Trading Strategy Analysis on Exchange Traded Fund

Jay Xiao
University College London

March 28, 2023

1 Introduction

Trading strategy on stock markets is a large area in computational finance. One of the early ideas is technical analysis, it can be traced back to the early 18th century the Japanese used the oldest version of candlesticksNison (2001). Technical analysis involves using technical indices on historical data with a set of rules to predict the future price trend and make decisions. Another active research area is time series analysis, by training a statistical model to predict the future trend.

Chio (2022) studied the simple trend following strategies based on different technical indicators in the U.S. stock markets. Strategies claimed positive returns, especially on the more liquid stock sets. On traditional statistical models, Vo and Ślepaczuk (2022) developed a new rolling window-based ARIMA-SGARCH trading strategy. The result shows that the new hybrid model outperforms the original ARIMA with relatively high robustness. The recent development of machine learning also raised studies on using machine learning to predict trading signals. Dash and Dash (2016) proposed a new trading strategy based on CEFLANN neural network to generate buy sell signal that provides both high efficiency and accuracy.

In this report we will explore three trading strategies on the second most traded ETF today, S&P500 exchange-traded fund(SPDR): The mean-reversion model relying on the technical indicator, the ARIMA model, and the ARIMA-GARCH model. In the end, we will also compare the performance of these models.

2 data and Methodology

2.1 assumption and problem statement

Here are the assumptions made in this paper:

1. Missing data are merged from the previous day's data.
2. There is no commission fee during the trading.
3. We assume that if the margin balance drops, we will immediately top up the loss from the remaining cash.

4. Trade executes after the day of the trading signal occurs. i.e. Trading at the close price.

5. We assume that the unused capital will be put into a money market and grow at a risk-free rate r_t from the time t market close to the next time $t+1$ market close.

6. The smallest unit of the S&P500 ETF we can buy is 1.

The question we want to investigate is: Given we have initial wealth $V_0 = \$200000$ with leverage $L = 5$ on the SPDR which can be long or short, we want to find a leveraged strategy (i.e. a sequence $\{\theta_t\}_{t=1}^{T}$ of dollar values of SPDR) such that the final wealth V_T growth.

2.2 DATA ANALYSIS

The historical data of S&P500 exchange-traded fund is downloaded from yahoo finance[1] and we will use effective federal fund[2] rate as our risk-free rate. The data will be split into a 70% training set and a 30% test set.

From figure 1, we can see that in several periods like the Aug of 2015, the beginning of 2018 and the beginning of 2019 shows relatively high volatility on returns, and we can also see that the effective federal fund rates are continuously increasing from 2014 to 2019 and dropped a little afterwards.

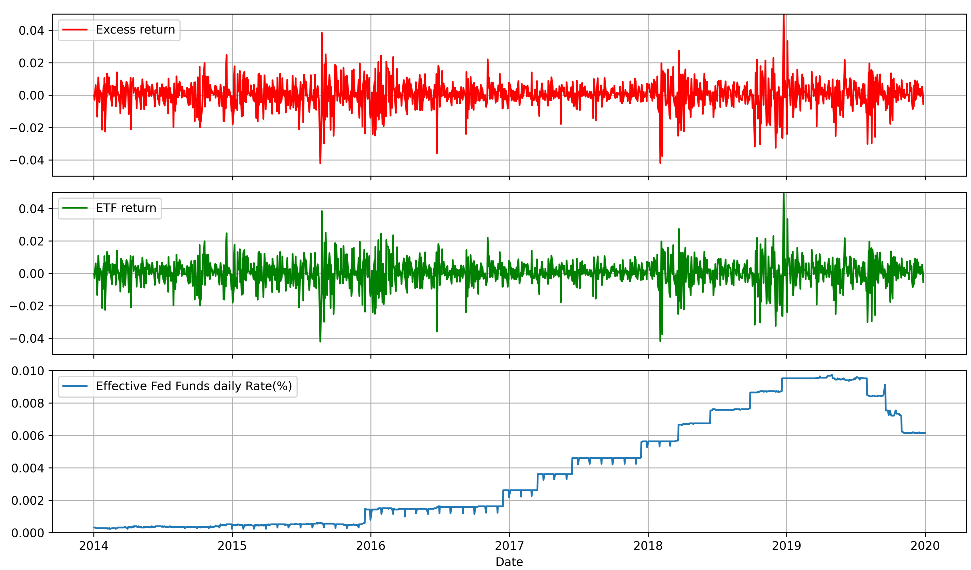

Figure 1: From below to top is the effective federal fund daily rate, S&P500 Exchange Traded Fund(SPDR) daily return and the excess return in the period between 1 Jan 2014 to 31 Dec 2019

