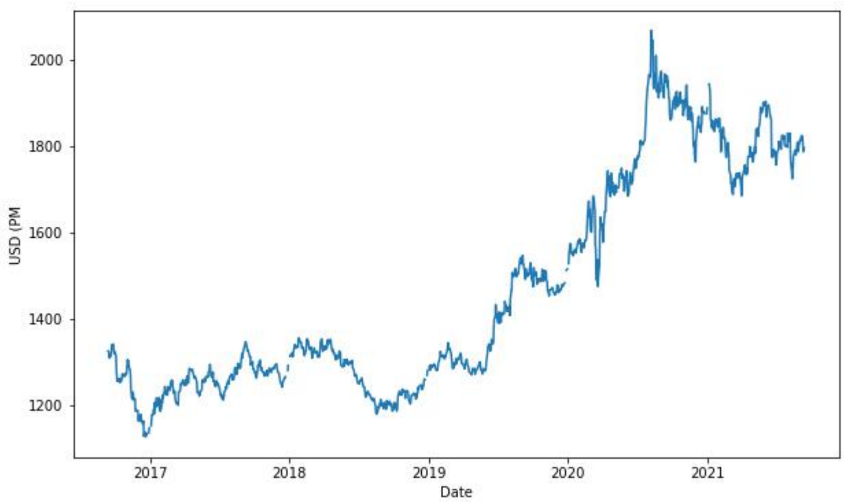


Figure 1: Daily price of gold (USD per troy ounce)

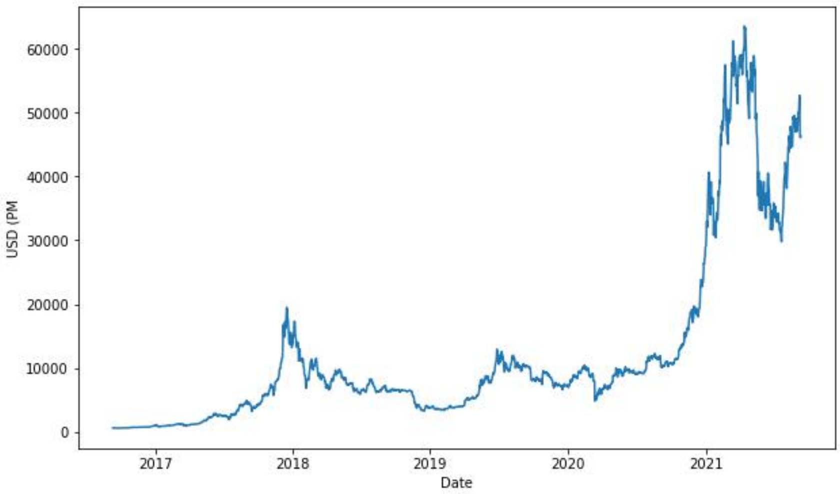


Figure 2: Daily price of bitcoin (USD per bitcoin)

1.3 Our Work

Aiming to build reasonable models that can provide the best trade strategies, we have carried out the following work:

1. **Based on ARIMA, a gold and bitcoin value forecasting model is developed.** By using five years of gold and bitcoin value data, time series analysis was employed to predict the future bitcoin and gold value on that day.
2. **Based on linear programming, a daily trading strategy planning model is developed.** Using the forecasted data, the daily trading strategies are planned using the linear programming model.
3. **Based on analytic hierarchy process, an investment risk evaluation model is established.** Minimize investment risk and optimize the model. Under certain objective conditions, the strategies to deal with risks are discussed.
4. **Verify the model and error analysis.** The predicted data are compared with the actual data, and the analysis of variance is carried out to prove the reliability of the prediction model. We can make a reasonable trade strategy for the characteristic data to prove the reliability of the planning model.
5. **Sensitivity analysis.** Analyze the sensitivity of transaction cost to trade strategy.

Intuitively, our workflow can be presented in the following diagram:

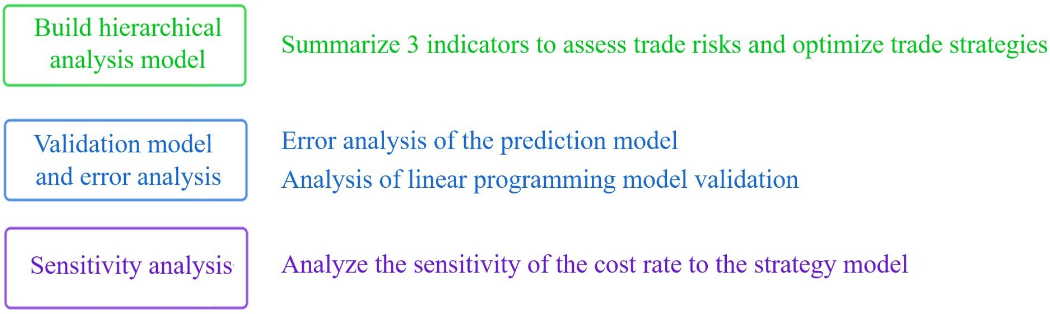


Figure 3: Flow Plot of Our Work

2 General Assumptions

Since the actual problem contains many complex factors, in order to streamline the problem, we have made the following assumptions, all of which are given with corresponding explanations.

Assumption 1: **Bitcoin and gold values are not affected by major international socio-political and economic events, such as world wars, economic crise sand systematic risk.**
Explanation: World wars and economic crises have a very large impact on the value of bitcoin and gold, causing the value of bitcoin and gold to rise and fall extremely dramatically while this model and indeed all currently known economic models cannot predict non-rational sudden changes.

Assumption 2: **Market traders are economic men.**
Explanation: Economic man imply that individuals act and make decisions in full accordance with the principle of revenue maximization, e.g., consumers fully pursue utility maximization and manufacturers fully pursue profit maximization. A economic man has complete knowledge about his environment and has stable and well-organized preferences. All economic theories are based on the assumption of economic man. If this condition is not satisfied, several models in this paper and the economic theories used will have large errors.

Assumption 3: **All the recent economic theories mentioned in this model are accurate.**
Explanation: If all the recent economic theories mentioned in the model are wrong, there will be underlying problems in the analysis and optimization of the paper, including risk estimation, and the whole model of the paper will be subject to unforeseen errors. Therefore, the model should be validated on the basis of the present assumptions.

*In addition, in order to simplify individual problems, other assumptions are also made in this paper. These assumptions will be discussed in the corresponding position of this article.

