

“Automated trading systems’ evaluation using d-Backtest PS method and WM ranking in financial markets”

AUTHORS	Dimitrios Vezeris  https://orcid.org/0000-0002-2251-5772 Themistoklis Kyrgos Ioannis Karkanis Vasiliki Bizergianidou
ARTICLE INFO	Dimitrios Vezeris, Themistoklis Kyrgos, Ioannis Karkanis and Vasiliki Bizergianidou (2020). Automated trading systems’ evaluation using d-Backtest PS method and WM ranking in financial markets. <i>Investment Management and Financial Innovations</i> , 17(2), 198-215. doi: 10.21511/imfi.17(2).2020.16
DOI	http://dx.doi.org/10.21511/imfi.17(2).2020.16
RELEASED ON	Wednesday, 17 June 2020
RECEIVED ON	Saturday, 25 April 2020
ACCEPTED ON	Friday, 29 May 2020
LICENSE	 This work is licensed under a Creative Commons Attribution 4.0 International License
JOURNAL	"Investment Management and Financial Innovations"
ISSN PRINT	1810-4967
ISSN ONLINE	1812-9358
PUBLISHER	LLC “Consulting Publishing Company “Business Perspectives”
FOUNDER	LLC “Consulting Publishing Company “Business Perspectives”



NUMBER OF REFERENCES

31



NUMBER OF FIGURES

11



NUMBER OF TABLES

8

© The author(s) 2024. This publication is an open access article.



BUSINESS PERSPECTIVES



LLC "CPC "Business Perspectives"
Hryhorii Skovoroda lane, 10,
Sumy, 40022, Ukraine
www.businessperspectives.org

Received on: 25th of April, 2020

Accepted on: 29th of May, 2020

Published on: 17th of June, 2020

© Dimitrios Vezeris, Themistoklis Kyrgos, Ioannis Karkanis, Vasiliki Bizergianidou, 2020

Dimitrios Vezeris, Ph.D., COSMOS4U,
Xanthi, Greece.

Themistoklis Kyrgos, M.Eng.,
COSMOS4U, Xanthi, Greece.

Ioannis Karkanis, B.Sc., COSMOS4U,
Xanthi, Greece.

Vasiliki Bizergianidou, M.Eng.,
COSMOS4U, Xanthi, Greece.



This is an Open Access article, distributed under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license, which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Conflict of interest statement:

Author(s) reported no conflict of interest

Dimitrios Vezeris (Greece), Themistoklis Kyrgos (Greece),
Ioannis Karkanis (Greece), Vasiliki Bizergianidou (Greece)

AUTOMATED TRADING SYSTEMS' EVALUATION USING D-BACKTEST PS METHOD AND WM RANKING IN FINANCIAL MARKETS

Abstract

Given the popularity and propagation of automated trading systems in financial markets among institutional and individual traders in recent decades, this work attempts to compare and evaluate such ten systems based on different popular technical indicators in combination – for the first time – with the d-Backtest PS method for parameter selection. The systems use the technical indicators of Moving Averages (MA), Average Directional Index (ADX), Ichimoku Kinko Hyo, Moving Average Convergence/Divergence (MACD), Parabolic Stop and Reverse (SAR), Pivot, Turtle and Bollinger Bands (BB), and are enhanced by Stop Loss Strategies based on the Average True Range (ATR) indicator. Improvements in the speed of the back-testing computations used by the d-Backtest PS method over weekly intervals allowed examining all systems on a 3.5 years trading period for 7 assets in financial markets, namely EUR/USD, GBP/USD, USD/JPY, USD/CHF, XAU/USD, WTI, and BTC/USD. To evaluate the systems more holistically, a weighted metric is introduced and examined, which, apart from profit, takes into account more factors after normalization like the Sharpe Ratio, the Maximum Drawdown and the Expected Payoff, as well as a newly introduced Extended Profit Margin factor. Among the automated systems examined and evaluated using the weighted metric, the Adaptive Double Moving Average (Ad2MA) system stands out, followed by the Adaptive Pivot (AdPivot), and the Adaptive Average Directional Index (AdADX) systems.