[1]https://finance.yahoo.com/quote/SPY/
[2]https://www.newyorkfed.org/markets/reference-rates/effr

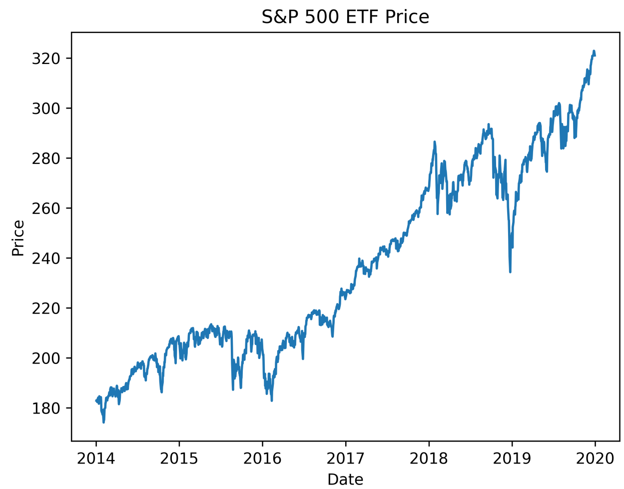

Figure 2: S&P 500 ETF Price

We also did an augmented Dickey–Fuller (ADF) test on the first difference SPDR close price series, it is a method to test the null hypothesis a time series exits a unit root (i.e. The time series is non-stationary)Shumway et al. (2000). The resulting test statistic is -0.146 and p-value 3.9×10^{-27} which is way smaller than the significance level 5%. So we reject the null hypothesis, the first difference of the SPDR close price series is stationary.

2.3 Methodology

2.3.1 Mean-reversion strategy using Bollinger bands

Bollinger Bands was created by John Bollinger in the 1980s. Bollinger bands indices are the combination of the moving average and standard deviation Bollinger (2001). The index includes calculating the moving average of the past 15 trading days (time window is determined by the training set performance), and plus/minus two standard deviations $\sigma(t)$ from that period. At time t, the moving average for the previous period n is defined as:

$$\mathrm{MA}(t) = \frac{1}{n}\sum_{i=1}^{n} x_{t-i+1}$$

and the upper bound of the Bollinger bands is $\mathrm{MA}(t) + 2\sigma(t)$ and the lower bound of the Bollinger bands is $\mathbf{MA}(t) - 2\sigma(t)$
The detailed rules are:

- Long signal: If the daily close price S_t goes below the lower bound and there is no long or short position hold currently, we will open a long position of SPDR dollar value $\theta_t = \lfloor 0.65 * V_t * L / S_t \rfloor * S_t$
- Short signal: If the daily close price S_t goes above the upper bound and there is no long or short position hold currently, we will open a short position of SPDR dollar value $\theta_t = -\lfloor 0.65 * V_t *$

$L/S_t \rfloor * S_t$

- Close long position: If the stock daily close price S_t passes through $\text{MA}(t) + 0.5 * \sigma(t)$, we will close the long position.
- Close short position: If the stock daily close price S_t passes through $\text{MA}(t) - 0.5 * \sigma(t)$, we will close the long position.
- Stop loss long: If the Loss is above 5% (i.e. $\frac{(S_{t_{long}} - S_t) * L}{\theta_{t_{long}}} < -5\%$ where t_{long} is the time that we open the long position.), we close the long position.
- Stop loss short: If the Loss is above 5% (i.e. $\frac{(S_{t_{short}} - S_t) * L}{-\theta_{t_{short}}} < -5\%$ where t_{short} is the time that we open the short position.), we close the short position.