3 Notations

Some important mathematical notations used in this paper are listed in Table 1.

Table 1: Notations used in this paper

Symbol	Description
$z_i^{(1)}$	Number of bitcoins owned on day i
$z_i^{(2)}$	Number of gold owned on day i
$x_i^{(1)}$	Number of bitcoin transactions on day i
$x_i^{(2)}$	Number of gold transactions on day i
b_i	Bitcoin price predicted on day i
g_i	Gold price predicted on day i
$a_1\%$	Transaction cost rate of bitcoin
$a_2\%$	Transaction cost rate of gold
c_i	Number of cash owned on day i
α_i	Coefficient of whether gold is open on day i

* Some variables are not listed here and will be discussed in the corresponding position.

4 Model I: Gold and bitcoin value prediction model

In order to get the best strategy for each day's trade, we need to predict the future stock value on that day.Time series analysis models can use the characteristics of the time of an event in the past period to predict the characteristics of that event in the future period. One of them is Autoregressive Integrated Moving Average Model (ARIMA) that helps us to predict the stock price in the coming days based on the price data as of that day.[1] [2] So we take the help of ARIMA model and python programming to make the prediction.

ARIMA (p,d,q), AR is autoregressive, p is the autoregressive term; MA is moving average, q is the number of moving average terms, and d is the number of differences made when the time series becomes smooth.

4.1 Difference processing

Initially, predict the value of bitcoin.The time series that can be applied to the ARIMA model for analysis and prediction must satisfy the condition of a smooth non-white noise series. According to the Daily price of bitcoin (USD per bitcoin)(2) we can see that the overall value of Bitcoin stock is highly variable and not smooth, so the time series should be differenced first and then tested for smoothness until it is smooth.

Then, after differencing the data once, a square root test is performed, which proves that it has been a smooth time series. Therefore $d = 1$.

$$\begin{cases} y_t + y_{t-1} = (\alpha - 1)y_{t-1} + \varepsilon_t \\ \triangle y_t = y_t - y_{t-1} = (\alpha - 1)y_{t-1} + \varepsilon_t \\ \triangle(\triangle y_t) = \triangle^2 y_t = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) \end{cases}$$

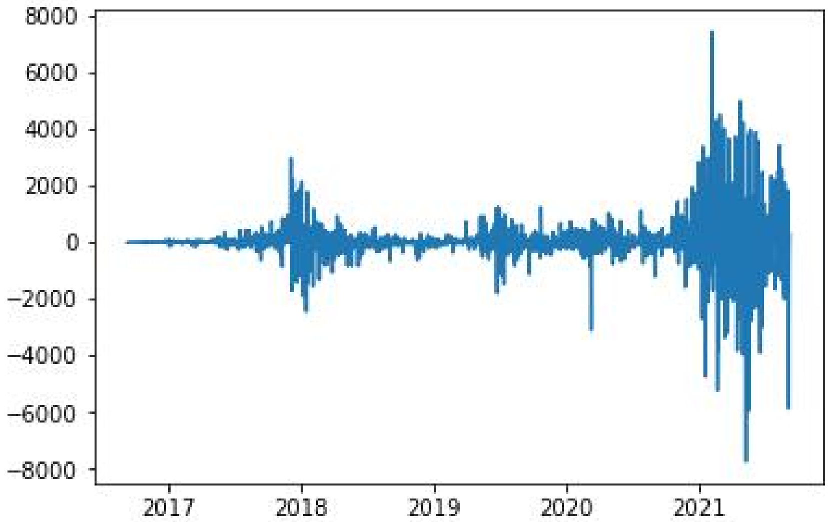
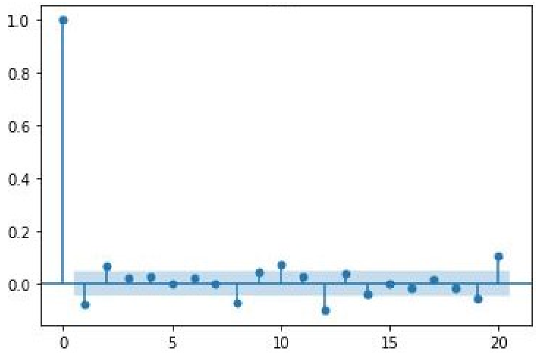


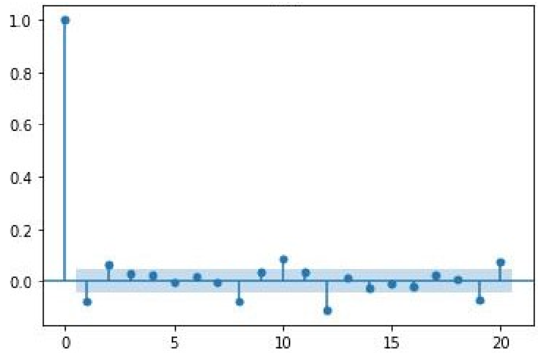
Figure 4: Difference plot

4.2 Selection of model according to ACF/PACF and parametric model

Since the p-value is small, it is able to prove that the data are correlated by q-validation. Since both ACF and PACF plots have trailing phenomenon, the order p and q cannot be determined by this.



(a) ACF plot



(b) PACF plot

Figure 5: ACF & PACF plot

Therefore, the AIC and BIC minimum criteria are used to determine the order p and q. AIC and BIC are used to avoid the overfitting problem by adding a penalty term for model complexity, and usually both tend to choose models with fewer parameters, and the penalty term of BIC is larger than that of AIC, taking into account the sample size, which can effectively prevent excessive model complexity caused by high model accuracy when the sample size is too large. However, in many cases, the model with the smallest AIC does not mean that the BIC will also be the smallest.

$$\begin{cases} AIC = 2k - 2ln(L) \\ BIC = kln(n) - 2ln(L) \end{cases}$$

k is the number of model parameters, n is the number of samples, and L is the likelihood function. The results obtained $p = 2$ and $q = 2$.

4.3 Validating the model: residual analysis

The model is trained using ARIMA, and the obtained results are subjected to residual tests. The validation of the model is also the independence test of the noisy series. If the residuals are white noise series, it means that the model passes the test. The results show that the model can pass the white noise test. Finally, according to the Q-Q plot(6) test, it is found that the black dots basically coincide with the blue line. the residuals of the ARIMA model are normally distributed with zero and constant variance.

