Traders buy and sell volatile assets to maximize their income. Recently, with the growing interest of the public in cryptocurrencies, gold and bitcoin become a feasible kind of portfolio. But what comes with interest is a risk. How can we do our best to increase our income? Our team is requested to determine the trading strategies for a trader which utilizes only the past daily prices.

As for problem 1, First, we plot the price data and found that they show non-stationarity in the sense of mean. We, therefore, utilize **ARIMA** to deal with the non-stationarity and predict the future prices of gold and bitcoin. Our choice of parameters is **ARIMA**$(4, 1, 4)$ for both bitcoin and gold. Our **ARIMA**$(4, 1, 4)$-based predictor shows the good capability to accurately predict the price of the future.

To find the best timing to sell and buy the two assets, we first rate them with our well-designed rating system by three important factors:

- Changes in value

- Moving averages

- Bias

Then from the factors our model further calculates our indicator for risk and trend. By linearly composing these two indicators, we get our main factor, which can give us enough information to make trading decisions. By utilizing the information we get with our main factor to make trading decisions, we furthermore set thresholds based on our observations and experiments in our data.

As for problem 2, we dig into the application of Markowitz portfolio theory to our task. By plotting a thermodynamic diagram for various asset allocation schemes, we find that our scheme yields nearly the best result that is possible for us to get. Moreover, it indicates that we achieve better risk management by maximizing the income for taking every unit of risk. Our model reached an **annualized return** of **1469.58**% and yielded **accumulated equity** of **1.03743 × 10⁶ USD**. The **Sharpe ratio** is **1.1358**, which reveals that our risk-taking strategies perform well compared to a risk-free asset.

As for problem 3, we test our model by changing the commissions for trading bitcoin and gold. By visualizing the result we find that commissions do not have a noteworthy impact on the amount of income of our model, which reveals great stability and scalability for more complicated applications.

Key words: ARIMA, Regression, Trading Strategy, Markowitz Portfolio Theory

To Harvest More: Achieve the Best Portfolio with Gold and Bitcoin

Contents

1   Introduction

1.1   Problem Background

With the popularity of cryptocurrencies and the simplification of trading methods, more of the population become market traders. Some of them expect to outperform inflation, while others want to create wealth. By buying and selling volatile assets frequently, market traders pursue a goal to maximize their total return. Gold and bitcoin enjoy great popularity these days for their complementary characteristics in risk and value. Gold is stable in price and has lower risk while the value of bitcoin varies greatly and thus has a higher risk, as is shown in Figure 1.

Figure 1: Gold and bitcoin daily prices, in U.S. dollars per troy ounce and U.S. dollars per bitcoin. **Source:** London Bullion Market Association, 9/11/2021 and NASDAQ, 9/11/2021

Regarding trading rules, Gold is only traded on days the market is open while bitcoin are traded every day. Commissions are charged to make each transaction. For market traders to achieve their goals, they need to build a model to determine the strategies to manage their portfolios well.

1.2   Restatement of the Problem

- Develop a model that gives the best daily trading strategy based only on price data up to that day, and calculate how much the initial $1000 investment is worth on 9/10/2021 using the model and strategy.

- Present evidence that this model provides the best strategy.

- Find how sensitive the strategy is to transaction costs and analyze how the costs affect the strategy and results.

1.3 Our Approach

1. **To predict the prices of bitcoin and gold and make decisions, we use $ARIMA$ (Autoregressive integrated moving average).** To predict future prices with existing data is a difficult problem because too many factors may influence the prices. The international situation, national policies, and even social media can have a considerable impact on the prices. Moreover, the data shows non-stationarity. To take as many factors as possible and predict accurately according to their inner laws, we adopt ARIMA to predict the prices, which is proved to give satisfying results.

2. **To make decisions on trading, design a comprehensive rating system to make decisions on trading.** Based on various factors drawn out from price data, we compute our main factor by composing the trend indicator and risk indicator. We then set thresholds based on both our observations and findings in data. Then we decide the final strategies Markowitz portfolio theory.

2 Assumptions and Justifications

2.1 Assumptions

To simplify the problem stated above, we make the following assumptions, each of which is justified properly:

1. **The trader does not have a bias towards a lower risk.** The two given assets, gold, and bitcoin, differ a lot in risk. To simplify the problem, we assume that the trader will not prefer lower risk than higher and only cares about higher income.

2. **The trader will have \$1000 in the beginning, and the transaction commissions for gold and bitcoin are $\alpha_{gold} = 1\%$ and $\alpha_{bitcoin} = 2\%$, respectively.**

3. **The return rate for cash is 0%** To simplify the problem, we assume that the cash that we hold does not produce any return at all.

4. **The market trader sells all of the gold and bitcoin by the end of the five-year trading period, i.e. on 9/10/2021.** Generally, investors care about funding liquidity. Among cash, gold ,and bitcoin, only cash can circulate unhindered in the market. So we make this assumption and thus measure the outcome in cash.

2.2   Symbols and Definitions

Table 1: Symbols and Definitions.

