Stockholms universitet

The predictive power of two different trading strategies on intraday Bitcoin data

Alexandros Dalaklis

Kandidatuppsats i matematisk statistik
Bachelor Thesis in Mathematical Statistics

Kandidatuppsats 2023:18
Matematisk statistik
September 2023

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Stockholm University

Mathematical Statistics
Stockholm University
Bachelor Thesis **2023:18**
http://www.math.su.se

The predictive power of two different trading strategies on intraday Bitcoin data

Alexandros Dalakis*

September 2023

Abstract

In this study, we investigated the predictive power of technical analysis involving a candlestick pattern strategy and an advanced strategy incorporating multiple indicators on intraday Bitcoin price movements from 2021 to 2023. Our analysis delved into assesing forecasting capabilities focusing on trend reversals, and assessing their profitability, aiming to discern potential market inefficiencies. Results showed that basic candlestick patterns did not have strong predictive power but were slightly profitable, although with unsatisfactory metrics. In stark contrast, the advanced strategy had strong predictive power with results ranging from 3.30 to 4.18 standard deviations away from the null hypothesis and presented a notable 79.2% portfolio return over 901 days, significantly outperforming the 0.19% portfolio return of the buy-and-hold benchmark. Given the vast difference in predictive power between the two strategies, this study highlights the potential for more sophisticated methodologies in exploiting short-term market inefficiencies, specifically in Bitcoin trading.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: dalaklis@gmail.com. Supervisor: Kristofer Lindensjö & Taras Bodnar.

Acknowledgements

I would like to thank my supervisors, Kristofer Lindensjö and Taras Bodnar, for their guidance and their invaluable advice that I received during this project. I would also like to thank my family, especially my parents, who motivated and inspired me to persist.

Contents

1 Introduction

1.1 Background

There are primarily two investment methodologies in stocks and other assets called fundamental analysis and technical analysis. Fundamental analysis involves evaluating a company's or asset's intrinsic value by examining economic and financial factors, including the industry conditions and future growth prospects. Technical analysis involves predicting future prices of stocks or other assets based on past price action and volume. Technical analysts (or chartists) use chart patterns, and indicators (different statistics based on price) to identify trends and to predict future price activity.

An important principle of technical analysis is that all information about an asset such as earning and future performance is reflected in the price [1]. Therefore, it is not essential for a technical analyst to have any fundamental information at all about the asset traded. However, there are investment strategies that rely both on fundamental analysis and technical analysis and the former is perhaps better for long-term investing while the latter is more suitable for short-term investing [2].

1.2 The Efficient Market Hypothesis

The efficient market hypothesis (EMH) was developed by economist Eugene Fama in the 1960s, who later won a Nobel Prize in Economics in 2013. The basic concept of the theory is that financial markets are informationally efficient, which suggests that prices on tradeable assets already reflect all known information and instantly change to reflect new information. This implies that it is impossible to consistently achieve higher than average market returns because it is not possible to predict price movements [3].

If the concept of the EMH is true, it then invalidates the whole technical analysis methodology in investing. Worth to note is that the EMH has been heavily discussed and is not universally accepted and that there are different forms of the concept.

The strong form states that no investor can achieve abnormal returns, even insiders. This form is widely rejected. The semi-strong form suggests that all publicly available information is immediately reflected in security prices. Hence, neither technical analysis nor fundamental analysis can be used to achieve abnormal returns. The weak form suggests that all past market prices and data are fully reflected in security prices. Therefore, technical analysis (which relies on the analysis of past prices) cannot be used to achieve abnormal returns.

We can see that the theory in all of its forms, the EMH rejects the use of technical analysis to consistently achieve higher than above average market returns. The reasoning is that once the market discovers a trading pattern using technical analysis that consistently gives above average returns, it will be exploited by the market (traders) until a point where it is not possible anymore to use the pattern for above average returns, thus making that inefficient part of the market efficient [4] [5].

1.3 Research Goal

The goal of this research is to statistically evaluate the predictive capabilities of basic candlestick charting analysis and a more complex strategy incorporating multiple indicators. In addition, the goal is to ascertain whether candlestick patterns or more advanced strategies can yield significant returns or forecast future price movements using Bitcoin prices from 2021 to 2023 and to assess whether the market at that period was weakly efficient or not.

1.4 Previous Studies

Most of the studies that have been done on technical analysis focus on patterns or different indicators that are derived using Japanese candlestick charts. These type of charts will be explained in the methodology section. They were invented in Japan in the 18th century and popularised in America and the West by Steve Nison in 1989. Nison presents a thourough investigation of these charting methods and discusses how they can be used as a tool for predicting market movements such as potential trend reversals. He does not argue that the patterns provide a foolproof method of predictions, but rather argues that they are a valuable tool that can be used in conjunction with other tools to form a trading strategy [6].

Caginalp and Laurent (1998) showed in their paper "The predictive power of price patterns" [7], that using candlestick patterns in technical analysis, it is possible to predict trend reversals and also generate profit. They found extremely strong statistical evidence to support their claim price data from S&P stocks between 1992 and 1996. Many subsequent researchers used their statistical methodology and similiar trading patterns to test for predictive ability and profitability in the stock market. Our own study uses their methodology as well.

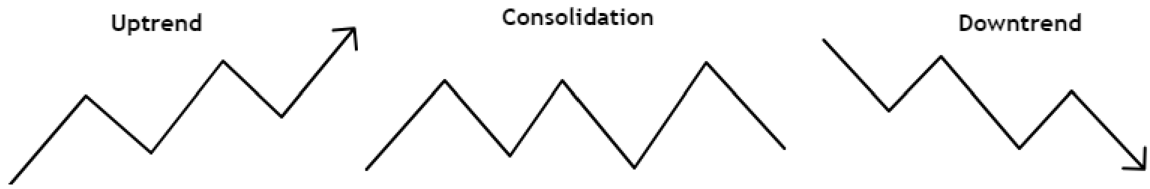
Mashall, Young and Rose (2006) [8] found no significant predictive power and no significant returns using data from 35 stock in the Dow Jones Industrial Average between 1992 to 2001. Horton (2009)[9] had similiar results on data from stocks on the S&P and found that the candlestick patterns did not provide better predictive power than any random pattern. Fock, Klein and Zwergerl (2005) [10] used intraday data from the German stock index (DAX) futures over the period of 2002-2003 testing 19 patterns used in previous papers for profitability. Their study showed that even without transactions cost included, that the results were poor and in most cases not significantly better than a benchmark test with randomized transactions. When they included other technical indicators in conjunction with the candlestick patterns they had higher returns but still not high enough to be statistically significant.

While most studies done in technical analysis post Caginalp & Laurent have not shown significant statistical results when it comes to predictive power and profitability there are few exceptions. Lu, Shiu and Liu (2012) [11] found that some patterns had significant results on the Taiwanese 50 component stocks with a slighly different methodology and a different strategy, where they entered trades on pattern confirmations and held the trades until a reversal pattern confirmation would occur. Y.Ni et all (2020) [12] found that it was possible to beat the market using a strategy based on Bollinger Band trading indicator by entering long positions as price hit the lower Bollinger band and entering short positions while it hit the upper Bollinger band. The Bollinger Band indicator will be used in one of the two strategies in this study and is described in the Appendix.

2 Defining Trends

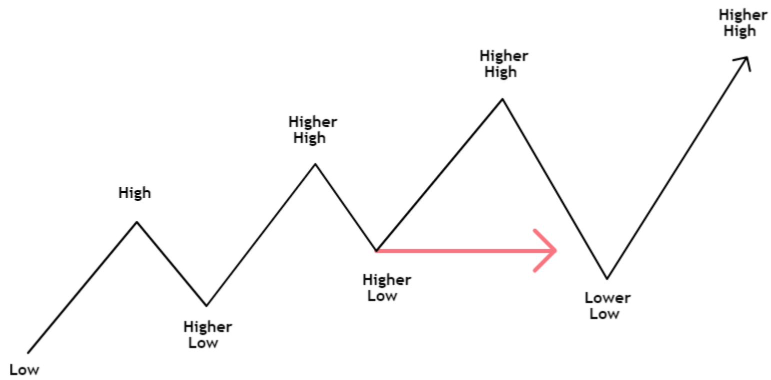
An up trend in the stock market is loosely defined as prices reaching higher highs and higher lows on the time frame of interest whereas a down trend is defined as prices reaching lower highs and lower lows. It should be clarified that there are two types of trends, uptrends and downtrends. When the price is not trending we call it consolidation.

Figure 1: Types of directions in price



During consolidation there is no obvious direction in price. While during trends, we can see larger movements in the leading direction of the price. These up and down trends can be often easily spotted while viewing a historical price chart of any stock or asset. However mathematically defining a trend is a bit more problematic. For example, during uptrends there can be exceptions to the higher highs and higher lows while the stock price is still maintaining an upwards movement over a specified time interval.

Figure 2: Lower Low during an uptrend



2.1 The Caginalp & Laurent MA_3 trend rule

A proposed solution to the problematic nature of mathematically defining a trend that allows for exceptions in price movement is the MA_3 trend rule (Caginalp and Laurent, 1998) which uses a three day moving average over a six day period. Let $P(t)$ denote the closing price on day t, then the three-day moving average is defined as

$$\mathrm{MA}_{3,t} = \frac{1}{3}\sum_{i=0}^{2} P(t-i).$$

We then define an up trend if at least five of the six following inequalities hold

$$\mathrm{MA}_{3,t-6} < \mathrm{MA}_{3,t-5} < ... < \mathrm{MA}_{3,t-1} < \mathrm{MA}_{3,t},$$

and similarly a down trend is defined as

$$\mathrm{MA}_{3,t-6} > \mathrm{MA}_{3,t-5} > ... > \mathrm{MA}_{3,t-1} > \mathrm{MA}_{3,t},$$

if at least five of the six inequalities hold. This is simple trend definition that has some advantages and is not dependent on many parameters. It is important to note that this method of deciding the trend status is only for the very short term. For example, there might be an overall long-term up trend in the market, but using this definition of trend, we can find local short-term down trends within a bigger overall up trend.