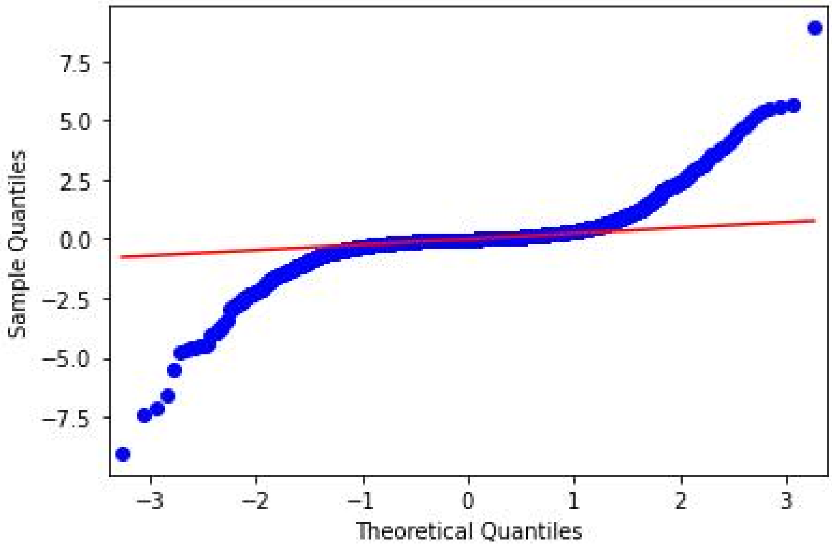


Figure 6: Q-Q plot

According to the Prediction plot(7), using the model to compare the predicted data with the original data found that the accuracy is high, indicating that the model is relatively successful in establishing. Therefore, we choose the ARIMA (2,1,2) model for value prediction of bitcoin.

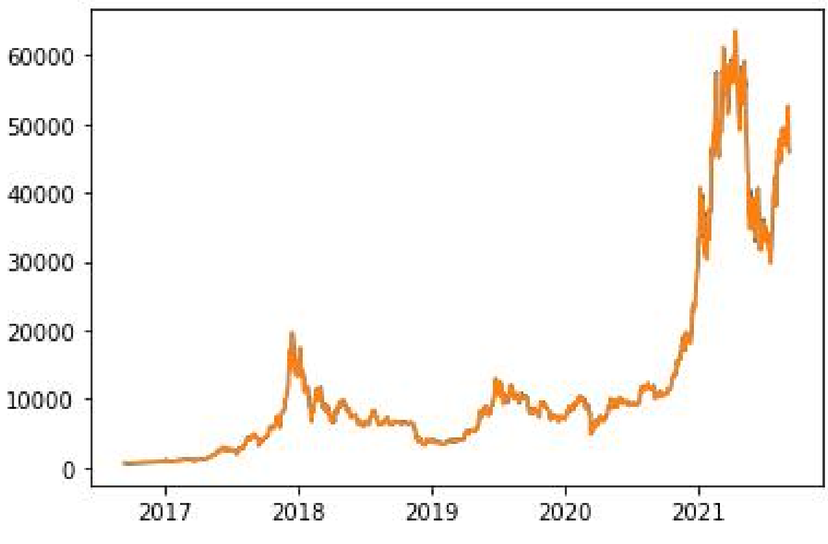


Figure 7: Prediction plot

Similarly, we apply the same method to forecast the value of gold and choose the ARIMA (0,1,0) model for forecasting.

5 Model II: Daily trading strategy planning model

To further decide on the best strategy for daily trade, we need to plan the daily trade strategy using a linear programming model.

To facilitate the calculation, the daily zero point is used as the daily trade decision time. When making the decision, the long term future should be considered as much as possible for the maximum benefit. However, due to the accuracy limitation of ARIMA forecasts, the further the forecast is from the day, the larger the error value is, so we decided that the daily trade decision should be influenced by the maximum value of total assets based on ARIMA forecasts for the next 15 days.

5.1 Basic quantity

The relationship between the quantities is calculated by the following equation:

$$\begin{cases} z_i^{(1)} = z_{i-1}^{(1)} + x_{i-1}^{(1)} \\ z_i^{(2)} = z_{i-1}^{(2)} + x_{i-1}^{(2)} \\ c_i = c_{i-1} - [(|x_{i-1}^{(1)}|b_{i-1}a_1\% + x_{i-1}^{(1)}b_{i-1}) + (\alpha_{i-1}|x_{i-1}^{(2)}|g_{i-1}a_2\% + x_{i-1}^{(2)}g_{i-1}\alpha_{i-1})] \end{cases} \tag{1}$$

With the gold only available for trading during market opening hours, it is controlled with coefficient α_k for each trade.

5.2 Objective function

In order to maximize the total asset value after 15 days, we sum the daily asset values for 15 days to obtain the objective function:

$$max\sum_{k=i}^{i+15}\{[(z_k^{(1)}+x_k^{(1)})b_k-|x_k^{(1)}|b_ka_1\%-x_k^{(1)}b_k]+[(z_k^{(2)}+\alpha_kx_k^{(2)})g_k-\alpha_k|x_k^{(2)}|g_ka_2\%-\alpha_kx_k^{(2)}]+C_k\} \tag{2}$$

$(z_k^{(1)} + x_k^{(1)})b_k$ and $(z_k^{(2)} + \alpha_k x_k^{(2)})g_k$ are the predicted values of bitcoin and gold after daily transactions for fifteen days, $|x_k^{(1)}|b_ka_1\%$ and $\alpha_k|x_k^{(2)}|g_ka_2\%$ are the costs consumed by daily transactions of bitcoin and gold, and $x_k^{(1)}b_k$ and $\alpha_k x_k^{(2)}$ are the impact on cash after daily transactions of bitcoin and gold.