Notations	Description
$\Delta^k y_t$	The k-order difference of the function y
PC	Percentage change
$P_{0,bitcoin}$	The p-value for ADF test of the raw data of bitcoin
$P_{0,gold}$	The p-value for ADF test of the raw data of gold
$P_{1,bitcoin}$	The p-value for ADF test of the first-order difference of bitcoin
$P_{1,gold}$	The p-value for ADF test of the first-order difference of gold
SS	Threshold for main factor to sell certain asset
BS	Threshold for main factor to buy certain asset
TI	The trend indicator
RI	The risk indicator
MF	The main factor

3   Data Preprocessing

3.1   Fill Missing Value

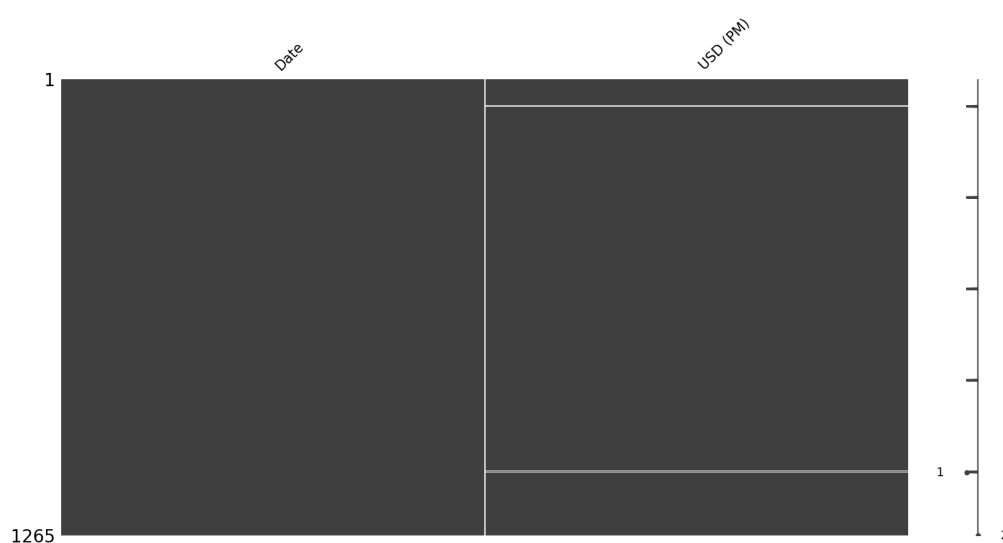There are some missing value in the provided data of gold prices, as is shown in 2

Figure 2: Gold and bitcoin daily prices, in U.S. dollars per troy ounce and U.S. dollars per bitcoin. **Source:** London Bullion Market Association, 9/11/2021 and NASDAQ, 9/11/2021

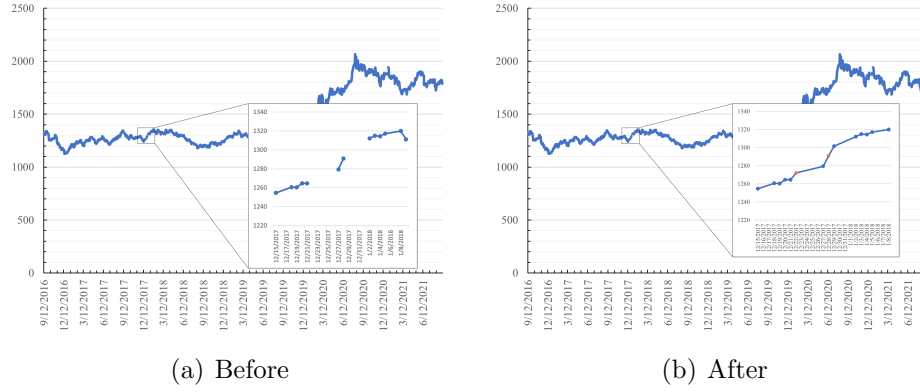(a) Before   (b) After

Figure 3: Missing Filling

The data-dense view reveals that in the price column named USD (PM), there are some missing values roughly at the end of each year. Carefully inspecting the data, we found there are two missing values each year, one is the last day and one is the fourth day from the last.

To make up the them, moving median method is used, where the length of window is 5. For each missing value p_0, we chose a window

$$p_{-2}, p_{-1}, p_0, p_1, p_2$$

Where p_{-2}, p_{-1} are the predecessors of p_0 and p_1, p_2 are the successors of p0. Then, we let the value of p0 be the median of $p_{-2}, p_{-1}, p_0, p_1, p_2$. The method is commonly used when only a few values are missing, which is fast and fits in the trend of the provided data well, as is shown in Figure 3.

We not only check the filled blanks by observing if they fit in the trend, but also compare it with the real-world data, as is shown in Table 2. The error is in an acceptable range, so in the following modeling procedure, we use the modified data.

Table 2: Compare with Real-world Values

Date	Interpolated Value	Real-world Value
12/22/2017	1271.975	1268.05
12/29/2017	1301.525	1296.5

3.2   Metric Calculating

3.2.1   Average Change

Using the trained predictor, we calculate the average of change in the period of 5,10,15,20,and 25 days.

The change each day is defined as

$$Percentage\ change = PC = \frac{New\ price - Old\ price}{Old\ price} \times 100\%$$

The result for the given data is in the Appendix: Figures and Tables. Figure 12 and Figure 13 show the average change we calculated for bitcoin and gold, respectively.

3.2.2   Bias

Bias measures how much the closing price shifts from the average price.

The formula to calculate the Bias in a n day period is

$$Bias = \frac{Closing\ price}{The\ mean\ price\ of\ n\ days} - 1$$

Bias can help to track the market for the rise and fall of the price, which helps us to decide whether and when to sell or buy.