2.2 Using Volume Weighted Average Price as an intraday trend indicator

The Volume Weighted Average Price (VWAP) is a trading indicator that gives an average price a security has traded at throughout the day, based on both volume and price. It is calculated as

$$\mathrm{VWAP(t)} = \frac{\sum_{t=0}^{T}(P_t \cdot V_t)}{\sum_{t=0}^{T} V_t},$$

where P_t is the weighted price of the high,low and closing price of the security at time interval t, i.e

$$P_t = \frac{P_{\text{t,high}} + P_{\text{t,low}} + P_{\text{t,close}}}{3},$$

V_t is the volume traded at time interval t and T is the total number of time intervals in the trading day. As the VWAP is an indicator used for intraday trading, these pre-chosen time intervals t are of shorter durations, such as 5 minutes, 15 minutes, or 1 hour. Intervals like one or two days would not be applicable in this context.

If you for example break down a standard 6.5-hour trading day in the U.S Market into 5-minute intervals, then $T = 78$ would represent the whole trading day. This means that the VWAP resets every day when trading opens. The VWAP of Bitcoin has $T = 288$ 5-minute periods since Bitcoin is open for trading at all times.

The VWAP is widely used both by day-traders, by algorithms and by institutional level traders. It is sometimes used as a benchmark for institutional traders that are evaluated on their performance to execute orders at prices better than the VWAP [13]. This means buy (or long) below the VWAP when price levels have stayed above it for a certain period of time and sell or (short) above the VWAP when prices generally has stayed below it.

Hence the VWAP will often act as support when price levels have been trading above as big market participants tend to buy (long) when prices reach near or below the VWAP. Conversely, the VWAP will many times act as resistance if prices have been trading below it for a while as institutional level traders will be inclined to sell (short) when the price nears it again. With this logic then, it is easy to see that it can also be used as an intraday trend indicator, indicating an intraday up trend if prices have been trading above it for a period of time and indicating an intraday downtrend if prices have been trading below it for a certain period of time.

The pros of using the VWAP as an intraday trend indicator is that it is widely used and as mentioned can act as support and resistance, meaning if price is either above or below it tends to stay that way. Another benefit is that it will require only one parameter to decide trend status, specifically the parameter for how many time intervals t prices have been trading above it (for an up trend) or below it (for a down trend).

The cons of using the VWAP as trend indicator is that it is only practically applicable to the lower time frames where the time interval t is one hour or less. While it can theoretically be used in all time frames below one day, there is not much meaning in specifying trend status if price has been above or below the VWAP for 2 or 3 time intervals, e.g if each t represents 12 or 8 hours. Another disadvantage is that if price has moved far from the VWAP due to a strong trend, when and if that strong trend breaks and price starts moving the other direction, that trend change will only be captured once price crosses to other side of the VWAP, which might even be the next day when the VWAP resets.

2.3 Break in trend & Reversal

Break in trend and reversal are sometimes used synonymously and can other times represent two slightly different concepts. Break in trend is when either a down trend or an up trend ceases and the subsequent price movement can be described as either non-trending (also known as consolidation) or as the opposite trend. Reversal is sometimes used to necessarily imply that the subsequent price movement is following an opposite trend. In this text, break in trend and reversals will be used synonymously to imply that the subsequent price direction follows either an opposite trend or a consolidation.

We denote the price at time t as $P(t)$. If $P(t)$ is in a down trend, we consider it a break in trend or a reversal if

$$P(t) \leq P_{AVG}(t+1, t+2, t+3) \qquad \textit{(A1)}$$

This means that when the time interval $t+3$ closes we will know if a break in trend has occured according to $(A1)$. For example, if each time interval t represents one day, we will know at the end of day $t+3$ if the downtrend which time interval t was a part of has been broken. Similarly if the price $P(t)$ is in an up trend, we consider it a break in trend if

$$P(t) \geq P_{AVG}(t+1, t+2, t+3) \qquad \textit{(B1)}$$

This is consistent with the break in trend definition used by Caginalp and Laurent.
To check that the results are robust, we vary the break in trend definition to

$$P(t) \leq P_{AVG}(t+2, t+3, t+4) \qquad \textit{(A2)}$$

$$P(t) \leq P_{AVG}(t+1, t+2, t+3, t+4, t+5) \qquad \textit{(A3)}$$

for the reversals during a down trend, and to

$$P(t) \geq P_{AVG}(t+2, t+3, t+4) \qquad \textit{(B2)}$$

$$P(t) \geq P_{AVG}(t+1, t+2, t+3, t+4, t+5) \qquad \textit{(B3)}$$

for reversals during an up trend. Something to consider is that in their paper, Caginalp and Laurent used the MA_3 trend rule described above together with these break in trend rules for their statistical tests, but the trend rule does not necessarily imply the logic of the break in trend rule and vice versa, e.g we can use their trend rule and use another break in trend rule or keep the break in trend rules $(A1)$...$(B3)$ and test them on a different trend definition.

3 Simulation of trends and reversals

Before using the trend and break in trend rules for real market data strategies, we will see how these rules perform on simulated data. We will be simulating Bernoulli random variables that either assume the value 1 or assume the value 0. If at time t the Bernoulli random variable assumes the value of 1 then the price moves up at time $t + 1$ and if the random variable assumes the value 0 then the price moves down at time $t + 1$. The parameters for the Bernoulli random variables used are $p_1 = 0.99$ and $p_2 = 0.01$ and $p_3 = 0.5$, emulating an uptrend, a down trend and a consolidation in price. Which parameter will be used at first will be simulated with probability $\frac{1}{3}$ for each.

Subsequently, we will be switching the probability parameter based on the exponential distribution with rate parameter $\lambda = 0.05$. Therefore the mean time to switch probability parameter will be $\frac{1}{\lambda} = 20$ time units. The purpose of this is to see if the trend definition is good enough to indeed capture trends in price movement which contain some random fluctuations but where trends are expected to last a significant period of time. When there is a switch of parameter it will switch to one of the other two parameters with probability 0.5.

Furthermore, the price change from time t to time $t+1$ will be simulated from a normal distribution. Let M denote price changes, then

$$M \sim \mathcal{N}(\mu\xi_t, \sigma),$$

where $\mu = 0.0245$ and $\sigma = 0.03$ are fixed and

$$\xi_t = \begin{cases} 1.0367 & \text{if } x_t = 1 \\ -1 & \text{if } x_t = 0, \end{cases}$$

and where x_t is the outcome of the Bernoulli random variable described above. That is, $X \sim Be(p)$ with $p = \{p_1, p_2, p_3\}$. Effectively, the mean parameters used for the up and down price changes are taken from historical daily price changes of Bitcoin from Yahoo Finance over the period 2014-09-17 to 2023-05-17 totaling 3165 days or 3164 daily price changes. The daily price changes were split into positive (up) and negative (down). Then the mean for the two categories were $\mu_{up} = 0.0254$ for the positive (up moves) and $\mu_{down} = 0.0245$ for the negative (down moves).

This will be an uptrending model for n amount of total up and down price changes where we assume that on average there are as many price changes up as there are down. We can see this the limit

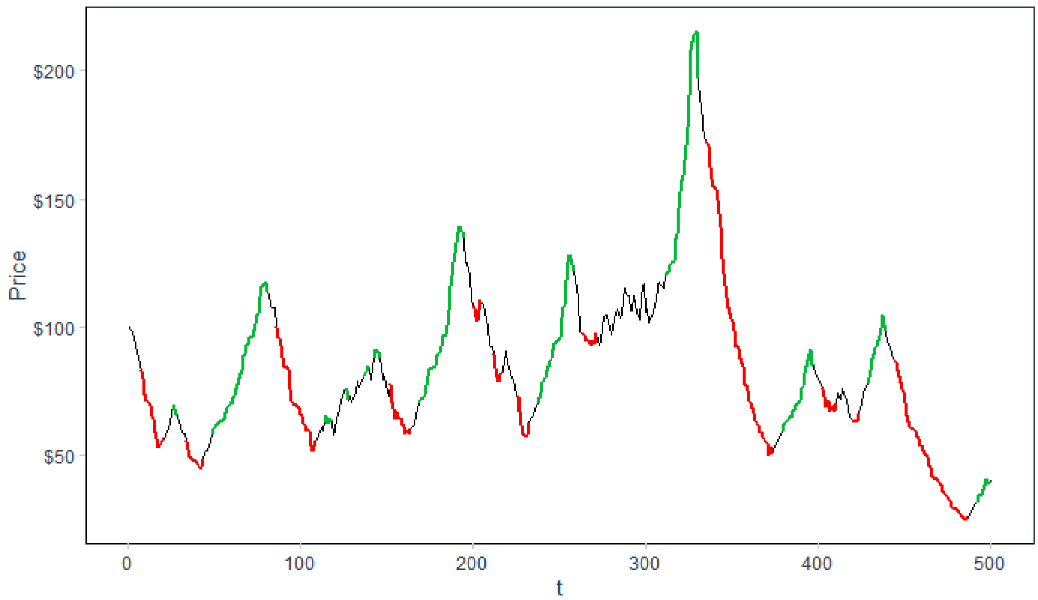
$$\lim_{n\to\infty} (1 + 0.0254)^{\frac{n}{2}} (1 - 0.0245)^{\frac{n}{2}} = \infty,$$

results in an up trending model in the long term, while using the same mean for both up and down moves results in a down trending model in the long term.