5.3 Constraints

Since the amount of bitcoin and gold held after daily trades should be no less than zero, the first constraint is achieved by the following equation:

$$\begin{cases} z_i^{(1)} + x_i^{(1)} \geq 0 \\ z_i^{(2)} + x_i^{(2)} \geq 0 \end{cases}$$

Meanwhile, the amount of cash held after the daily trade should also be no less than zero, and the second constraint is achieved by the following equation:

$$c_i - [(|x_i^{(1)}|b_ia_1\% + x_i^{(1)}b_i) + (\alpha_i|x_i^{(2)}|g_ia_2\% + x_i^{(2)}g_i\alpha_i)] \geq 0$$

5.4 Establish planning model

objective function: $max \sum_{k=i}^{i+15}\{[(z_k^{(1)}+x_k^{(1)})b_k-|x_k^{(1)}|b_k a_1\%-x_k^{(1)}b_k]+[(z_k^{(2)}+\alpha_k x_k^{(2)})g_k-\alpha_k|x_k^{(2)}|g_k a_2\%-\alpha_k x_k^{(2)}]+C_k\}$

$$s.t.\begin{cases} z_i^{(1)}+x_i^{(1)}\geq 0 \\ z_i^{(2)}+x_i^{(2)}\geq 0 \\ c_i-[(|x_i^{(1)}|b_i a_1\%+x_i^{(1)}b_i)+(\alpha_i|x_i^{(2)}|g_i a_2\%+x_i^{(2)}g_i\alpha_i)]\geq 0 \end{cases} \tag{3}$$

6 Model III: Investment risk assessment model

Due to the uncertainty of the market and the possible error of the model, investors' investment is inevitably risky. In order to avoid the risk as much as possible, after comparing similar evaluation methods, this report finally adopts the hierarchical analysis method to establish a risk evaluation model for investment risk assessment and achieve the minimum investment risk.[6]

6.1 Establish hierarchical structure model

According to the nature of trade risks and the optimal decision objectives to be achieved, this report decomposes the problem into 3 different components, [9]and according to the association of these 3 components and their affiliation, they are cohesively combined at different levels to hierarchize the risk assessment problem, and finally form a multi-level analysis structure model. Among them, these 3 indicators are:

1. **Indicator 1: Ratio of current value to historical one-year average value.**

   Compare this indicator to 1. The greater the difference from 1, the greater the difference between the current value and the historical value, the greater the fluctuation in value, and the greater the trade risk.

2. **Indicator 2: Current value.**

   The indicator is percentile divided and compared with the same side data of the indicator. All trade value data is taken as a reference, and the larger the gap, the greater the value fluctuation and the greater the trade risk.

3. **Indicator 3: Forecasted growth rate in the next 15 days.**

   Taking the indicator as an absolute value, the higher the absolute value of the growth rate, the greater the value fluctuation and the greater the trade risk.

According to the relationship between them, the decision objectives, decision criteria and decision schemes are divided into the highest level, middle level and lowest level, and the hierarchical structure diagram(8) is drawn:

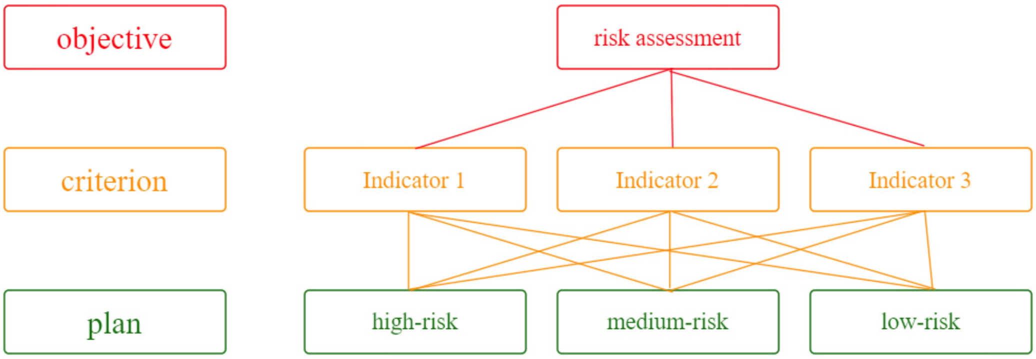


Figure 8: Hierarchical structure model

6.2 Construct judgment matrix

Analytic hierarchy process requires that the judgment value R obtained after pairwise comparison of each index is used to construct the judgment matrix, and the comparison scale range of pairwise comparison is $R \in [1, 9]$; When an indicator is equally important, slightly important, obviously important, most important and absolutely important relative to another indicator, R is taken as 1, 3, 5, 7 and 9 respectively. When the comparison results of the two indicators are in the adjacent judgment, take the middle value between the two numbers, i.e. 2, 4, 6 and 8.

This paper holds that the importance of indictor 1 to indictor 2 is between equally important and slightly important, so $R = 2$; Indicator 3 is slightly important for indicator 1, so $R = 3$; Indicator 3 is obviously important to indicator 2, so $R = 5$.

	Indicator 1	Indicator 2	Indicator 3
Indicator 1	1	2	1/3
Indicator 2	1/2	1	1/5
Indicator 3	3	5	1

Figure 9: Judgment matrix

6.3 Hierarchical single ranking and its consistency test

The elements of W are the ranking weights of the same hierarchical factor for the relative importance of a factor in the previous hierarchical factor, and this process is called hierarchical single ranking. That whether the hierarchical single ranking can be confirmed, a consistency test is needed, and the so-called consistency test is to determine the allowable range of inconsistency for the pairwise comparison matrix.

Let matrix A (9) be a square matrix of order n. Normalize each column vector of A to obtain B. Then sum up B by rows to obtain C. Normalize C to obtain W,where

W_i is the gold wire eigenvector. Finally, $\lambda_{max} = \frac{1}{n}\sum_{i=1}^{n}\frac{(AW)_i}{W_i}$ is calculated as the approximation of the maximum eigenvalue. The final calculation gives $\lambda_{max} = 3.0037$.

$CI = \frac{\lambda_{max} - n}{n-1}$ is the indicator that defines consistency.The smaller the value of CI, the better the consistency. The $CI = 0.0018$ is obtained after calculation.

n	1	2	3	4	5	6	7	8	9	10	11
RI	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49	1.51

Figure 10: Random consistency index RI

To measure the magnitude of CI, the random consistency index RI is usually introduced.

According to the figure(10), since $n = 3$, it can be found that $RI = 0.58$. Therefore $CR = \frac{CI}{RI} = 0.0032 < 0.1$. Eventually this evaluation model passed the consistency test.

6.4 Calculate the index score of trade date and optimize the model

Given the weight $W = [0.2299; 0.1222; 0.6479]$, this report uses the three indicators to calculate and sort the final scores in ascending order.The top ten scores are intercepted below:

Table 2: Top 10 trade risk scores

Bitcoin Sorting	Gold Sorting
0.4869575	0.1454333
0.4789235	0.1431969
0.4672042	0.0848206
0.4619203	0.0665937
0.4600188	0.0627274
0.4588271	0.0564713
0.4562471	0.055744
0.4410997	0.0536866
0.4128295	0.0512171
0.3981310	0.0509693

In this report, the blue data (i.e. No.1 and No.2) are identified as high-risk trade, and the number of transactions is adjusted by 80% to avoid high-risk transactions and prevent large losses in trade.