3.2.3   Moving Averages (MA)

An n day's MA is the average of the price today and the previous $n-1$ days. Plotting all the MAs in a chart can reveal the trend of the price of the given period. Combined with the current price, it helps us to find the favorable timing to increase our outcome.

4   Mathematical Modeling

4.1   ARIMA Predictor

ARIMA (autoregressive integrated moving average) is widely used for forecasting time series data, which is a generalization of the ARMA (autoregressive moving average) model[3]. ARMA requires the mean function of the data to be stationary because history data of stationary series can be used to predict the future. But time series are not always stationary, so ARIMA takes an initial differencing step to eliminate the non-stationarity of the mean function. In this case, we find the first-order difference is stationary, so we take first-order

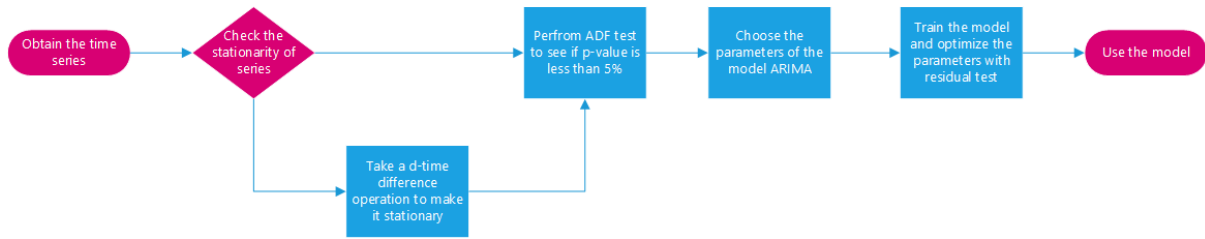difference. Figure 4 shows the procedure we take to build the ARIMA model.

Figure 4: Procedure to Build ARIMA Model

To apply the model, we first have to test the stationarity.

4.1.1 Augmented Dickey-Fuller Test

An augmented Dickey–Fuller test (ADF) is a kind of hypothesis test. The null hypothesis is

$$H_0 \ : \ A \ unit \ root \ is \ present \ in \ a \ time \ series \ sample.$$

Subject to the χ^2 distribution, we accept it when the observed value of the waviness of the series, P, also known as the $p-value$, satisfies

$$P \geq 0.05$$

It can test whether the data is stationary or trend-stationary. If the unit root is present in the series, the time series is non-stationary.

In the raw data, we found

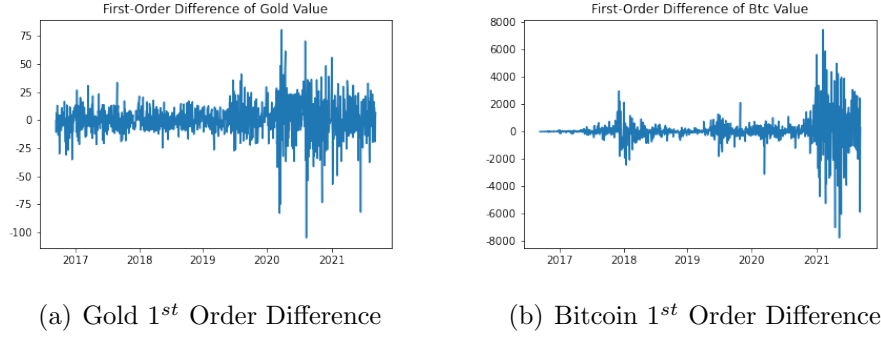$$P_{0,bitcoin} = 0.8420802907198858 > 0.05$$
$$P_{0,gold} = 0.9042384812941653 > 0.05$$

So we accept H_0. The raw time series is non-stationary. And in the first-order difference, we found

$$P_{1,bitcoin} = 9.276161468079189e - 13 < 0.05$$
$$P_{1,gold} = 9.269711421536524e - 13 < 0.05$$

The first-order difference is stationary, So we plot them in Figure 5.

(a) Gold 1st Order Difference

(b) Bitcoin 1st Order Difference

Figure 5: The 1st Order Difference

We also can see from the plot that it is stationary. Therefore, we use it for our ARIMA model. The parameter d of $ARIMA(p, d, q)$ is $d = 1$.

4.1.2   Finite Difference

The difference of a function y is defined as

$$\Delta^k y_t = \Delta(\Delta^{k-1} y_t) = \Delta^{k-1} y_{t+1} - \Delta^{k-1} y_t = \Sigma_{i=0}^k (-1)^i C_k^i y_{t+k-1} \ (k = 1, 2, 3, \dots)$$

Specially, the first-order difference, which is commonly used in ARIMA model, is defined as

$$\Delta y_t = y_{t+1} - y_t$$

Finite difference is useful in processing non-stationary data[4].

4.1.3   Choice of Parameters

To select the best parameters of the ARIMA model, we first rely on observation on PACF and ACF plots. And then, if it does not fall in a range that makes sense, we search for possible parameters by grid searching the feasible domain.

The parameter p of ARIMA(p,d,q) is the number of lag observations in the model, which is also known as the lag order. To assign p, the ACF plot in Figure 6 is inspected. We find that $p = 4$ is the best choice for both bitcoin and gold.