3.1 Simulation Results

Perhaps the best performance test for the trend definition on simulated data is simply plotting the data and see how well it depicts trends and no trends.

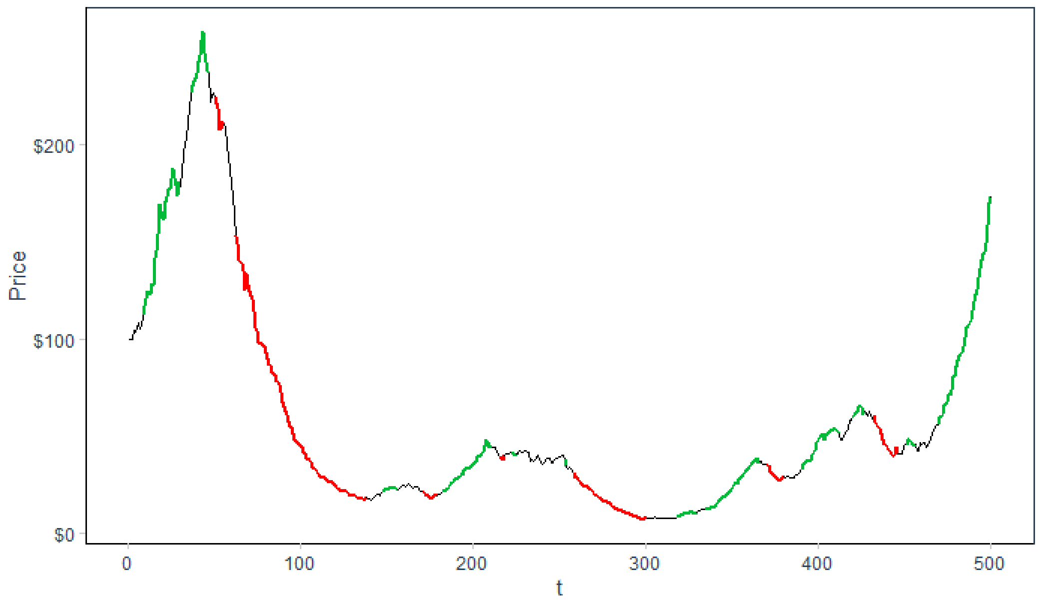
Figure 3: Capturing trends with simulated prices



In Figure 3 we can see the price movement for 500 time units containing up trends, down trends and consolidations. The green line represents up trends as per the MA_3 trend definition. Similarly the red lines represent down trends and the black lines no-trends. We can see that both trends and no-trends are captured very well on simulated data.

In Figure 4 we can see another outcome of the simulation with the exact same parameters. We see once gain that both up and down trends are generally captured well - but not perfect. If we for example observe the first green line on the chart, i.e the first time an up trend appears, we can see at the end of it that there is a very small retracement to the downside and then it continues to the upside to form the peak which is in green. Between the peak which is in green, and the first green line there is a segment which is in black, that is interpreted as a no-trend according to our MA_3 rule while the inclination is high and price is clearly uptrending. Overall though, we can see that on another outcome of simulated data that trends are generally captured very well.

Figure 4: Another outcome of the simulation



To test the break in trend rules ($A1$) and ($B1$) we cannot use plots but will need to test some statistical properties. We will use the break in trend definition to first test if we can accurately infer where and when the probability parameter had switched while the price movement was being on a downtrend according to the definition used.

The benchmark comparison that will be used will be the most naive guessing where with probability $\lambda = 0.05$ that at each time unit t a parameter switch has occured. It can be expected that on average, this naive guessing will be right $\frac{1}{20}$ of the time. Out of $n = 10000$ simulated time units and prices, $n_D = 3487$ were classified as being a on a downtrend. Simulation results using the uniform distribution and number of $n_D = 3487$ gives a result of 10 correct guesses out of 183 guesses in total or 5.18% which is close enough to λ.

To repeat, the break in trend rule will classify that a parameter switch has occured at time t on the simulated data if the following two conditions hold

- $P(t)$ is classified as a downtrend
- $P(t) \leq P_{AVG}(t+1, t+2, t+3)$,

for a reversal to the upside or the following two conditions hold

- $P(t)$ is classified as an uptrend
- $P(t) \geq P_{AVG}(t+1, t+2, t+3)$.

for a reversal to the downside. Testing on the same data as above, we have $n_D = 3487$ time units that were classified as a down trend. Using the above rule, 140 correct predictions were made and 254 incorrect which gives a 35.53% success ratio which is significantly higher than the benchmark test. If we use the classes *switch* and *no-switch* abbreviated S and NS respectively we can get a better picture of the performance in the confusion matrices below.

Predicted Class	Actual Class: S	Actual Class: NS
S	140 72%	254 8%
NS	55 28%	3038 92%

Figure 5: Switch classifier on a downtrend

Predicted Class	Actual Class: S	Actual Class: NS
S	122 71%	277 8%
NS	49 29%	3171 92%

Figure 6: Switch classifier on an uptrend

The diagonal from the top left corner to the bottom right represent the correct predicted class while the other two values show an incorrect predicted class. Looking at the left figure first which represents data on downtrends 189 parameter switches in total. Of these, 136 were correctly classified also known as TP (true positive) and 53 of these were incorrectly classified, also known as FN (false negative). The recall of the test, also known as sensitivity is then calculated as

$$\text{Recall}_{down} = \frac{TP}{TP + FN} = \frac{140}{140 + 55} = 0.7179.$$

The accuracy of the test is

$$\text{Accuracy}_{down} = \frac{TP + TN}{TP + TN + FP + FN} = \frac{140 + 3038}{140 + 3038 + 254 + 55} = 0.9114,$$

where TN and FP are the true negatives and false positives. Lastly the precision of the test has already been mentioned and is

$$\text{Precision}_{down} = \frac{TP}{TP + FP} = \frac{140}{140 + 254} = 0.3553.$$

On the uptrending prices the results are the following. Out of $n = 10000$ simulated time units there were $n_U = 3619$ classified as an uptrend. We can se the results on the right hand side figure. $\text{Recall}_{up} = 0.7135$, $\text{Accuracy}_{up} = 0.9099$ and $\text{Precision}_{up} = 0.3058$. The Recall and Accuracy scores are similar to the downtrending ones but we can observe a significantly worse result on Precision.

3.2 Analysis

The purpose of simulation was to examine if the MA_3 trend rule and the $(A1)$ and $(B1)$ break in trend rules are appropriate to use for our purposes of capturing short-term trends and trend reversals or if the rules to a greater extent capture random fluctuations. Our result show that even on simulated data that these rules are able to identify trends and reversals and seem fit to use as an indicator for real-world data.

While the model generally seems to perform well on simulated data, its practical implications could be discussed. Given the high amount of false positives, not comparing to the benchmark used, traders might end up making premature trades if relying too much on this model, that is if the results had been similar on real data. On the other hand, the model's high recall (or sensitivity) could be leveraged to identify potential trend break zones, which can then be used together with other technical indicators to arrive at a trading decision.

There are some limitations to consider. The different of variations of break in trend rules were not used in the simulations, only $(A1)$ and $(B1)$ which diminishes robustness in the tests. It might be valuable to further delve into how these variations influence predictive power and if there exist optimal definitions that maximize accuracy and minimizes false positives.

It can also be problematic using the same mean for the up and down moves and we will illustrate that with an example. Lets say we use 1% for both up and down moves and set the price at time t to $P(t) = 100$. If the price makes a down move and two up moves, the price would read 99 at $t = 1$,99.99 at $t = 2$ and 100.9899 at time $t = 3$. The average of these three prices is $P_{AVG}(t+1, t+2, t+3) = 99.9933 \leq P(t)$, meaning even two up moves out of three total moves does not qualify as a trend reversal according to our definition. A slightly bigger up move would seem to fix that, and its not entirely unreasonable since most markets are uptrending in the long term. However a slightly bigger up move would instead negatively impact the reversals on uptrending prices $(B1)$ where two down moves and one up move would not be a successful reversal, which is perhaps the reason why the Precision score was significantly worse in uptrending data, given we used a slightly bigger mean for the up moves taken from Bitcoins historical daily price changes.

4 Methodology

4.1 Japanese Candlestick charts

Candlestick charts are a type of financial chart used to describe price movements of a security, derivative, or currency. They were developed by the Japanese rice trader Munehisa Homma in the 18th century [6] . He noticed there was a relationship between supply and demand and that emotions played a significant role in price movement and was able to profit from his discovery. Candlestick charts provide more information than a simple line chart and are a powerful tool to visualize the four data points for each time period, i.e the open, high, low, and close prices within a certain time period. A line chart would have one data point for each time period and that would be the closing price.

Each candlestick on the chart represents a prespecified time period. This could be any duration, e.g 5 minutes, 4 hour, 1 day, or even 1 week. The rectangular shape on the chart is called the body of the candle and represents the opening and closing prices, while the lines protruding from it are called wicks or shadows and indicate the high and low prices for that time period.

For a bullish candlestick (a price increase), the bottom of the body represents the opening price and the top represents the closing price. For a bearish candlestick (a price decrease), the top of the body represents the opening price, and the bottom represents the closing price.