6.5 Result

After reducing the corresponding transaction volume, we solve the first question again and get the final result, as shown in the following table:

Cash	Bitcoin transaction	Gold transaction	Bitcoin Holdings	Gold Holdings Holdings	Date
1000	1.57708	0	0	0	2016/9/11
2.21E-04	-1.275853	0.5700491	1.57708	0	2016/9/12
0	-0.3012274	0.1349658	0.3012274	0.5700491	2016/9/13
0	1.80E-06	-9.52E-07	0	0.7050149	2016/10/26
0	-1.80E-06	9.38E-07	1.80E-06	0.7050139	2016/10/27
0	1.116772	-0.7050149	0	0.7050149	2016/11/28
0	7.27E-07	-5.01E-07	0	0.7209349	2016/12/14
0	1.046781	-0.7209344	7.27E-07	0.7209344	2016/12/15
0	-1.046781	0.9942984	1.046781	0	2017/1/4
0	1.424312	-0.9691249	0	0.9942984	2017/1/11
0	3.69E-02	-2.52E-02	1.424312	2.52E-02	2017/1/12
0	-1.461169	1.0652	1.461169	0	2017/1/17
0	1.15E-06	-9.16E-07	0	1.0652	2017/1/30
0	-1.15E-06	9.07E-07	1.15E-06	1.065199	2017/1/31
0	1.238408	-1.0652	0	1.0652	2017/2/17
0	-1.01E-06	1.01E-06	1.238408	0	2017/3/3
0	-1.28E-02	1.29E-02	1.238407	1.01E-06	2017/3/6
0	-1.22561	1.237064	1.22561	1.29E-02	2017/3/7
0	1.274649	-1.24998	0	1.24998	2017/3/8
0	-2.68E-07	2.72E-07	1.274649	0	2017/3/15
6	-1.274649	1.297698	1.274649	2.72E-07	2017/3/16
0	0	-1.297698	0	1.297698	2017/3/17
1578.923	1.625615	0	0	0	2017/3/18
0	-1.625615	1.432182	1.625615	0	2017/3/21
0	1.84028	-1.432182	0	1.432182	2017/3/24
7.82E-03	8.26E-06	0	1.84028	0	2017/3/25
0	-1.840288	1.485248	1.840288	0	2017/3/27
0	1.05E-06	-8.94E-07	0	1.485248	20017/3/30
0	1.738669	-1.485247	1.05E-06	1.485247	2017/3/31
0	-1.73867	1.544435	1.73867	0	2017/4/4
0	1.628957	-1.544435	0	1.544435	2017/4/17
0	1.628957	-1.544435	0	1.544435	2017/4/18
0	-1.628957	2.982729	1.628957	0	2017/5/25
0	1.3213	-2.982729	0	2.982729	2017/5/26
0	-0.908	2.494348	1.651687	0	2017/6/6
0	1.150116	-2.494348	0.5167049	2.494348	2017/6/7
0	-1.067165	2.319948	1.666821	0	2017/6/9
0	1.080243	-2.319948	0.5996555	2.319948	2017/6/12
0	-1.679899	3.514567	1.679899	0	2017/6/14
0	1.742972	-3.514567	0	3.514567	2017/6/15
0	-1.742972	3.67417	1.742972	0	2017/6/23
0	1.778756	-3.67417	0	3.67417	2017/6/26
0	-1.778756	3.25488	1.778756	0	2017/7/14
0	2.053241	-3.25488	0	3.25488	2017/7/17
0	-0.4615306	0.9874401	2.053241	0	2017/7/24
0	0.4719018	-0.9874401	1.59171	0.9874401	2017/7/25
0	-2.063612	6.045519	2.063612	0	2017/9/13
0	2.313761	-6.045519	0	6.045519	2017/9/14
0	-2.313761	9.5157	2.313761	0	2020/3/12
0	3.52302	-9.5157	0	11.89457	2020/3/13
0	0	0	3.52302	0	2021/9/10

*(Blue data is the optimized date. Red data is the optimized transaction data. Orange data is the final decision data on September 10, 2021.Trade dates without transactions are omitted.)

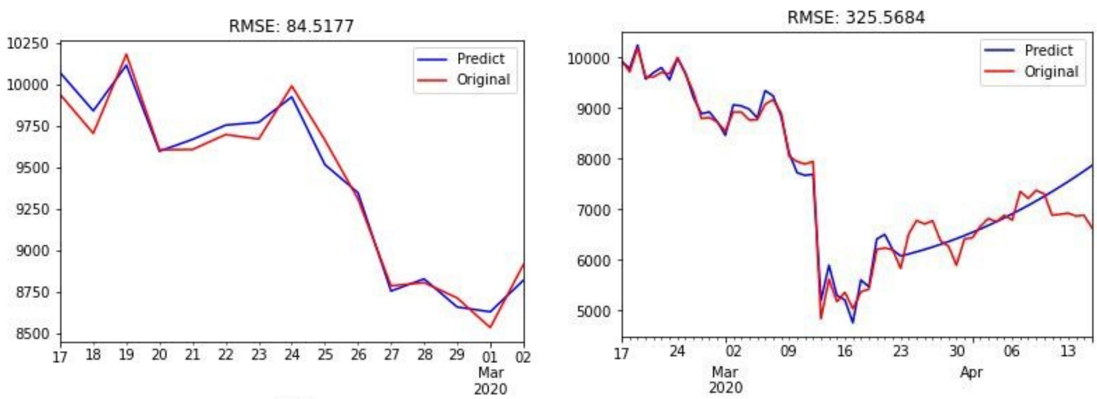
After our processing, on 2016/11/9 we have $1,000 base and following our derived trade strategy over the 5 year trade period, on 2021/9/10 we will have $163,357.8.

7 Model test and error analysis

7.1 Test and error of prediction model

After establishing the value prediction model of gold and bitcoin, the future value of gold and bitcoin is obtained. Comparing the predicted gold and bitcoin values with the actual gold and bitcoin values can verify the correctness of the prediction results and calculate the error of the model.