The parameter q is the size of the moving average window, which is also called the order of moving window. To assign q, the PACF plot in Figure 6 is inspected. We find that $q = 4$ is the best choice for both bitcoin and gold.
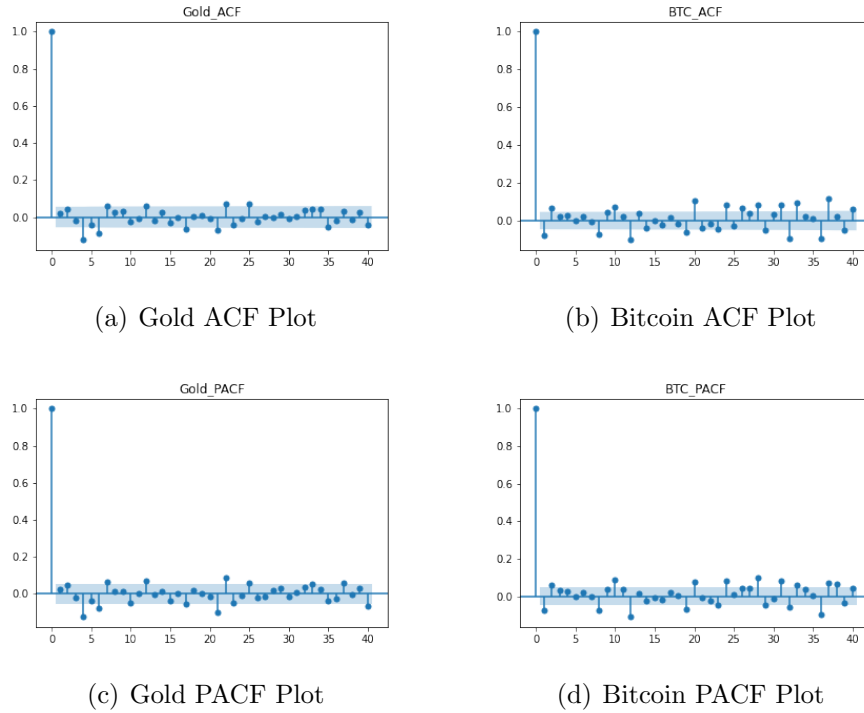
(a) Gold ACF Plot

(b) Bitcoin ACF Plot

(c) Gold PACF Plot

(d) Bitcoin PACF Plot

Figure 6: ACF and PACF Plot for Gold and Bitcoin

4.1.4 Residual Test and Durbin-Watson Test

To test whether our model has extracted all information in the data, we inspect the residuals and test if they look like white noise. White noise residuals reveal that all information to predict has been extracted so that we can utilize them to predict the future. And what is left is only random perturbation and is of no use for our ARIMA predictor. To carry out the test, we first plot the residual Q-Q plot and then carry out a Durbin-Watson(DW) test. The test results are as follows.

1. **Residual Q-Q Plot** A Quantile-Quantile plot (Q-Q plot) shows the "match" of an observed distribution with a theoretical distribution, almost always the normal distribution. We plot the Q-Q plot and find that the residual distribution is roughly the same as the normal distribution, or is similar to white noise.

   We can see in Figure 7 that the scatters appear to be around a straight line, which indicates our parameters are good enough to predict the future.

2. **DW Test** DW test is commonly used for testing whether autocorrelation presents in the residuals from a regression analysis. The DW statistics has a null hypothesis.

$$H_0 \; : \; \rho = 0$$
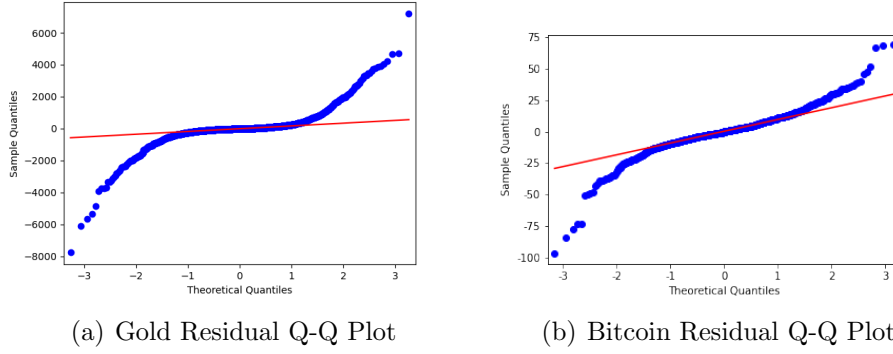
And the test statistic is

(a) Gold Residual Q-Q Plot   (b) Bitcoin Residual Q-Q Plot

Figure 7: Residual Q-Q Plot for Gold and Bitcoin

$$d \; = \; \frac{\Sigma_{t=2}^{T}(e_t - e_{t-1})^2}{\Sigma_{t=1}^{T} e_t^2} \tag{1}$$

We calculate d in equation 1, which is the approximation of $2(1 - \hat{\rho})$, where $\hat{\rho}$ is the sample autocorrelation of the residuals. So $d = 2$ indicates no autocorrelation. In our case, we calculate the values in by equation 1, and find

$$d_{gold} \; = \; 2.0179203733750732$$
$$d_{bitcoin} \; = \; 2.0000896547882885$$

which are very close to 2, indicating very low autocorrelation.

4.2 Trade Decision

4.2.1 Position Management

According to Markowitz portfolio theory[6], the final return of using a multi-asset portfolio investment can reach the weighted average of each asset's return, but the portfolio can significantly reduce systematic risk. The aim is to flatten the return curve and reduce retracements.