2.3.2 Rolling ARIMA model

The ARIMA model also known as Box–Jenkins method was developed by Box et al. (2015) to predict the time series. The ARIMA model is a combination of:

$$ARIMA(p, d, q) = \begin{cases} AR(p) & \text{p-order Autoregressive process} \\ MA(q) & \text{q-order Moving Average process} \\ d & \text{The number of differences (order) to make it a stationary sequence} \end{cases}$$

Definition: Shumway et al. (2000) Given a time series $\{Y_t\}$,with $\{\epsilon_t\} \sim \mathit{WhiteNoise}(0, \sigma^2)$. And we define lag operator $\mathbf{B}$ as $\mathbf{B}^d Y_t = Y_{t-d}$

$\{Y_t\}$ is an autoregressive process of order p, if

$$Y_t = \sum_{i=1}^{p} \phi_i Y_{t-i} + \epsilon_t$$

$$(1 - \sum_{i=1}^{p} \phi_i \mathbf{B}^i) Y_t = \epsilon_t$$

,$\{Y_t\}$ is a moving average process of order q, if

$$Y_t = \epsilon_t + \sum_{i=1}^{q} \theta_i \epsilon_{t-i}$$

$$Y_t = (1 - \sum_{i=1}^{q} \theta_i \mathbf{B}^i) \epsilon_t$$

Where ϕ_i and θ_i are constants $\forall i$. Then, an ARIMA(p, d, q) model can be described as

$$(1 - \sum_{i=1}^{p} \phi_i \mathbf{B}^i)(1 - \mathbf{B})^d Y_t = (1 - \sum_{i=1}^{q} \theta_i \mathbf{B}^i) \epsilon_t$$

The Rolling ARIMA algorithm is inspired from Siami-Namini and Namin (2018). We will first train the ARIMA model in a time window, the trained ARIMA model will predict the next-day price, then

add the actual next-day price into the historical data to refit the ARIMA model and make predictions for the day after the next day. Optimal p and q will be chosen by the minimum mse in the training set.
The trading rules are:
To avoid noise in the prediction, both long and short signals have been set at a limit (See below the limit i,j are regarded as the hyperparameter, which will be tuned in the training set).

- Long signal: If the predicted close price rise more than i% than the current day close price and no long positions hold currently, we will open a long position of SPDR dollar value $\theta_t = \lfloor 0.65 * V_t * L/S_t \rfloor * S_t$
- Short signal: If the predicted close price drops more than j% than the current day close price and no short position hold currently, we will open a short position of SPDR dollar value $\theta_t == \lfloor 0.65 * V_t * L/S_t \rfloor * S_t$
- If we receive a short signal but we currently hold a long position, we will close the long position first and then follow the above rule. vice versa.
- Stop loss long: If the Loss is above 5% (i.e. $\frac{(S_{t_{long}} - S_t) * L}{\theta_{t_{long}}} < -5\%$ where t_{long} is the time that we open the long position.), we close the long position.
- Stop loss short: If the Loss is above 5% (i.e. $\frac{(S_{t_{short}} - S_t) * L}{-\theta_{t_{short}}} < -5\%$ where t_{short} is the time that we open the short position.), we close the short position.

2.3.3 Rolling ARIMA-GARCH Model

The Generalized AutoRegressive Conditional heteroskedasticity(GARCH) model is proposed by Bollerslev (1986). Unlike the ARIMA model, which assumes the time series is homoscedastic, the GARCH model assumes the time series has non-constant variances. It models the volatility of the time series. The

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

where ϵ_t is the error term at time t, σ_t^2 is the conditional variance at time t, ω is a constant, α_i are the coefficients of the lagged squared error terms, and β_j are the coefficients of the lagged conditional variance terms. The GARCH(p,q) model is a popular method for modelling volatility in financial time series data. As a discrete version of the stochastic volatility model, GARCH can also capture the fat tail effect of the stock market. Therefore, combining ARIMA and GARCH is expected to be more suitable than either model alone when simulating stock prices.
The combination model act in the following order Kaur et al. (2022):
We will first feed in the time window T into the ARIMA model, the best p and q is identified by minimising training set rmse, it will be fixed in the test set. After we fit the ARIMA model, we will do residual diagnosis by using the Augmented Dickey–Fuller test to analyse whether the residual plot is stationary or not. If the statistical test indicates that we cannot reject the residual series is

non-stationary, we will fit the residual series to the GARCH model to fit the volatility. Then combine both of the model predictions together to get the final prediction. The rolling mechanism is the same as the Rolling ARIMA model.
The trading strategy is also the same as the Rolling ARIMA model.