According to the comparison between the predicted value and the actual value, we find that the two curves have good approximation within 15 days. After more than 15 days, the error increases. After more than 35 days, the prediction curve gradually approaches the linear curve and basically loses the prediction ability.



(a) 15days plot (b) 60days plot

Figure 11: Comparison between predicted value and actual value

In order to further analyze the error, the error of the prediction model can be expressed by the deviation coefficient f between the future value predicted by the model and the real value. The deviation coefficient f satisfies the following formula:

$$f = \sigma/\overline{x}$$

Among them, σ is the standard deviation of the experimental results, and $\overline{x}$ is the average value.

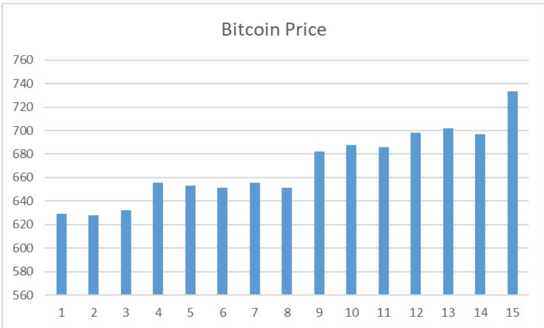
Through python programming calculation, the deviation after standardization is $f = 0.450604 < 0.5$. It shows that the data predicted by the model is close to the real value, and the correctness of the prediction results has been verified.

7.2 Test and error of linear programming model

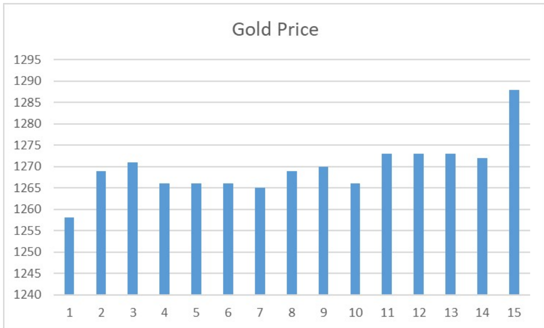
In this report, some data with characteristics from 2016/9/11 to 2021/10/9 are selected and analyzed as a way to test the established linear programming model.

7.2.1 When the price trend is stable

Below is an example of the 15 days from 2016/10/19 to 2016/11/2. The following chart shows the price of bitcoin and the price of gold, it can be seen that the price of bitcoin and gold did not fluctuate much in these 15 days, and their prices are relatively stable from the trend. From the point of view of maximizing trade profits, the number of trade should be reduced in the case of smooth price fluctuations, so as to reduce trade costs and prevent the situation of not being able to make ends meet due to the excessive number of transactions.



(a) Bitcoin Price



(b) Gold Price

Figure 12: Bitcoin Price vs Gold Price from 2016/10/19 to 2016/11/2

The decision derived from the linear programming model at this point is the following:

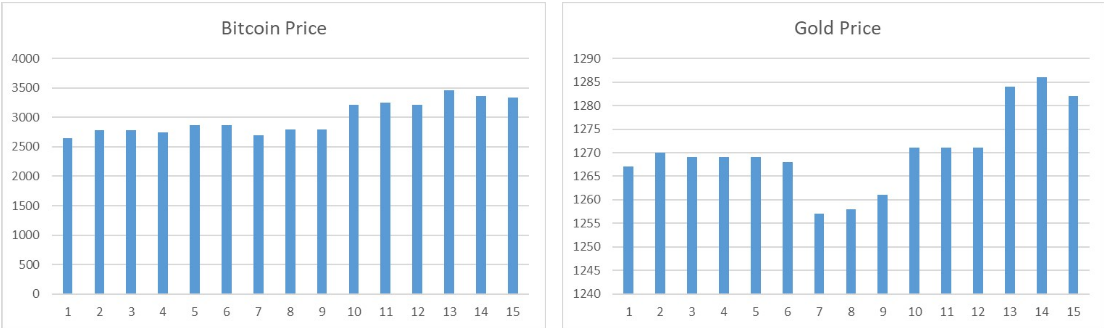
Cash	Bitcoin transaction	Gold transaction	Bitcoin Holdings	Total Value Holdings	Gold Holdings	Date
0	0	0	0	0.7050149	886.9087	**16/10/19**
0	0	0	0	0.7050149	894.6639	**16/10/20**
0	0	0	0	0.7050149	896.0739	**16/10/21**
0	0	0	0	0.7050149	892.5488	**16/10/22**
0	0	0	0	0.7050149	892.5488	**16/10/23**
0	0	0	0	0.7050149	892.5488	**16/10/24**
0	0	0	0	0.7050149	891.8438	**16/10/25**
0	1.80E-06	-9.52E-07	0	0.7050149	894.6639	**16/10/26**
0	-1.80E-06	9.38E-07	1.80E-06	0.7050139	895.3689	**16/10/27**
0	0	0	0	0.7050149	892.5488	**16/10/28**
0	0	0	0	0.7050149	897.4839	**16/10/29**
0	0	0	0	0.7050149	897.4839	**16/10/30**
0	0	0	0	0.7050149	897.4839	**16/10/31**
0	0	0	0	0.7050149	896.7789	**16/11/1**
0	0	0	0	0.7050149	908.0591	**16/11/2**

As can be seen from the table, the number of trade days in these 15 days is only

2 days, which is in line with the prediction just made and indicates that the linear programming model in this report has the ability to provide the best strategy.

7.2.2 When the price trend continues to rise

Take the 15 days from 17/7/27 to 17/8/10 as an example, as shown in the chart of bitcoin price and gold price, the price of bitcoin increased significantly in these 15 days, in order to maximize the trade profit, you should buy a large amount of bitcoin with cash and hold it afterwards and wait for the bitcoin to increase in value.