Markowitz utilizes mean-variance analysis in his model to measure the expected return and risk of each asset, and in the following, we will also use this model to quantify the expected return and risk of this portfolio of bitcoin and gold.

To measure the risk of the portfolio, we calculate the standard deviation. The standard deviation of the portfolio is defined as

$$\sigma = \sqrt{\mathbf{w_T} \cdot \Sigma \cdot \mathbf{w}}$$

Where σ is the standard deviation of the portfolio, Σ is the covariance matrix of the portfolio, and vector **w** is the weight of the portfolio. So to solve the given problem, the vector **w** is defined as

$$\mathbf{w} = (w_{cash}, w_{bitcoin}, w_{gold})$$

Where w_{cash}, $w_{bitcoin}$, and w_{gold} is the weight in the portfolio for cash, bitcoin and gold, respectively.

To measure the return of the portfolio, we use the annualized return of the portfolio.
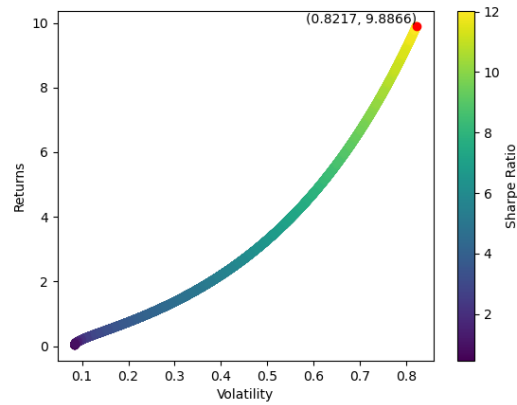

Figure 8: Application of Markowitz Theory

Using the mean-variance analysis, we try to find a portfolio that minimizes the risk and maximizes the return. In our model, we use the Monte-Carlo method to simulate the positions of bitcoin and gold. The result of the simulation is plotted in Figure 8 [2]. We can conclude from the figure that if the trader is looking for a higher Sharpe Ratio[1], we should look for the point with the largest value of the vertical coordinate over the horizontal coordinate, or on the top-left area of the figure. Extracting the portfolio data behind this point, we find that most portfolios with a Sharpe Ratio of 10 or higher have a portfolio share of 95% or more in bitcoin. Therefore, we conclude to invest bitcoin as much as possible in the whole portfolio. When both gold and bitcoin are on an upward trend, traders are supposed to buy more bitcoin for their portfolio to achieve a high expected return.

4.2.2 Rating Gold and Bitcoin

Based on the closing price every day, we first calculates the required normalized factors using including the change today, and the average change of 15 days and the bias, for both bitcoin and gold. Then our model predict the price the next day with ARIMA model.

Then, we calculate the indicators[5] with the formula below for bitcoin and gold.

$$Trend\ indicator = TI = 0.7 \times 30\ day's\ averagechange\ +\ 0.3 \times 20\ day's\ average\ bias$$
$$Risk\ indicator = RI = 0.4 \times Normalized\ 15\ day's\ average\ change + 0.6 \times bias$$

Then we compose TI and RI linearly with ratios chosen according to Markowitz portfolio theory and then normalize it to get our main factor (MF) for bitcoin and gold.

We set following thresholds for selling or buying bitcoin and gold in Table 3.

Table 3: Trading Thresholds.

	Bitcoin	Gold
Buy(BS)	0.775	0.57
Sell(SS)	0.55	0.3

Then, we trade bitcoin and gold subject to the following rules.

	$MF_{bitcoin} > BS_{bitcoin}$	$Others$	$MF_{bitcoin} < SS_{bitcoin}$
$MF_{gold} > BS_{gold}$	**Further Judgment Required**	Buy gold	Sell bitcoin and buy gold
Others	Buy bitcoin	Do nothing	Sell bitcoin
$MF_{gold} < SS_{gold}$	Sell gold and buy bitcoin	Sell gold	Sell both

When further judgment is required, we come up with the following factor F

$$F = 2.5 \times (MF_{bitcoin} - BS_{bitcoin}) - (MF_{gold} - BS_{gold})$$

Then, our model recommend buying gold when $F < 0$ and buying bitcoin when $F > 0$. In the above formula, the ratio 2.5 comes from the weights that reach the maximum in return that we find in the previous sub-section Position Management.
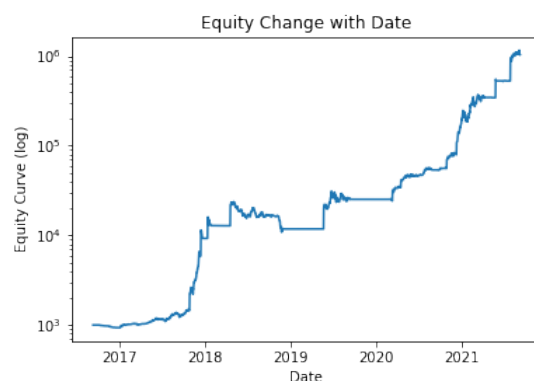
5   Results and Solutions


Figure 9: Equity Change Result

As is plotted in Figure 9, our equity change rises a lot with time going by. From the plot we can know the following facts.

- **Our strategies work well to maximize income.** As is seen in the plot, the equity rises in the order of magnitude from 10^3 to 10^6, which is a considerable achievement.