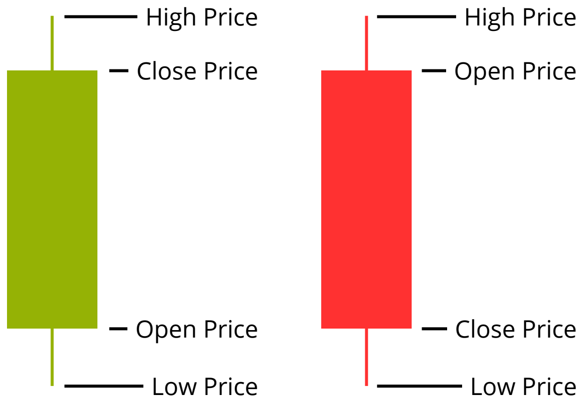


Figure 7: An up closing "bullish" candlestick and a down closing "bearish" candlestick

Candlesticks are color-coded to make it easy to understand the price movement direction at a glance. Usually, if the closing price is higher than the opening price, the candlestick will be colored in green or hollow, indicating it is bullish. Conversely, if the closing price is lower than the opening price, the candlestick will be filled with a different color usually red or black, indicating it is bearish.

The terminology "bullish" and "bearish" come from the way these two animals attack their target. A bull lunges its horns up into the air when attacking, which is symbolizes of a rising market. When traders or investors say that they are bullish, they express their belief that a particular asset or market is going to rise in price. Conversely a bear swipes down with its paw when attacking which symbolizes a declining market. When traders say they are bearish on the market they except prices to fall. These two terms have been used for more than one century in the financial markets and will also be used throughout this study.

The length of the body and the wicks can provide further insight into the market sentiment. Long bodies suggest strong buying or selling pressure. Short bodies suggest little price movement and are often treated as a consolidation pattern.

The wicks, on the other hand, represent the price extremes for the period. A long upper wick shows that buyers pushed prices up during the period but sellers brought the price back down to close near the open. This can be a sign of a potential bearish reversal, meaning prices are expected to fall. Conversely, a long lower wick shows that sellers pushed prices lower, but buyers brought the price back up to close near the open, signaling a potential bullish reversal, meaning prices are expected to rise.

Candlestick charts are popular among traders because they can help identify market trends and potential reversal points, offering the possibility to anticipate future price action.

4.2 Candlestick patterns

Candlestick patterns can be of different length. We will use three candlestick patterns in our first strategy where the main objective is to test their predictive power for trend reversals. The first pattern is in fact a single candlestick which is called the hammer.

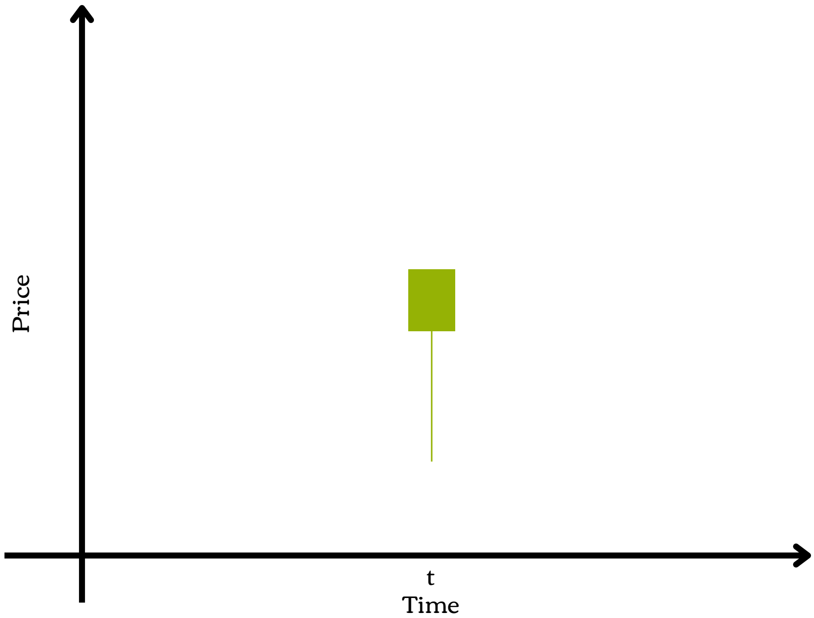


Figure 8: Bullish Hammer candlestick

The bullish hammer which we can see in Figure 8 consists of a small body with a long lower wick and no or a small upper wick. It shows prices went down at some point due to selling pressure but reversed and closed higher than it opened. This shows that the demand in that particular timeframe of the candlestick was stronger and overcame the selling pressure to make a short term reversal within the timeframe of a single candle. The bullish hammer is especially strong if there is no upper wick, because at the moment of closing there has not been a sign of rejection to the upside due to supply, and it is possible and perhaps even likely that the up move will continue on the next candle. We require that the wick is at least twice as large as the body size. The bearish hammer is the exact opposite.

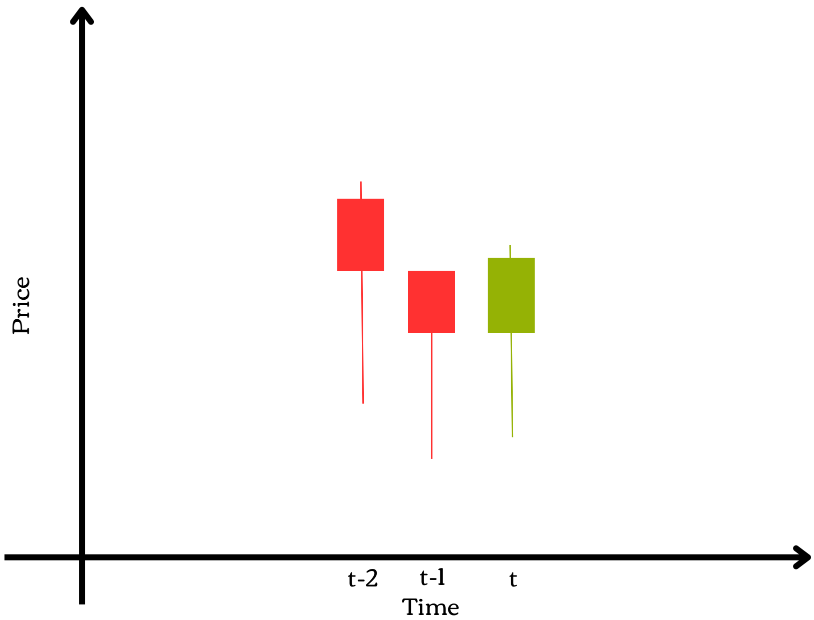


Figure 9: Example of the Three Wicks Pattern

The second pattern we will test consists of three candles and is similar to the hammer which we will call it the Three Wicks Pattern. We have three large wicks in row similar to the hammer. The difference is that we do not require the wicks to be twice as large as the candle body but only 50% larger. Also we do not require the candles to be green, i.e the closing price higher than the opening price of the candle. The logic with this pattern is similar to hammer, sellers now try to push down the price three times but the selling is absorbed by the buyers indicating a possible sellers exhaustion. The bearish version is again the opposite where we instead have three upper wicks on three candles in a row.

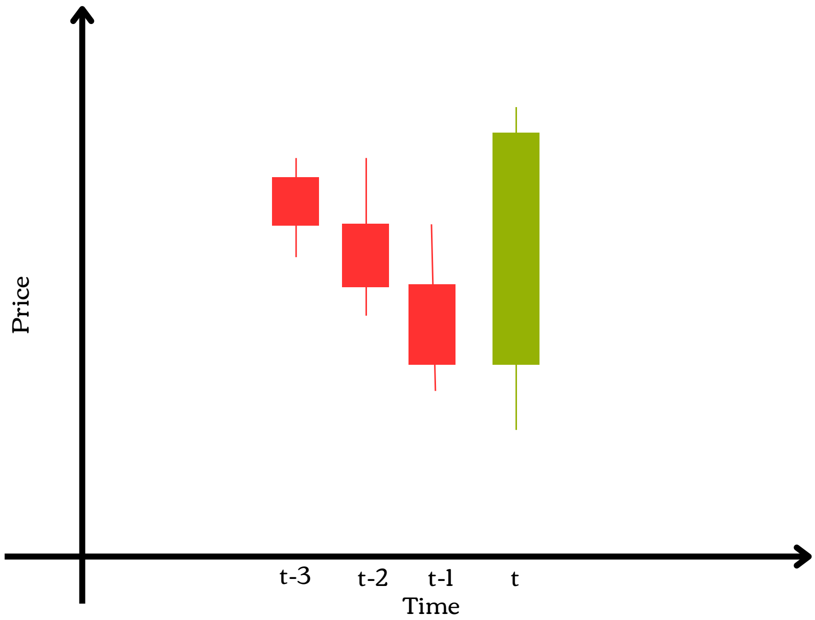


Figure 10: Bullish Three Lines Strike pattern

The third and final pattern is called Three Line Strike. We have three red candles in a row, where each candle closes below the previous one and then we have one green candle that closes above the high of each of the three red candles. This shows a powerful move which might be signaling a reversal on the next few candles.

The bearish version is once again the exact opposite, with three green candles with each closing above the previous one, and the one red candle which closes below the three previous candles low. The exact definitions of these three patterns can be found in the appendix.

For each of the three candlestick patterns we refer to the last candle of the pattern as candle t. It is the closing price at candle t that will be tested for reversals according to the break in trend rules discussed in section 2.3. Moreover, we require the first candle in the pattern to be in the correct trend. For the bullish reversals, where we try to predict a reversal from a downtrend, we require the first candle to be in a downtrend. For example, in the Three Lines Strike pattern, we require candle $t-3$ to be in a downtrend. Since the hammer consists of one candle only, in that case we require candle t to be in a down trend. We have no requirements for the candlesticks that precede any of our three patterns.