3 Result

By tuning the rolling window, we choose $T = 300$ as the optimal value in the training set. But this means we do not have trading signals at the beginning of the 300 trading days. So some adjustments are needed, in order to make the mean-reversion strategy and the two ARIMA models comparable in the training set, the first 300 days for all three strategies cannot make long and short positions. For the test set, because the test set only has 453 trading days, we will take 300 days at the end of the training set as the input to predict the first day's close price in the test set and so on. (Note that the profit of the mean-reversion strategy in the train set should be larger if we start trading on day 0.)

3.1 Strategies position Analysis

θ_t is the dollar value of SPDR holding at time t. Figure 3 shows the position of the three strategies we developed over time. The plot shows that the mean-reversion strategy is the most active strategy, it frequently changes its position to both long and short, Table 1 also proves that the turnover units of the mean-reversion strategy are 5 times more than the turnover units of the ARIMA strategy and 10 times more than the ARIMA-GARCH strategy in the test set. As we can see in the test set between the end of 2018 and the beginning of 2019, there was a significant value drop in the rolling ARIMA strategy, whereas the ARIMA-GARCH model successfully detect the potential volatility and choose not making any decision which prevent a huge loss. However, the ARIMA-GARCH model also ignores the small up trend at the beginning of the test set.

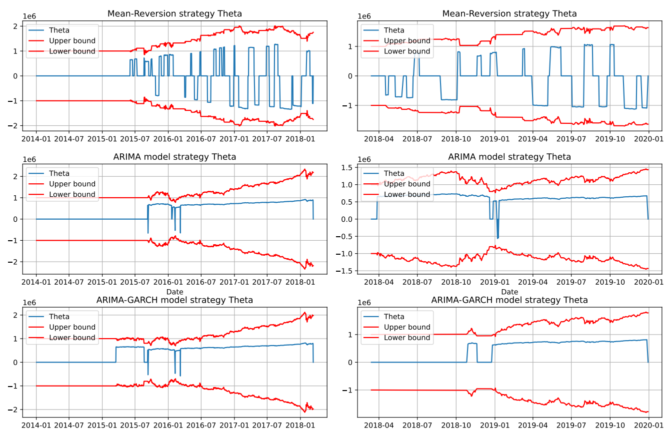

Figure 3: The positions of each strategy with its upper and lower bound (Left: training set. Right: test set)

Table 1: Turnover value of the trading strategies

Trading Strategies	Training set		Test set	
	TurnoverDollars ($)	TurnoverUnits	TurnoverDollars ($)	TurnoverUnits
Mean-Reversion	59693540	259442	31465052	106554
ARIMA	12324235	50526	7441522	22084
ARIMA+GARCH	11284485	44160	3904009	10030

3.2 Profit&Loss Analysis

From Figure 4, we can see that the mean-reversion strategy outperform at the beginning, but we can see from Figure 2 the SPDR price went up very fast in the period of 2017-07 to 2017-12, it lost a lot of money in this strong up trend period. We can conclude that the mean reversion strategy does not suit a strong market trend. The capital growth of all strategies is steady growth in the training set, this is because of the unused capital growth as we make more money.