- **Our strategy has a distinct feature that it has very low retracement.** As is seen in the plot, only there are only a few retracements, all of which have small value changes. It reveals that we have a good performance in dealing with risks.
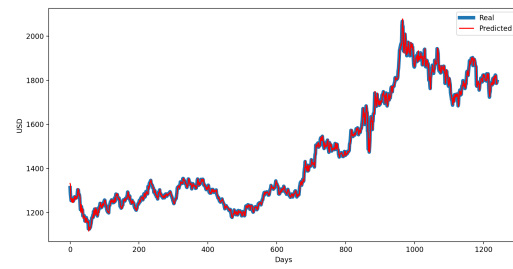
The detailed performance data is shown in Table 4. From the table we can have a quantitative view on the brilliant performance of our model.

Table 4: Performance Data

Performance Indicator	Measurement
Accumulated Equity	1.03743×10^6
Annualized Return	1469.58%
Maximum Retracement	-53.90%
Annualized Return/Retracement Ratio	27.77
Sharpe Ratio	1.1358
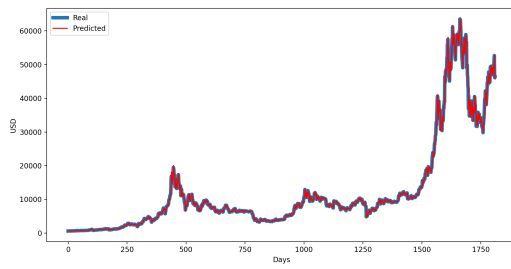
(a) Bitcoin Prediction   (b) Gold Prediction

Figure 10: Evaluation of Prediction for Gold and Bitcoin

6   Model Evaluation and Sensitivity Analysis

6.1   Model Evaluation

First, we test our ARIMA predictor. Our predictor predicts the price the next day according to the history data. So we test its function by comparing the prediction with the real data. We plot the result day by day in Figure 10.

As is seen in the plot, our predictor provides a very accurate prediction for us to make decisions on trading, making a good strategy possible.

6.2   Sensitivity Analysis

For each transaction, a commission is charged. The commission can also have an impact on our model. We tested our model with different commissions and found that the level of commission does not significantly affect our model.
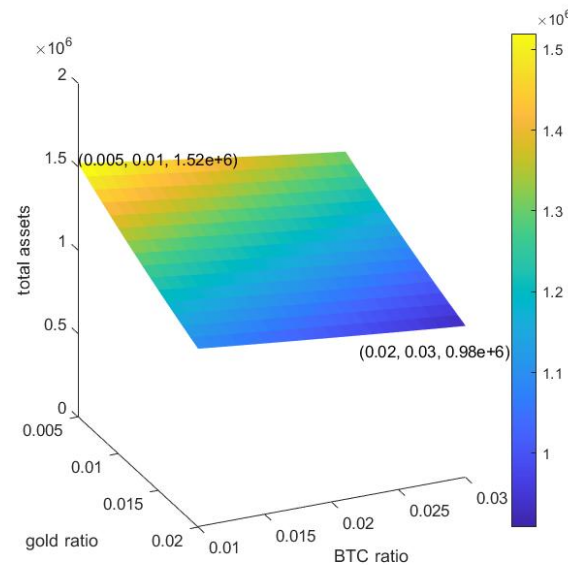
Figure 11: Equity Change with Commissions

By changing commissions and recalculating the total assets we get, we plot Figure 11, from which we can draw the following conclusions

1. Commissions may influence the equity change in the end, but the range is considered small, which reveals the great stability of our model.

2. From the figure we plot, we find that commissions for bitcoin have a stronger impact on the income. It is because to reach higher income, we hold more bitcoin so that when commissions go higher, we are charged much more.

3. The given commissions are higher than the commissions in the real world, 0.02% for bitcoin and 0.05% for gold. When choosing the thresholds for trading, we can choose more suitable parameters and trade in a lower frequency to reduce frictional costs.

What is worth mentioning is that our model has a quick responds to accidents in the investment market. In the investment market, there are many potential risks. In this paper, we verify the validity of the model by changing the commissions. we draw the conclusion that the commission does not have a significant impact on the model. But in fact, the investment market can be affected by many factors in the real world, such as wars, government regulations, and other negative news. It is worth mentioning that the model performs well in unexpected negative situations. For example, Bitcoin plunged on 21/5/19, and we reduced the position of Bitcoin through the portfolio model before the plunge, thus reducing the impact of systematic risk on the final return of the portfolio.

7   Strength and Weakness

7.1   Strength

The models have the following strengths:

- **Our model yields considerable income**

- **Our model does not involve artificial neural network, which saves computational resources and improves interpretability.**

- **Our model builds up a fair system to assess a financial asset.**

- **Our model can be generalized to all problems in trading.**

7.2   Weakness

Though our model performs well, the models have following weaknesses:

- **We do not consider factors other than prices.**

- **The choice of our model's parameters relies on experience and grid search.**

8   Conclusions

In this paper, we build an ARIMA based predictor and the rating-based decision-maker to enhance the income of trading gold, and bitcoin. First, we preprocess the data and calculate various factors. Then, we do feature engineering by finite difference to deal with non-stationarity. After that, we use an ARIMA(4,1,4) predictor to predict the price in the future. Finally, we design a rating system and then derive our main factor to decide whether to buy, hold, or sell bitcoin and gold, and our strategies is tested with Markowitz portfolio theory. Sensitive tests are carried out to find that our model will remain stable when commissions change. We measured our model by performance indicators like annualized return, retracement, and Sharpe ratio. Our model yields satisfying results.