4.3 Volume SMA14

Another indicator we will use in confluence with the candlestick patterns is the volume indicator. The volume indicator shows how many units of Bitcoin where bought and sold on every given candlestick. For further strength of our candlestick patterns, we will only consider the patterns that had a significiant spike in volume. The rationale of this is that a pattern on its own, traded with an insignificant amount of volume would not be as indicative of a reversal since that pattern is more likely to have been a result of random noise, than the same pattern but with a significant spike in volume. E.g if we know that that the bullish hammer candle had a significant spike in volume, we know for a fact that many buyers stepped in and pushed the price up which suggesting increased demand at those price levels. This will be the basis for our first strategy.

We will be comparing the volume of candle t (the last candle of the pattern) to the average volume of the 14 candles prior and requiring a 50% spike in volume, i.e

$$Volume(t) > 1.5 \sum_{i=1}^{i=14} \frac{Volume(t-i)}{14}.$$

This is the last condition we will use for the candlestick patterns strategy.

4.4 Predicting Trend Reversals

4.4.1 Classical Candlestick Pattern Strategy

In summary we have three conditions for trying to predict a price reversal to the upside with our first strategy which is based on classical candlestick patterns:

- One of the three discussed candlestick patterns is present
- The first candle of the pattern is on a downtrend according to MA_3 trend definition
- The last candle of the pattern has at least 50% higher volume than the average of the 14 previous candles.

In other words, when these three conditions coincide at time interval t, they collectively define our candlestick pattern strategy. Initially, we will evaluate this strategy for predicting reversals and subsequently assess its profitability. To further clarify, if candle t, which is the last candle of any of the three candlestick patterns fullfills the above three conditions, then candle t will be part of the classical candlestick pattern strategy where the closing price at time t, $P(t)$, will be tested as a reversal point against the break in trend rules discussed in section 2.3. Furthermore, the same candle or time interval t will be considered a long (buy) or a short (sell) signal in testing for profitability. A long signal at time t means a long order will be executed at time t for the closing price $P(t)$ and the equivalent applies to a short order.

While we will be discussing the case where price is on a downtrend, the analogous applies to price reversals on an uptrend. Specifically, we then use the bearish version of the candlestick patterns, and we require that the first candle of the pattern is on an uptrend.

4.4.2 Advanced multi-indicator strategy with layered trends

A more recent approach to trading, specifically to algorithmic trading is the use of several indicators to generate a trading strategy [14]. In the advanced multi-indicator strategy we will be using the VWAP for an intraday trend confirmation, where 15 closed candles in a row above the VWAP will indicate that we are on a uptrend and 15 candle closes in a row below the VWAP will indicate that we are on a downtrend. Important to note is that a lot of the times neither condition will be fullfilled and we are then neither in a uptrend nor a downtrend but in a consolidation phase.

When we are on an uptrend according to the VWAP rule, we will be looking for a shorter-term downtrend according to the MA_3 rule to try to predict a reversal to the upside to assess predictive power while also consider it a long (buy) signal to evaluate profitability. The strategy also uses the two indicators Bollinger Bands and RSI (see the Appendix for an explanation on these and how they are calculated).

On an uptrend we are looking for a candle low below the lower Bollinger Band which is 2 standard deviations below the simple moving average (SMA) of the closing price for the last 14 time points or candles, i.e $n = 14$ will be used. Furthermore, we will need an RSI below 40 wich a RSI lookback parameter of $n = 14$ and as already discussed a downtrend according to the MA_3 trend rule.

The logic of this strategy is to buy or long when the price is on an intraday uptrend but has deviated more than 2 standard deviations below its 14-period SMA, i.e price touching the lower Bollinger Band, while also it is slightly oversold with an RSI of 40 and is on a very short-term downtrend.

For trying to predict a reversal to the downside and finding a short (sell) signal we will be looking for an intraday downtrend as per the VWAP rule above, and a candle high above the upper Bollinger Band while the RSI value is above 60 which indicates it has been slightly overbought in the 14 last time periods or candles and lastly, an uptrend according the the MA_3 rule.

In summary, when we are looking for reversal to the upside (bullish reversal) with the advanced strategy we need the following conditions fullfilled:

- The short-term trend is down (up) according to the MA_3 rule
- The intraday trend is up (down) according to the VWAP trend rule

- Price has touched the lower (upper) Bollinger Band
- The RSI is less or equal to 40 (above or equal to 60)

When these conditions coincide at the same time interval t, they collectively define the advanced strategy which will be used to try to predict short term break in trends to the upside and as a long signal to assess profitability. The equivalent conditions that define when we use the advanced strategy to try to predict break in trends to the downside and look for short (sell) signals are shown in parenthesis.

With this strategy we use both trend rules together and require them to show the opposite trends. An example of why layered opposite trends are employed is the following. Consider a trader who expects the market to be in a strong up trend on the daily time frame and is interested in long positions. The trader might want to see a retracement of price, and possibly a down trend on a lower time frame as to execute the long order on a more favourable price. Similarly, in this strategy we use the VWAP intraday rule to indiciate whether we look for longs or shorts, and the opposite MA_3 trend together with the other discussed indicators for the the exact timing of executing the trade.

4.4.3 Testing for predictive power

In the test for predictive power, we will be using a binomial distribution approximated by the normal distribution. The statistical test will be performed by a one-sided Z-test using a 95% confidence level. The first step in our test for predictive power is to determine the overall probability where the condition $(A1)$ is occuring among all t's that are in a downtrend according to the MA_3 trend rule. Denote this estimate $\hat{p_0}$, then

$$\hat{p_0} = \frac{n_D}{n_A},$$

where n_D is the number of times the reversal condition $(A1)$ is appearing among all points t that are in a downtrend, and n_A the total number of times t is in a down trend. For a very large sample size which will be used, we can consider that the estimate $\hat{p_0}$ assumes its hypothetical mean. We will therefore assume $\hat{p_0} = p_0$ and perform the statistical test from the perspective of a single sample population. Next, we determine the number n of points t that are in a downtrend and also part of a strategy. Within this subgroup we determine the fraction p, and the number np for which the reversal condition $(A1)$ is true.

The standard deviation of the mean is given by

$$\sigma = \sqrt{np_0(1-p_0)}.$$

The right sided hypothesis test can then be described as

$$H_0 : p = p_0$$
$$H_1 : p > p_0.$$

We can then calculate the $Z-$score of the test as

$$Z = \frac{n(p - p_0)}{\sigma},$$

where np is the number of actual trend reversals while the the strategy is present versus the expected number of trend reversals np_0, which is the total number of trend reversals whether the strategy is present or not. We will be comparing the Z-score obtained to the critical value for a one-tailed z-test at $\alpha = 0.05$ which is $z = 1.6445$.

This test will be performed for each of the two strategies seperately, and for each strategy we will also use the variations of the break in trend rules discussed in section 2.3.

4.5 Testing profitability

We have two strategies that we test for predictive power in short-term reversals. For testing profitability of the strategies we also need entry and exit points while we are in a trade. The entry will simply be at the candle close when either strategy gives the confirmation long or short signal. The exit is a bit more complicated as it requires two values, the take profit which is when we exit the trade with profit, meaning the price went in the direction we predicted with our strategy, and the stop loss which is when we exit when price went to the opposite direction of what we predicted and we exit the trade with a loss.

Our take profit and stop loss, i.e our exit points when we are in a trade will be based on an indicator called the Average True Range (ATR) which is a measure of volatilty of the previous n time intervals or candles. The ATR is calculated as follows. First we calculate the True Range (TR) at time t as:

$$TR(t) = \max\left\{P(t)_{\text{High}} - P(t)_{\text{Low}}, \left|P(t)_{\text{High}} - P(t-1)_{\text{Close}}\right|, \left|P(t)_{\text{Low}} - P(t-1)_{\text{Close}}\right|\right\}.$$

Then, the ATR at time t is calculated as an exponential moving average of the TR, specifically,

$$\text{ATR(t)} = \left(1 - \frac{1}{N}\right) \cdot \text{ATR}(t-1) + \frac{1}{N} \cdot TR(t)$$

where common values of the lookback period N is 7 or 14. We will use $N = 7$. The stop loss will be set at 0.75 times the ATR at time t where we have an entry signal from a strategy and the take profit will be set at 2.25 times the ATR. This means for example, if we have a long signal at time point t, the long or buy will be at the closing price of t, $P(t)$, and if the price falls to $P(t) - 0.75ATR(t)$ we will exit the trade with a loss and if price increases to $P(t) + 2.25ATR(t)$ we will exit the trade with a win. This is a risk-vs-reward ratio of 3 and will require on average only an above 25% success rate to show profitability.

Our methodology for the statistical test on profitability will be the most simple, we will be comparing the profitability of the strategy with a buy-and-hold strategy over the entire dataset. This is viable statistical test in itself, as Reschenhofer & Sinkovics (2017) [15] argue that the most effective tool to assess the profitability of a trading strategy is the cumulative plot of its returns to some benchmark. They further argue that there is no need to supplement the graphical results with a statistical test as visual inspection already indicates statistical significance. Yet, this might not always be the case. Lets say that the buy-and-hold strategy generated negative returns due to recession. A "strategy" that would simply involve staying out of the market would perform better with no returns.

Even a strategy that would involve only a few trades with very limited exposure to the market and end up with small returns would seem to perform very well against the negative returns of the buy-and-hold strategy. However that strategy could perform much worse than a buy-and-hold strategy over a longer period of time. An ideal scenario would to make the comparison between two strategies on a large enough data that cover both market recessions and expansions.