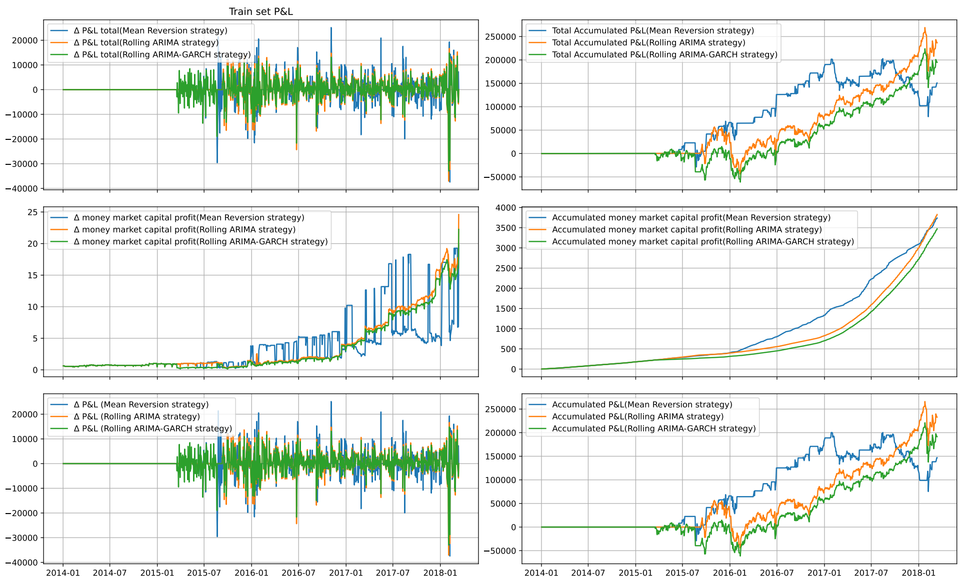

Figure 4: The P&L in the training set

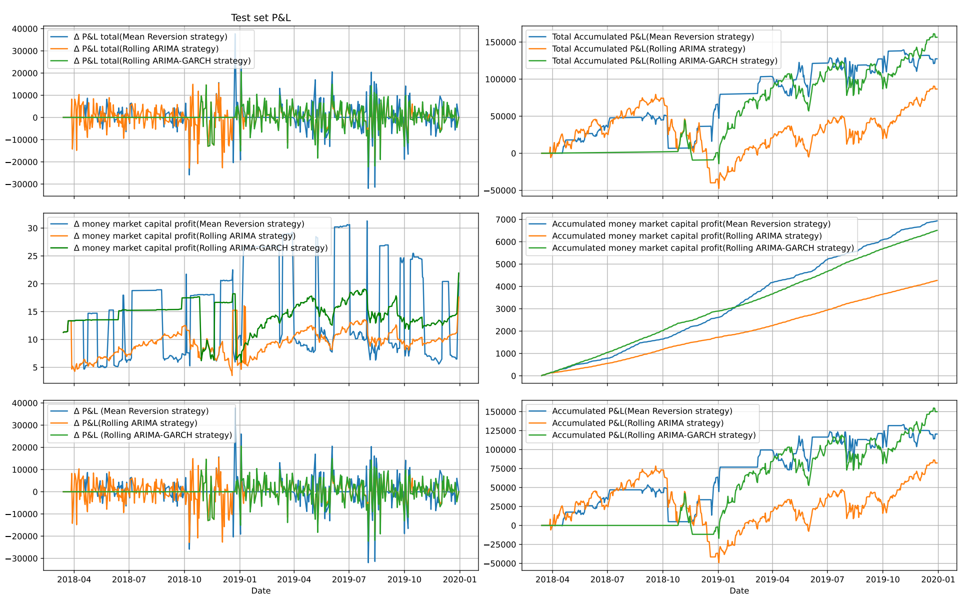

Figure 5: The P&L in the test set

3.3 Performance measure

3.3.1 performance indicators

We will take the excess return ΔV_t as the daily trading P&L. The financial measurement we use includes the Sharpe Ratio, the Sortino Ratio, Maximum Drawdown and the Calmar Ratio, the formulas are[3]:

$$\textbf{Sharpe Ratio} = \frac{\mathbb{E}(\Delta V_t)}{\delta_t} \times \sqrt{252}$$

$$\textbf{Sortino Ratio} = \frac{\mathbb{E}(\Delta V_t)}{\sqrt{\frac{1}{N-1}\sum_{t=1}^{N} \Delta V_t^2}} \times \sqrt{252}$$

$$\textbf{Maximum Drawdown} = \frac{\textbf{MinThroughValue - MaxPeakValue}}{\textbf{MaxPeakValue}}$$

$$\textbf{Calmar Ratio} = \frac{\mathbb{E}(\Delta V_t)}{\textbf{Maximum Drawdown}} \times 252$$

where δ_t is the standard deviation of the daily trading P&L.