References

[1] Lin Hongmei,Du Jinyan,Zhang Shaodong. Sharpe ratio: estimation method, applicability and empirical analysis[J]. Journal of Statistics,2021,2(06):73-88.DOI:10.19820/j.cnki.issn2096-7411.2021.06.006.

[2] Sun, Lipo. An empirical study of Makowitz's portfolio theory based on Python[J]. Times Finance,2020(25):46-47+50.

[3] Pierre Rostan,Alexandra Rostan,Mohammad Nurunnabi. Options trading strategy based on ARIMA forecasting[J]. PSU Research Review,2020.

[4] Marcos Lopez dai Prado, *Advances in Financial Machine* Learning[M]:75-82. ISBN 978-1119482086

[5] Li F. Modelling the stock market using a multi-scale approach[D]. University of Leicester, 2017.

[6] Marling H, Emanuelsson S. The Markowitz Portfolio Theory[J]. November, 2012, 25: 2012.

Memorandum to the Trader

Considering the intensely changing financial markets and the difficulty of handling your portfolio, using an appropriate model to make predictions and strategies to trade are of vital importance to improve income. It is a great honor for us to develop the model for you to buy, hold and sell your assets. Here is our model based on **ARIMA** and strategies for you to trade your assets effectively.

1. Our model calculates the required normalized factors using the closing price every day, including the change today, the average change of 15 days, and the bias, for both bitcoin and gold. Then our model predicts the price the next day with the ARIMA model.

2. Based on the calculated values and our research, we calculate the indicators with the formula below for bitcoin and gold.

$$Trend\ indicator = TI = 0.7 \times 30\ day's\ averagechange\ +\ 0.3 \times 20\ day's\ average\ bias$$
$$Risk\ indicator = RI = 0.4 \times Normalized\ 15\ day's\ average\ change + 0.6 \times bias$$

Then we compose TI and RI linearly with ratios chosen according to Markowitz portfolio theory and then normalize it to get our main factor (MF) for bitcoin and gold.

3. We set following thresholds for selling or buying bitcoin and gold in Table 5.

Table 5: Trading Thresholds.

	Bitcoin	Gold
Buy(BS)	0.775	0.57
Sell(SS)	0.55	0.3

Our model trades bitcoin and gold subject to the following rules.

	$MF_{bitcoin} > BS_{bitcoin}$	$Others$	$MF_{bitcoin} < SS_{bitcoin}$
$MF_{gold} > BS_{gold}$	***Further Judgment Required***	Buy gold	Sell bitcoin and buy gold
Others	Buy bitcoin	Do nothing	Sell bitcoin
$MF_{gold} < SS_{gold}$	Sell gold and buy bitcoin	Sell gold	Sell both

When further judgment is required, we come up with the following factor F

$$F = 2.5 \times (MF_{bitcoin} - BS_{bitcoin}) - (MF_{gold} - BS_{gold})$$

Then, our model recommends buying gold when $F < 0$ and buying bitcoin when $F > 0$.

With above strategies, our model achieve great income. The performance of our model is as follows

Table 6: Performance

Performance Indicator	Measurement
Accumulated Equity	1.03743×10^6
Annualized Return	1469.58%
Maximum Retracement	-53.90%
Annualized Return/Retracement Ratio	27.77
Sharpe Ratio	1.1358

We appreciate this opportunity to help you to build up a trading strategy for cash, gold, and bitcoin, and we firmly believe that our model can be utilized in your tradings to maximize your income. Feel free to contact us for further information on the proposal.