4.6 Portfolio metrics

If a trading strategy outperforms a certain benchmark in realized profit, it is essential to consider metrics that take into account volatility or maximum drawdown of the portfolio as it gives a measure to understand the risk taken to achieve that outperformance. Maximum drawdown (MDD) is a risk metric that measures the maximum cumulative loss from a peak to a following bottom which does not necessarily imply that the bottom is below the starting capital of the portfolio. It is defined [16] as

$$MDD = \sup_{t \in [0,T]} \Big[\sup_{s \in [0,t]} P(s) - P(t) \Big],$$

in terms of the maximal price drop. In calculating the Calmar ratio which will be defined later, we will instead use the MDD in terms of percent, which is defined as

$$MDD_{\%} = \sup_{t \in [0,T]} \left[\frac{\sup_{s \in [0,t]} P(s) - P(t)}{\sup_{s \in [0,t]} P(s)} \right].$$

The reason why the maximum drawdown is such an important metric in assessing trading and investment strategies can be illustrated with an extreme example. Let us consider a strategy that outperformed a benchmark but in the meantime had a very high volatility with a maximum drawdown of 90%. That strategy would likely result in total loss of funds in future trials. Also a winning investment strategy that has a very high but not an extreme maximum drawdown, e.g 50% will cut the returns in half everytime a big drawdown such as that occurs, therefore the total returns cannot be considered as reliable.

Daily returns for a portfolio at time t are calculated as

$$R_t = \frac{E_t - E_{t-1}}{E_{t-1}}$$

where E_t and E_{t-1} are the portfolio equities on day t and day $t-1$ respectively. The sample daily mean return is then defined as

$$\bar{R}_d = \frac{1}{N}\sum_{t=1}^{n} R_t,$$

where n is the total number of days in our sample period. The log-returns are calculated as

$$r_t = \ln\left(\frac{E_t}{E_{t-1}}\right).$$

Consider a dataset of realized daily log-returns, $r_1, r_2, ..., r_n$ where n is the total number of days in our sample period. The geometric mean of these daily returns log-returns, denoted $\hat{\mu}_d$ is given by

$$\hat{\mu}_d = e^{\left(\frac{1}{n}\sum_{t=1}^{n} r_t\right)} - 1,$$

and the sample variance is given by

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{t=1}^{n}(R_t - \bar{R}_d)^2.$$

Then the annual returns $\hat{\mu}_A$ are calculated as

$$\hat{\mu}_A = (1 + \hat{\mu}_d)^T - 1,$$

with $T = 365$. The annual volatility is according to A.E Weber (2017) [17] calculated as

$$\hat{\sigma}_A = \sqrt{(\hat{\sigma}^2 + (1 + \hat{\mu}_d)^2)^T - (1 + \hat{\mu}_d)^{2T}}.$$

4.6.1 Sharpe Ratio

The Sharpe ratio measures the excess return (return above the risk-free rate) achieved for each unit of risk taken (as measured by volatility). We will use annualized units to calculate it as

$$\text{Sharpe Ratio} = \frac{R_A - R_f}{\hat{\sigma}_A}$$

where R_A is the annualised return of the portfolio in percent, R_f is the risk-free return and $\hat{\sigma}_A$ the annual volatility as described above. We will set $R_f = 2\%$ in our calculations. A Sharpe ratio of 1 can be considered acceptable or even good but more importantly is to compare Sharpe ratios of other portfolios with similiar returns.

4.6.2 Calmar Ratio

The Calmar Ratio considers the relationship between the annualized return and the maximum drawdown of the portfolio and calculates simply as

$$\text{Calmar Ratio} = \frac{R_A - R_f}{MDD_{\%}}.$$

A Calmar ratio above 1 is considered good while one above 3 is considered excellent.It should be noted that there are different ways of calculating the Calmar ratio. Sometimes the risk-free rate, R_f in the numerator is omitted. Another thing to consider is the time period of the maximum drawdown which can be annualized or represent the whole time period which can be more than one year. In our case we will use annual returns R_A and the Maximal Drawdown will be during the whole time period which will be more than one year of data.

5 Data Overview

In this study we exclusively use data from the BTC/USDT Spot pair on Binance. The pair lets your directly trade Bitcoin for USDT (and vice verca) at current market prices. It is essentially a way to buy or sell Bitcoin in exchange for the digital asset USDT (Tether) that closely mirrors the value of the U.S dollar. Spot trading refers to the purchase or sale of a cryptocurrency for immediate delivery and settlement, as opposed to futures trading where settlement happens at a later date. The reason this pair was chosen is that it is the pair with the highest volume and liquidity among the spot trading pairs on any crypto exchange. This has the benefit of long and short prices in the analysis would accurately reflect real prices a buyer and seller would have received.

Another benefit of studying the asset Bitcoin is its market availability. Cryptocurrency markets operate 24/7, 365 days a year including weekends and holidays. This means there are no gaps in the data, which was double checked. It is also easier to process the data without making errors.

The study focuses on intraday data using the 5-minute time frame, meaning each interval t or candle is 5 minutes long. The reason a lower time frame was used was to be able to increase the frequency of a potential winning strategy but also to have a larger sample size for statistical analysis. The data is from the 1st January 2021 to the 21st of June 2023 which is roughly two and half years or 901 days, totalling 259 427 intervals t or 5-minute candles. The data was pulled from Binance using their API.

5.1 Data Processing

The data was processed using the Python. There were no missing data. The candlestick patterns used could potentially be overlapping, e.g, the one of the candles of the Three

Wicks Pattern could be a Hammer, making it fit into both patterns. In the bullish patterns there were three cases of overlap between these two patterns. The three cases were left in the Hammer pattern category and deleted in the Three Wicks Pattern as to not count them or execute the same strategy twice. There were no overlaps between these two patterns and the third Three Lines Strike pattern. In the bearish version of the patterns, there was one overlap between the Hammer pattern and the Three Wicks Pattern and the duplicate was deleted from the latter. Again, there was no overlap between these two patterns and the Three Lines Strike pattern. Since in the statistical analysis we will only differentiate between the downtrending patterns (bullish reversal patterns) and uptrending patterns (bearish reversals patterns) and not between the patterns themselves, it would make no difference from which of the three patterns a duplicate would be removed.

6 Results

6.1 Reversal Results

Out of 259 427 intervals t, $n_0 = 53124$ were classified as downtrending with the MA_3 rule. Out of these, 28620 satisfied reversal condition $(A1)$, which turns out to be 53.87% of the candles. With the candlestick reversal strategy present, there is n total points of which the fraction p satisfy $(A1)$.

The successful reversals, meaning the reversal condition $(A1)$ was satisfied while our first strategy was present were 56.39% or 128 versus the expected number of reversals of 122, i.e the estimated mean number of reversals, np_0, if the parameter p_0 is the true mean. With a standard deviation of $\sigma = 7.51$ the result is 0.76 standard deviations away from the null hypothesis and is not significant.

The same type of analysis was made by varying the bullish reversal to conditions to $(A2)$ and $(A3)$, and also to the bearish reversal conditions $(B1), (B2)$ and $(B3)$. The results can be summarised best in the following table.

Table 1. Statistics for intraday Bitcoin price reversals using candlestick patterns

	Equation	$A1$	$A2$	$A3$	$B1$	$B2$	$B3$
Overall number	$\frac{n_0 p_0}{n_0}$	$\frac{28620}{53124}$	$\frac{28766}{53124}$	$\frac{28649}{53124}$	$\frac{29471}{53839}$	$\frac{29456}{53839}$	$\frac{29319}{53839}$
Percent	p_0	53.87%	54.15%	53.93%	54.74%	54.17%	54.46%
Percent with strategy	p	56.39%	58.15%	55.95%	53.89%	56.29%	55.09%
Expected number of reversals	$[np_0]$	122	123	122	91	91	91
Actual reversals	np	128	132	127	90	94	92
Standard deviation	$\sigma = \sqrt{np_0(1-p_0)}$	7.51	7.51	7.51	6.43	6.43	6.44
Number of standard deviations from H_0	$Z = \frac{n(p-p_0)}{\sigma}$	0.76	1.21	0.61	-0.22	0.41	0.16
P-value	$1 - \Phi(Z)$	0.2237	0.1132	0.2709	0.5870	0.3412	0.4318

We can see that none of the results were significant or close to our pre-specified 95% confidence level. The predicted bearish reversals with our strategy using reversal condition ($B1$) were lower than the average number of reversals during an uptrend, which signifies a poor result in predictive power.

It is perhaps interesting to note that the three candlestick patterns used performed very differently. The worst performing candlestick pattern was the three line strike with a weight average across all tests of 10 reversals and 18 non-reversals which give a 35.71% reversals rate that is well below a random pattern. The three line strike pattern was also the the most uncommon of the three. The best performing pattern was the three wicks pattern which had 35 reversals and 19 non-reversals which give a reversal rate of 64.81%. However even in isolation this pattern would not give significant results with p-value of 0.064. The hammer pattern had on average 66 reversals and 49 non-reversals or 57.39% which was higher than the mean. We can observe quite a few occurences of the patterns given a large data set, specific candlestick patterns can be rare, even more so when we add extra conditions such as up and down trending and spike in volume.