(a) Bitcoin Price (b) Gold Price

Figure 13: Bitcoin Price vs Gold Price from 2017/7/27 to 2017/8/10

And the data according to the linear programming model of this report is derived from the following table:

Cash	Bitcoin transaction	Gold transaction	Bitcoin Holdings	Total Value Holdings	Gold Holdings	Date
0	0	0	2.063612	0	5463.672	**17/7/27**
0	0	0	2.063612	0	5740.22	**17/7/28**
0	0	0	2.063612	0	5746.748	**17/7/29**
0	0	0	2.063612	0	5666.587	**17/7/30**
0	0	0	2.063612	0	5915.204	**17/7/31**
0	0	0	2.063612	0	5907.916	**17/8/1**
0	0	0	2.063612	0	5558.616	**17/8/2**
0	0	0	2.063612	0	5765.976	**17/8/3**
0	0	0	2.063612	0	5757.478	**17/8/4**
0	0	0	2.063612	0	6640.942	**17/8/5**
0	0	0	2.063612	0	6712.028	**17/8/6**
0	0	0	2.063612	0	6624.608	**17/8/7**
0	0	0	2.063612	0	7134.68	**17/8/8**
0	0	0	2.063612	0	6928.22	**17/8/9**
0	0	0	2.063612	0	6893.043	**17/8/10**

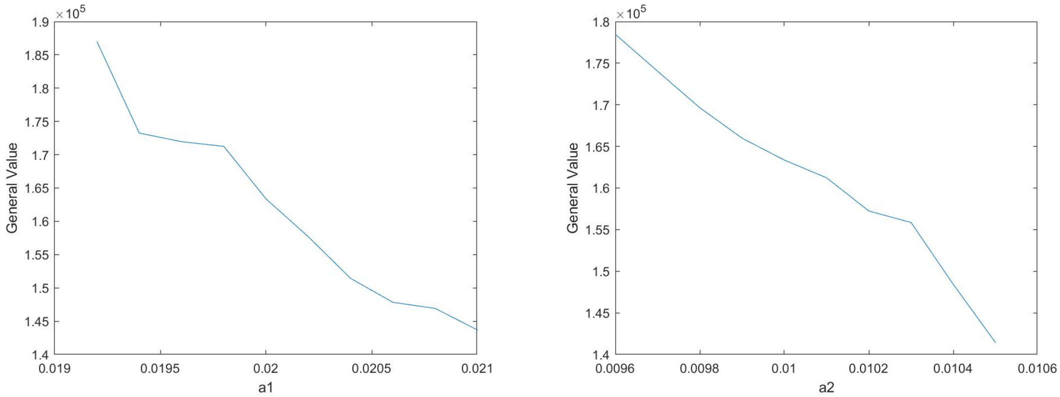
Where it is visible that there is no trading during these 15 days, but rather holding the bitcoin and waiting for its appreciation. This is consistent with the assumptions of this report and is also conducive to maximizing trade profits, suggesting that the linear programming model in this report has the ability to provide the best strategy.

8 Sensitivity analysis

In this model, we default that the transaction cost rates $a_1\%$ and $a_2\%$ of bitcoin and gold are fixed at 2% and 1%, and decide the best trade strategy according to the fixed transaction cost rate. The transaction cost is directly proportional to the transaction cost rate. Next, we will analyze the sensitivity of transaction costs to trade strategies.

First, keep the initial value of $\$1000$ and the gold transaction cost rate $a_2\%$ unchanged, and make the bitcoin transaction cost rate $a_1\%$ fluctuate continuously by 5%. Use matlab to draw the impact diagram of bitcoin transaction cost rate on the final asset value.

Similarly, keep the initial value of $\$1000$ and the bitcoin transaction cost rate $a_1\%$ unchanged, and make the gold transaction cost rate $a_2\%$ fluctuate continuously by 5%. Use matlab to draw the impact diagram of gold transaction cost rate on the final asset value.



(a) the bitcoin transaction cost rate $a_1\%$ (b) the gold transaction cost rate $a_2\%$

Figure 14: Impact of transaction cost rate on final asset value

The impact of transaction cost rate on the final asset value. For the cost rate $a_1\%$ of bitcoin, we know that the cost rate $a_1\%$ increases, and the cost required for each bitcoin transaction also increases. It can be seen from the figure that with the increase of $a_1\%$, the final value commitment decreases monotonically from 2016 to 2021. This phenomenon shows that parameter $a_1\%$ is sensitive to the linear programming model established in this report. Similarly, for the cost rate $a_2\%$ of gold, it can be seen from the figure that with the increase of $a_2\%$, the final value of the cost from 2016 to 2021 will also decline monotonically. Therefore, parameter $a_2\%$ is sensitive to the linear programming model established in this report.

9 Model evaluation and improvement

9.1 Strength

- The model uses grayscale model to forecast when there is too little data (the first 100 days), and uses ARIMA model to forecast when there is a certain amount of data accumulation, avoiding the uncertainty prediction error generated by ARIMA model when there is too little data, and predicting data accuracy.

- The investment planning model in this report is based on linear programming combined with hierarchical analysis, which can handle a large amount of data with universality and process it efficiently and accurately, and make concise and easy-to-operate decision recommendations.
- The investment planning model is supplemented by risk analysis after the initial decision, and can minimize the risk of investment loss in the case of high return on investment decisions.

9.2 Shortage

- The forecasting model in this report is partially over-fitted(about 5%), and the forecasting model has no realistic practice, and its true real-world capability is doubtful.
- The investment planning model is optimized using hierarchical analysis, and its robustness in dealing with investment-type data is slightly lacking.

9.3 Possible improvements

- The forecasting model can be combined with the LSTM neural network model for integrated forecasting. ARIMA-LSTM model is currently more mature in the academic community and can be used in combination thus reducing the overfitting of data and making the model forecasting more realistic.
- The selection of evaluation indexes of the investment planning model needs to be improved, and the unification of evaluation and risk parameters can be improved.

10 Memorandum

A memorandum to Investment managers

–How to sail through the thunderstorm?

Dear Investment managers,

In 2022 and 2021, Bitcoin trading has become a hot commodity in the trading market, creating a huge "Bitcoin storm" among investors. Its high returns and high risk rate have made investors love and fear it. At the same time, gold trading still holds a high market share and is still favored by a great number of investors as one of the most valuable commodities in the world. However, as I mentioned at the beginning, bitcoin & gold trading has a high risk rate, and investing in bitcoin and gold is often at risk of losing money, in fact, it's like sailing a boat on a stormy sea trying to reach the other side safely, and my team and I will provide you with a boat of steel and an accurate map: a great decision and risk assessment system that can predict the price of both in the short term future. Together with your ingenuity, I think you will be able to sail safely through the storm without fail.