Sincerely yours

MCM 2022 Team
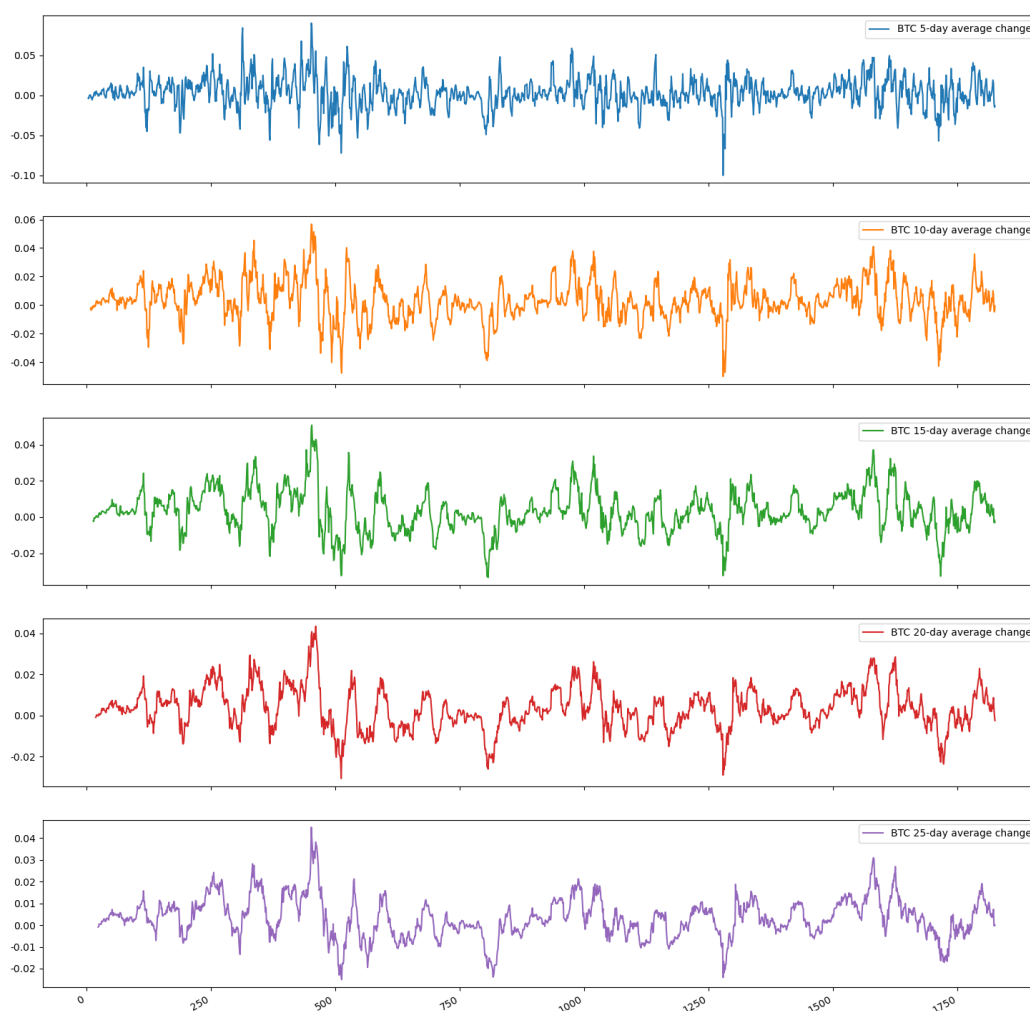
Appendix: Figures and Tables
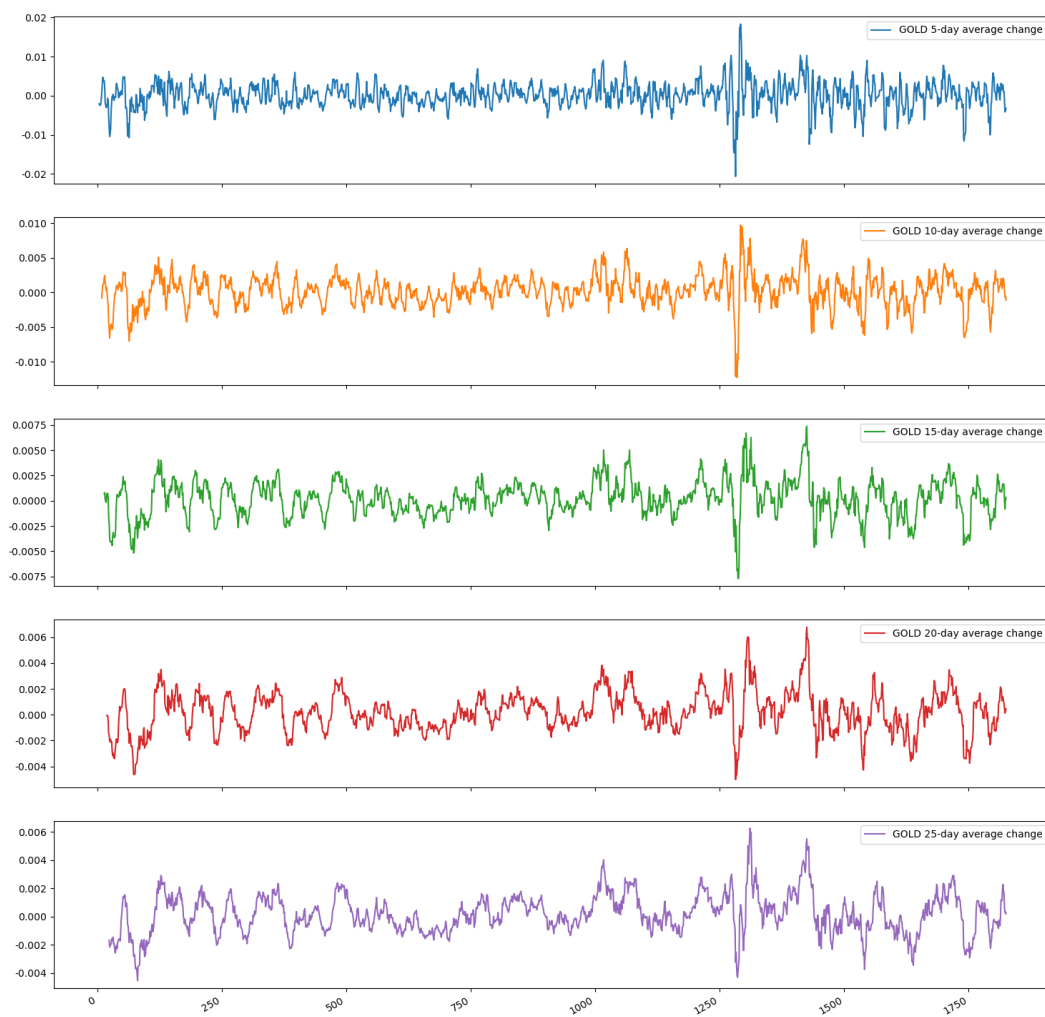

Figure 12:   Average Change in 5,10,15,20,25 days for bitcoin

Figure 13:   Average Change in 5,10,15,20,25 days for gold