The results for the advanced strategy were very different. All results were significant on our pre-specified 95% confidence level. The least significant result was the strategy used for bearish reversal with reversal condition ($B1$) which was 3.30 standard deviations away from the null hypothesis with a p-value of **0.00048** and the most significant result was the strategy used on reversal condition ($A2$) with 4.18 standard deviations away from the null hypothesis and a p-value of **0.00001**. The results can be seen on Table 2:

Table 2. Statistics for intraday Bitcoin price reversals using the advanced strategy

	Equation	$A1$	$A2$	$A3$	$B1$	$B2$	$B3$
Overall number	$\frac{n_0 p_0}{n_0}$	$\frac{28620}{53124}$	$\frac{28766}{53124}$	$\frac{28649}{53124}$	$\frac{29471}{53839}$	$\frac{29456}{53839}$	$\frac{29319}{53839}$
Percent	p_0	53.87%	54.15%	53.93%	54.74%	54.17%	54.46%
Percent with strategy	p	58.61%	60.05%	59.57%	59.90%	60.10%	60.20%
Expected number of reversals	$[np_0]$	673	676	674	556	555	553
Actual reversals	np	732	750	744	608	610	611
Standard deviation	$\sigma = \sqrt{np_0(1-p_0)}$	17.62	17.61	17.62	15.86	15.86	15.87
Number of standard deviations from H_0	$Z = \frac{n(p-p_0)}{\sigma}$	3.36	4.18	4.00	3.30	3.45	3.67
P-value	$1 - \Phi(Z)$	**0.00040**	**0.00001**	**0.00003**	**0.00048**	**0.00028**	**0.00012**

With significant results for every reversals condition ($A1$)...($B3$) we can confirm very robust results for the predictive power of the advanced strategy. The strategy has significantly higher chance of predicting very short-term reversals compared to a random strategy. Given these results, it is very likely that it is possible to use this strategy to have above average profits which shows an inefficiancy in the chosen market.

6.2 Profitability results

The first strategy that was based on simple candlestick patterns in confluence with a volume spike generated a profit of 3.98% with 394 trades in 901 days. The annual volatility was at 5.14% and the maximum drawdown −7.85%. Both the Sharpe Ratio and the Calmar ratio were below 0.25 which is considered poor for both metrics. While 3.98% is higher then the buy and hold return at 0.19%, it is insignificantly higher over such a large sample size and with generally poor metrics and lower return than the maximum drawdown we can not conclude that this is a winning strategy.

The advanced strategy generated a portfolio return of 79.2% with 1701 trades in 901 days. The annualized portfolio return was 26.72% , the annualized volatility 15.82% and the maximum drawdown −11.71%. The return of 79.2% can be considered very good comparing it to the chosen benchmark of a buy-and-hold strategy which would have a portfolio return of 0.19% over 901 days. We can also consider it significant with such a large sample size and also considering the large amount of trades taken. A better overview of the advanced strategy versus the buy-and-hold strategy will be best observed by viewing a plot of the returns.

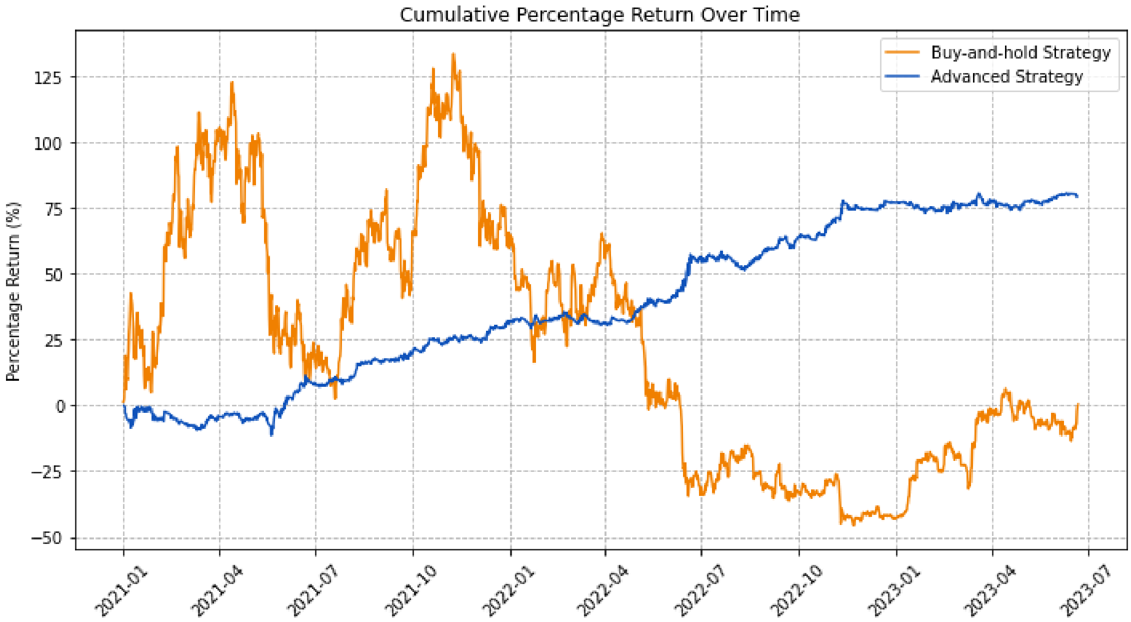


Figure 11: A comparison of two investments strategies

In Figure 11 we can see a big difference between the advanced strategy and the buy-and-hold strategy. While the advanced strategy had higher returns in the end, the buy-and-hold strategy was up much more initially during the phase of expansion and clearly outperformed the advanced strategy. During the phase of expansion of Bitcoin, which aligned in time with the surges in major markets like the Nasdaq and S&P500, the price rose to $69000 which can be seen as the second peak in orange. From that peak, to the low of $15476 which price reached in the end of 2022 was a the maximum drawdown of the buy-and-hold strategy which is a about 77.57%.

We can clearly see, judging from the plot, that Bitcoin was a very volatile asset during the time period captured by the data. While the advanced strategy was outperformed in the expansion phase, it clearly outperformed Bitcoin during the recession, rising steadily with a comparably much lower maximum drawdown. Taking into account that both an expansion phase and a recession phase were present in the data spanning 901 days, we can conclude that the advanced strategy outperformed the buy-and-hold strategy, and we consider the results significant.

The Sharpe ratio was 1.56 and the Calmar ratio 2.11. Both the Sharpe ratio and the Calmar ratio can be considered good to very good. A risk-free return of 2% was chosen as an estimation to calculate the Sharpe and Calmar ratios. The chosen rate might not be entirely correct but it was estimated in comparison with the US 1 year Treasury rate which very close to 0% for the entirety of 2021, increased to about 2.9% in mid 2022 and further increased in 2023.

Another thing to consider is that we have not taken into account trading fees which would have negatively affected the performance of both the strategies but not the benchmark strategy. The first strategy would have very likely shown a negative return and the second would have shown a smaller return then 79.2% and would also affect metrics such as the volatility, maximum drawdown and Sharpe and Calmar ratios. With that being said, the fees for the particular pair traded on the particular exchange, i.e Bitcoin/USDT on Binance had zero fees since mid 2022 till the end of the period on the data set.

7 Conclusion and Discussion

The primary goal of this study was to test the predictive power of two strategies, one based on classical candlestick patterns that initially showed strong results with the Caginalp and Laurent (1998) paper, but subsequently failed to replicate those findings in later studies; and one more advanced strategy that is more in line with how algorithms are set up to trade the markets today.

We found that the strategy based on classical candlestick patterns, even with the addition of volume spikes did not have strong predictive power. This is not surprising but rather consistent with market dynamics, even if these candlestick patterns once held significant predictive power enabling above-market returns, they would inevitably be exploited over time, neutralizing their advantage and eliminating market inefficiencies. Since candlestick patterns are very easy to program and algorithmic trading is estimated to compromise above 80% or even as high as 92% of all equity volume [18], it is definitely a likely scenario.

Nonetheless, only three candlestick patterns were used in the study while the original study by Caginalp and Laurent used eight. Many of the patterns they used which were also used in subsequent studies could not be used on Bitcoin data since many the patterns have gaps from the closing price of interval t to the opening price of $t + 1$. These gaps can occur on in the traditional markets because when they close on a given day, the futures market or high impact news can affect the opening price on the next day, making it higher or lower then the closing price of the previous day. This is not possible on Bitcoin since the market is open at all times. Although we are inclined to believe that candlestick patterns alone lack strong predictive power, it's important to acknowledge the limited number of patterns tested in this study.

The advanced strategy that consisted of several components such as layered trends, showed strong predictive power in all tests; both bullish and bearish and in all variations of the reversal conditions with the least significant result 3.30 standard deviations away from the null hypothesis and the most significant 4.18 standard deviations away from the null hypothesis. With this significant result, and if not taking into account fees and transaction costs, there must exist a trading strategy that can take advantage of the strong predictive power and generate above market returns. Testing profitability was the secondary goal of this study.

The candlestick pattern strategy yielded a portfolio return of 3.98% over 394 trades in a span of 901 days. While this performance surpasses the benchmark buy-and-hold strategy, which generated a return of 0.19%, the results are not deemed significant. The modest advantage over the benchmark is further diminished by subpar Sharpe and Calmar ratios, both falling below 0.25. When accounting for trading fees, the net result would most likely turn negative.

The advanced strategy yielded a portfolio return of 79.2% with 1701 trades in 901 days. Taking into consideration sample size, amount of trades and the fact that the strategy clearly outperformed the buy-and-hold benchmark in a long period of time of time which consisted of both growth and recession phases, we can conclude that the results are indeed significant. This suggests that we have found a market inefficiency. The Sharpe and Calmar ratios which were 1.56 and 2.11 respectively, further affirm our conclusion.