Table 2: Strategies Performance

Trading Strategies	Training set				Test set			
	Sharpe Ratio	Sortino Ratio	Maximum Drawdown (%)	Calmar ratio	Sharpe Ratio	Sortino Ratio	Maximum Drawdown (%)	Calmar ratio
Mean-Reversion	0.502	0.502	-62.4	-56241	0.769	0.768	-95.1	-70447
ARIMA	0.832	0.830	-178.7	-31092	0.541	0.540	-162.6	-28175
ARIMA+GARCH	0.723	0.722	-37.7	-12249	1.159	1.156	-140.1	-59545

The maximum DrawDown is calculated using cumulative P&L, I found there is no meaning to calculate daily P&L DrawDown

Figure 2 shows the annualized Sharpe ratio of three strategies varies in the training set and the test set, both Mean-reversion strategy and rolling ARIMA-GARCH strategy have a large increased Sharpe ratio compared to the training set, but the ARIMA strategy have very small Sharpe ratio in the test set. Because the strategies are leveraged, and we only long or short 65% of our V_t every time, we allow profit to drop below 0. Even if we set stop loss condition, two rolling ARIMA strategies still show very bad Maximum Drawdown on both the training and the test set. Comparing the rolling ARIMA strategy and the rolling ARIMA-GARCH strategy, the Maximum Drawdown of the ARIMA-GARCH is significantly small compared to the other one in the training set, which means the GARCH model is a good measure of volatility, but in the test set it does not show good results, so maybe the model is overfitting.

[3] from week 5 notes

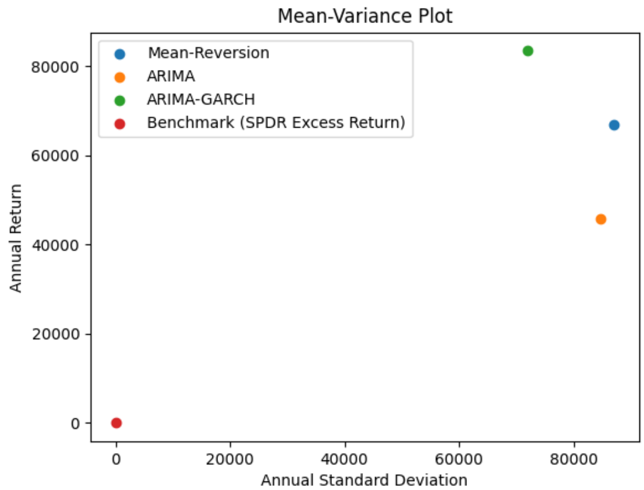

Figure 6: Mean versus Variance of Three strategies

Based on Figure 6, the rolling ARIMA-GARCH strategy shows a higher mean and lower standard deviation compared to the other two strategies, so there is no linear combination of trading strategies with a higher Sharpe Ratio than the rolling ARIMA-GARCH strategy. But If we invest 60% of our capital into the rolling ARIMA strategy and 40% of our capital into the Mean-Reversion strategy, we can get a Sharpe ratio of 0.772[4] which is slightly higher than the original Mean-reversion strategy's Sharpe ratio.

3.3.2 Drawdown Analysis

From the Drawdown figure, we can see that all three strategies' Drawdowns are highly correlated to the SPDR 90-day's historic volatility. The Drawdown of the rolling ARIMA strategy reaches to the highest as the volatility reaches the maximum. For the Mean-Reversion strategy and the rolling ARIMA-GARCH strategy, the Drawdown reaches to the maximum in the middle of uptrend volatility and back to zero when the volatility went to the maximum, so we can say that these two strategies have predicting power on high volatility. One way of using the volatility to control margin is when the volatility is at the downtrend part, we increase our margin, when the volatility is in the uptrend part we gradually reduce our margin to minimize the loss. This could reduce the overall volatility of the strategy but may also influence the total profit you will earn.