With these wishes and ideas in mind, my team and I have created a Bitcoin & Gold Investment Planning Model. This model is based on ARIMA forecasting and grey forecasting models, and uses the currently available bitcoin & gold price data to predict the price direction of both in the next 15 days, which has been verified to be over 95% accurate. After obtaining the forecast data, this planning model combines linear programming principles to make detailed planning recommendations for investors' daily investment plans with the goal of maximizing long-term profits. Based on the last five years of bitcoin & gold price trends provided by Nasdaq, we tried and trained several investment planning models and conducted simulations with the training set as the subject of the experiments, and finally selected the optimal model, which proved to be very effective.

In a simulation relying on our planning model for bitcoin and gold investments, we assume that an investor with $1,000 who faithfully follows the investment recommendations of our bitcoin & gold investment planning model, after 50 buys and sells, after five years, the total value of the virtual investor's money and holdings of bitcoin and gold exceeds $160,000. million dollars. I'm sure you, as experienced professional investment managers, know what this means: a return of over 16,000% in 5 years - enough to wow everyone who sees it. Since the results were obtained in a laboratory setting, the above dizzyingly amazing results may be slightly exaggerated for they, but they can still be quite helpful to you, and we are very happy to share our findings with you, and if this can help you in some way - we will be even happier.We are more than delighted to help you sail through the thunderstorm with our Bitcoin & Gold Investment Planning Model.

Sincerely,

MCM Team Members

References

[1] Yiqing Hua. Bitcoin price prediction using ARIMA and LSTM[J]. E3S Web of Conferences,2020,218;

[2] Makala D,Li Z. Prediction of gold price with ARIMA and SVM[J]. Journal of Physics: Conference Series,2021,1767(1);

[3] Zhang Xiaoduo,Zhai Rui,Gao Wen. Analysis and Economic Prediction of Stock Index Based on Index Tracking and ARIMA Model[J]. Journal of Physics: Conference Series,2021,1903(1):

[4] Khalifa Hamiden Abd El Wahed,Kumar Pavan. Solving fully neutrosophic linear programming problem with application to stock portfolio selection[J]. Croatian Operational Research Review,2020,11(2);

[5] Grima Simon,Thalassinos Eleftherios. Editorial: Risk Management Models and Theories Volume II[J]. Frontiers in Applied Mathematics and Statistics,2021,7;

[6] Cai, Huiping, Cheng, Qian-sheng. Application of attribute hierarchy model AHM in stock selection decision[J]. Mathematical Practice and Understanding,2005(03):55-58;

[7] Yang, Chun-Ling, Zhang, Chuan-Fang. Application of attribute hierarchy model (AHM) in stock selection decision[J]. University Mathematics,2006(05):27-30;

[8] Lu Pugui. Application of AHP hierarchical analysis in financial investment: the selection of bank stock investment targets in A-share market as a benefit [J]. Mall Modernization,2007(31):288-289;

[9] Nie Xiangge. Excellent stock decision based on information entropy decision model[J]. China Management Informatization,2017,20(06):108-109.

12 appendices

12.1 Software and tools

- LaTeX
- Python
- Lingo
- MATLAB

12.2 The Codes

12.2.1 The code of prediction model

```
# -*- coding: utf-8 -*-
"""
Created on Fri Feb 18 14:00:25 2022
ARIMA attempt1
@author: 86136
```

```
"""
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
df=pd.read_csv(r'D:\bit_1.csv',index_col=0, parse_dates=True)
df.head()
data = df['Value']
dt=pd.read_csv(r'D:\date.csv',index_col=0, parse_dates=True)
dt.head()
date = dt['Date']
plt.figure(figsize = (10, 6))
plt.plot(df.index, data)
plt.show()
data_diff = data.diff()
data_diff = data_diff.dropna()
plt.plot(data_diff)
plt.show()
from statsmodels.tsa.stattools import adfuller
print(adfuller(data_diff))
from statsmodels.stats.diagnostic import acorr_ljungbox
print(acorr_ljungbox(data_diff, lags = 20))
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
pacf = plot_pacf(data_diff, lags=20)
plt.title('PACF')
pacf.show()
acf = plot_acf(data_diff, lags=20)
plt.title('ACF')
acf.show()
import statsmodels.tsa.stattools as st
model = st.arma_order_select_ic(data_diff, max_ar=5, max_ma=5, ic=['aic', 'bic', 'hqic']
print(model.bic_min_order)
from statsmodels.tsa.arima_model import ARMA
model_arma = ARMA(data_diff, order = (2,2))
result_arma = model_arma.fit(disp = -1, method = 'css')
resid = result_arma.resid
from statsmodels.graphics.api import qqplot
qqplot(resid, line='q', fit=True)
plt.show()
import statsmodels.api as sm
print(sm.stats.durbin_watson(resid.values))
forecast=result_arma.forecast(15)
forecast=pd.Series(forecast[0],index=pd.period_range(start=str(date[0]),end=str(date[14]
for j in range(1,15):
    forecast[j]=forecast[j-1]+forecast[j]
forecast
```

12.2.2 The Code of planning model

$model:$
$sets:$
$factory/1..15/ : b, g, alpha, x1, x2, z1, z2, c, y;$
$endsets$
$data:$
$a1 = 0.02;$
$a2 = 0.01;$
$b = @file(b.txt);$

$g = @file(liu.txt); alpha = @file(alpha.txt);$
$enddata$
$max = @sum(factory(i) : ((z1(i) + x1(i)) * b(i) - @abs(x1(i)) * b(i) * a1 - x1(i) * b(i)) + ((z2(i) + x2(i)) * g(i) - alpha(i) * @abs(x2(i)) * g(i) * a2 - alpha(i) * x2(i) * g(i)) + c(i));$
$@for(factory(i) : (z1(i) + x1(i)) >= 0);$
$@for(factory(i) : (z2(i) + x2(i)) >= 0);$
$@for(factory(i) : (c(i) - @abs(x1(i)) * b(i) * a1 - x1(i) * b(i) - alpha(i) * @abs(x2(i)) * g(i) * a2 - x2(i) * g(i) * alpha(i)) >= 0);$
$z1(1) = 0;$
$z2(1) = 0;$
$c(1) = 1000;$
$@for(factory(i) : @free(x1(i)));$
$@for(factory(i) : @free(x2(i)));$
$@for(factory(i) : y(i) = (z1(i) + x1(i)) * b(i) - @abs(x1(i)) * a1 - x1(i) * b(i) + (z2(i) + x2(i)) * g(i) - alpha(i) * @abs(x2(i)) * g(i) * a2 - alpha(i) * x2(i) * g(i) + c(i));$
end