7.1 Futher Studies

Based on the results, there are several areas that merit further research. The first would be to measure how big of an impact trading fees would have on profitability since in our own analysis we did not take them into account. Given that the take profit and the stop loss were set to small multiples of the Average True Range of the last 7 candles, which generally is in close proximity to the entry taken, the average profit of the trades would perhaps not be much higher than the fees paid which would greatly diminish the returns presented. If dimished by a lot, it would perhaps warrant another set of trading rules when to exit a trade, i.e changing the take profit and the stop loss, which is our next area of interest.

Parameter optimization could perhaps improve the strategies by a lot. While the advanced strategy showed promising results, there is potential for enhancing its performance through parameter optimization with or without the use of machine learning. The parameters used the take profit and stop loss of both the strategies were set to the same values. Using the same values for two entirely different strategies is most likely sub-optimal. This was done on purpose to avoid overfitting and overtuning the parameters to find the best results given the data, but which would be hard or almost impossible to replicate on other data.

One additional area deserving further attention is cross-asset validation. It would be insightful to evaluate how the advanced strategy, or one similar to it, performs on other asset classes like forex pairs and market indices such as the S&P and Nasdaq. This could increase the robustness of the strategy and reduce the risk of overfittning. However, it is worth noting that a particular strategy might excel in one asset class and perform poorly on another.

8 Appendix

8.1 Candlestick pattern definitions

Here we define the three candlestick patterns used. We denote the open, high, low and close at time t as o_t,h_t,l_t and c_t.

Definition Hammer
Bullish
$c_t > o_t$
$c_t = h_t$ or $c_t - o_t > 4(h_t - c_t)$
$c_t - o_t < 2(o_t - l_t)$
Bearish
$o_t > c_t$
$c_t = l_t$ or $o_t - c_t > 4(c_t - l_t)$
$o_t - c_t < 2(h_t - o_t)$

Definition Three Wicks Pattern
Bullish
$\min(c_t, o_t) - l_t > 1.5|o_t - c_t|$
$\min(c_{t+1}, o_{t+1}) - l_{t+1} > 1.5|o_{t+1} - c_{t+1}|$
$\min(c_{t+2}, o_{t+2}) - l_{t+2} > 1.5|o_{t+2} - c_{t+2}|$
Bearish
$h_t - \max(c_t, o_t) > 1.5|o_t - c_t|$
$h_{t+1} - \max(c_{t+1}, o_{t+1}) > 1.5|o_{t+1} - c_{t+1}|$
$h_{t+2} - \max(c_{t+2}, o_{t+2}) > 1.5|o_{t+2} - c_{t+2}|$

Definition Three Line Strike
Bullish
$c_t < o_t$
$c_{t+1} < o_{t+1}$
$c_{t+2} < o_{t+2}$
$c_{t+1} < c_t$
$c_{t+2} < c_{t+1}$
$c_{t+3} > \max(h_t, h_{t+1}, h_{t+2})$
Bearish
$c_t > o_t$
$c_{t+1} > o_{t+1}$
$c_{t+2} > o_{t+2}$
$c_{t+1} > c_t$
$c_{t+2} > c_{t+1}$

$c_{t+3} < \min(l_t, l_{t+1}, l_{t+2})$

8.2 Bollinger Bands

Bollinger Bands, created by John Bollinger in the 1980s, are a tool used in technical analysis. They are depicted as three lines on a price chart. The central line represents a moving average of the price, while the two outer lines, or bands, signify a measure of the asset's volatility. The bollinger bands are calculated as follows,

$$\text{Middle Band} = \frac{\sum_{i=1}^{n} P_i}{n},$$

which as mentioned is a a simple moving average (SMA) of the price, usually over $n = 14$ periods (e.g, 14 days). We then calculate the standard deviation as

$$\text{SD} = \sqrt{\frac{\sum_{i=1}^{n}(P_i - \text{Middle Band})^2}{n}}.$$

Then, the upper and lower bands are calculated as,

$$\text{Upper Band} = \text{Middle Band} + (k \cdot \text{SD}),$$

and

$$\text{Lower Band} = \text{Middle Band} - (k \cdot \text{SD}),$$

where $k = 2$ is typically chosen.

8.3 RSI

The RSI is a momentum index invented by Welles Wilder in 1978. To get an intuition of what the RSI is we can break down the calculation into several steps. Let $close_t$ be the closing price of candle t and $close_{t-1}$ be the closing price of candle $t-1$. Then the size of the upward move for candle t is denoted by U_t and the size of the downward move D_t. We then have for $t = 1, ..., n$

$$U_t = \begin{cases} close_t - close_{t-1} & \text{if } close_t > close_{t-1} \\ 0 & \text{if } close_t \leq close_{t-1} \end{cases} \tag{1}$$

$$D_t = \begin{cases} close_{t-1} - close_t & \text{if } close_{t-1} > close_t \\ 0 & \text{if } close_{t-1} \leq close_t. \end{cases} \tag{2}$$

We calculate a simple moving average for the up moves with period α as

$$SMA_{U_{t+1-\alpha}} = \sum_{t=1}^{\alpha} \frac{U_t}{\alpha},$$

and one for the downmoves

$$SMA_{D_{t+1-\alpha}} = \sum_{t=1}^{\alpha} \frac{U_t}{\alpha},$$

for $t \geq \alpha$ where α can also be interpreted as the lookback period for the moving average. The relative strength RS is calculated as the quotient of the two moving averages

$$RS_{t+1-\alpha} = \frac{SMA_{U_{t+1-\alpha}}}{SMA_{D_{t+1-\alpha}}}.$$

Finally the Relative Strength Index, RSI is a normalised RS ranging between 0 and 100

$$RSI_{t+1-\alpha} = 100 - \frac{100}{1 + RS_{t+1-\alpha}}$$

The next values of the RSI can be calculated using the same simple moving average SMA, or as suggest by Wilder [19] using what has come to be known as *Wilders' Smoothed Moving Average (WSMA)* which is calculated for the up moves as

$$WSMA_{U_{t+2-\alpha}} = \frac{WSMA_{U_{t+1-\alpha}} \cdot (\alpha - 1) + U_t}{\alpha}$$

for $t \geq \alpha$ where $WSMA_{U_{t+1-\alpha}}$ is equal to the Simple Moving Average $SMA_{t+1-\alpha}$ if $t = \alpha$, i.e for the first lookback period. We will be using Wilders' Smoothed Moving Average as opposed to the simple moving average and we will also be using a $\alpha = 14$ which on the daily timeframe is a two-week lookback period.

9 References

[1] Malkiel, B.G. (1973, rev 2014). A Random Walk Down Wall Street. W. W. Norton & Company.

[2] Murphy, J.J. (1999). Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications. New York Institute of Finance.

[3] Fama, E.F. (1965). The Behavior of Stock-Market Prices. The Journal of Business, 38(1), 34-105.

[4] Max Jönsson (2016), The Predictive Power of Candlestick Patterns: Department of Economics, Lund University

[5] Elton, et al. (2009). Modern Portfolio Theory and Investment Analysis (9th Edition). Wiley.

[6] Nison, S. (1991). Japanese Candlestick Charting Techniques: A Contemporary Guide to the Ancient Investment Techniques of the Far East (1st Edition). New York Institute of Finance.

[7] Caginalp, G. & Laurent, H. (1998). The Predictive Power of Price Patterns. Applied Mathematical Finance, 5(3-4), 181-205. DOI: 10.1080/135048698334637

[8] Marshall, B.R., Young, M.R., & Rose, L.C. (2006). Candlestick Technical Trading Strategies: Can They Create Value for Investors? Journal of Banking & Finance, 30.

[9] Horton, M.J. (2009). Stars, Crows, and Doji: The Use of Candlesticks in Stock Selection. The Quarterly Review of Economics and Finance, 49, 283–294.

[10] Fock, H., Klein, C., & Zwerger, B. (2005). Performance of Candlestick Analysis on Intraday Futures Data. Journal of Derivatives, 13(1), pg. 28.

[11] Lu, T-H., Shiu, Y-M., & Liu, T-C. (2012). Profitable Candlestick Trading Strategies—The Evidence from a New Perspective. Review of Financial Economics, 21, 63–68.

[12] Ni, Y., Day, M.-Y., Huang, P., & Yu, S.-R. (2020). The Profitability of Bollinger Bands: Evidence from the Constituent Stocks of Taiwan 50. Physica A, 551, 124144.

[13] Madhavan, A. (2002). VWAP Strategies. Transaction Performance, Spring 2002 issue.

[14] Faijareon, C. & Sornil, O. (2019). Evolving and combining technical indicators to generate trading strategies. J. Phys.: Conf. Ser., 1195, 012010. DOI: 10.1088/1742-6596/1195/1/012010, School of Applied Statistics, National Institute of Development Administration, Bangkok, Thailand.

[15] Reschenhofer, E. & Sinkovics, T. (2017). Examining the Profitability of Automatic Trading Strategies with a Focus on Trend Indicators. Quantitative Finance, 17(7), 979-991. DOI: 10.1080/14697688.2016.1262055

[16] Magdon-Ismail, M., Atiya, A., Pratap, A., & Abu-Mostafa, Y. (2004). On the maximum drawdown of a Brownian motion. Journal of Applied Probability, 41(1), 147-161. doi:10.1239/jap/1077134674

[17] Webe, A.E. (2017). Annual Risk Measures and Related Statistics. Journal of Performance Measurement, Spring 2017 issue.

[18] Kissell, R. L. (2021). Algorithmic Trading Methods (Second Edition). Academic Press. DOI: 10.1016/B978-0-12-815630-8.00001-6.

[19] Wilder, J.W. (1978). New Concepts in Technical Trading Systems. Trend Research.