[4] 3 significant figures

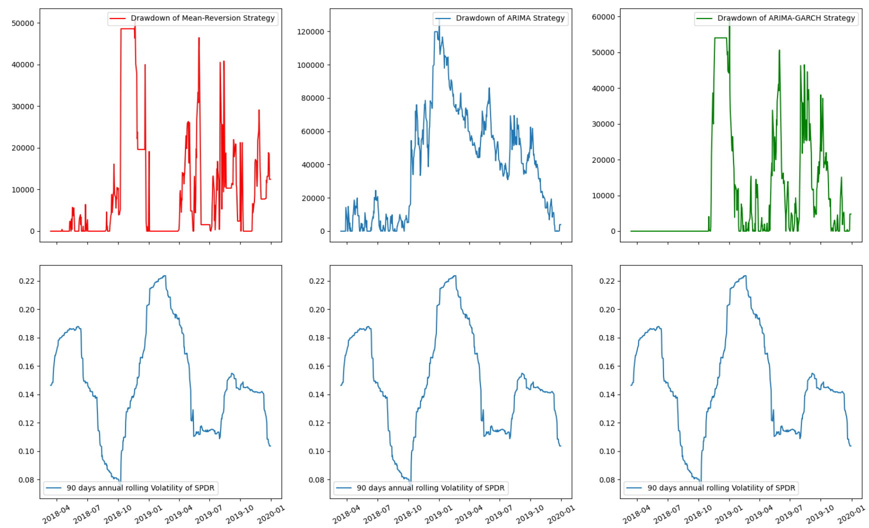

Figure 7: DrawDown for all strategies

4 Discussion and Conclusion

In this report, I evaluate a Mean-Reversion strategy, and two ARIMA-based forecasting strategies, I use performance indicators to analyse the strategies and conclude that the rolling ARIMA-GARCH strategy gets the best result. I found that the GARCH model has an early warning of high volatility. In some cases, the performance even gets worse indicating that the trained models are being over-fitted. As for the Mean-reversion strategy, it performs badly during some strong trends. There is still room for improvement:

- We fixed the amount of margin we can use in each time period, further development should consider adding different levels of signals to control the margin we should use. We can also consider combining two strategies' signals together to make a trading decision.
- Both the training set and the test set are in an uptrend, so the performance of the strategies is unclear in the volatile market or the downtrend market.

References

Bollerslev, Tim. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics* **31**(3) 307–327.

Bollinger, John. 2001. *Bollinger on Bollinger bands*. McGraw-Hill.

Box, George EP, Gwilym M Jenkins, Gregory C Reinsel, Greta M Ljung. 2015. *Time series analysis: forecasting and control*. John Wiley & Sons.

Chio, Pat Tong. 2022. A comparative study of the macd-base trading strategies: evidence from the us stock market. *arXiv preprint arXiv:2206.12282* .

Dash, Rajashree, Pradipta Kishore Dash. 2016. A hybrid stock trading framework integrating technical analysis with machine learning techniques. *The Journal of Finance and Data Science* **2**(1) 42–57.

Kaur, Parminder, Ravi Singla, et al. 2022. Modelling and forecasting nifty 50 using hybrid arima" garch model. *The Review of Finance and Banking* **14**(1) 7–20.

Nison, Steve. 2001. *Japanese candlestick charting techniques: a contemporary guide to the ancient investment techniques of the Far East*. Penguin.

Shumway, Robert H, David S Stoffer, David S Stoffer. 2000. *Time series analysis and its applications*, vol. 3. Springer.

Siami-Namini, Sima, Akbar Siami Namin. 2018. Forecasting economics and financial time series: Arima vs. lstm. *arXiv preprint arXiv:1803.06386* .

Vo, Nguyen, Robert Ślepaczuk. 2022. Applying hybrid arima-sgarch in algorithmic investment strategies on s&p500 index. *Entropy* **24**(2) 158.