Keywords

algorithmic trading, d-Backtest PS method, FOREX, weighted metrics, extended profit margin

JEL Classification

F31, G17

INTRODUCTION

In the past couple of decades, the scientific and financial community has tackled an ever-increasing amount of work to develop software systems that trade automatically in the financial and stock markets around the world. Thus, these systems have become ubiquitous in the global trading scene, and a need to evaluate them becomes apparent.

The current research examines and attempts to evaluate ten automated trading systems that are based on widely used technical indicators. For this purpose, the d-Backtest PS method is used to select the systems' parameters, and a weighted metric based on five factors regarding the consistency, risk efficiency, and profitability of the systems is introduced.

Taking the above into consideration and in order to evaluate the systems, weekly backtests were conducted on 7 assets of financial markets for a total period of 3.5 years. Along with the d-Backtest PS method, a

more traditional constant-period backtesting method was also used for parameter selection. The results of each backtesting method were compared using a range of weights for each factor of the weighted metric, that produced a classification of the automated trading systems performance.

1. LITERATURE REVIEW

As financial data became more accessible and consistent, and information technology tools became more capable, automated trading systems could utilize various techniques and indicators. Some systems use machine learning techniques to process the time series data of an asset and decide on trading actions (Henrique, Sobreiro, & Kimura, 2018; Booth, Gerding, & McGroarty, 2014), while other systems use the help of technical indicators and rules based on these technical indicators to make decisions when trading an asset (Chong & Ng, 2008; Wilder, 1978). There are also more complex systems that combine two or more technical indicators, including artificial neural networks, fuzzy logic, or other advanced machine learning techniques (Silva et al., 2014; Osunbor & Egwali, 2016). Apart from technical data and indicators, automated trading systems can also utilize information from outside the financial markets captured in news articles or social media trends (Azhikodan, Bhat, & Jadhav, 2019).

In parallel with the development of automated trading system, there has been a search for better means for evaluating such systems. The most commonly used criterion for evaluating the trading strategies has been the return on investment or the similar growth rate and profit rate (Y.-H. Chou, Kuo, Chen, & H.-C. Chao, 2014). These metrics regarding profits have usually been accompanied by other metrics such as Drawdown ratio and Sharpe ratio (Wang, Lee, Xiang, Liu, Lei, & Chau, 2019; Liu, Yu, & Han, 2002). Other metrics used in the evaluation of trading systems can include the profit factor, the expected payoff, and the profit margin (Yong, Rozaini, & Abdullah, 2017; Wu, Tseng, Chan, Huang, Chu, & Chen, 2012). Apart from the profitability, the combination of these metrics can give valuable insights regarding the overall behavior of an automated trading system, such as the risk taken, or the consistency of its returns. Different metrics can be weighted and combined to form a single metric for evaluating an automated trading system (Svoboda, 2012) who created a

5F metric using 5 metrics with equal weights. In the current work, 5 different metrics were combined into a new metric, examining the use of different weights for each one to compare and rank several automated trading systems, continuing on the work done by Vezeris and Schinas (2018).

Since most of the automated trading systems utilize rules and components with various parameters, choosing optimal values for each parameter is of great importance. One of the simplest ways to determine an optimal value for a parameter is to test an automated trading system for different values of each parameter for a set period and then use the best values for a future period. This forward testing technique has been employed by Marcus (2013). Automated trading systems that use neural networks, or other components that have to be trained, usually allow for a training period for training and configuring the system and an evaluation period for evaluating the system (Silva, Castilho, Pereira, & Brandao, 2014; Trippi & Desieno, 1991). Another approach is to choose a variable period of time to use as a back-testing period (Vezeris, Schinas, & Papaschinopoulos, 2016) with the introduction of the d-Backtest PS method. In the current work, both a constant back-testing period and the d-Backtest PS method at weekly intervals are used (Vezeris, Kyrgos, & Schinas, 2018a) to optimize the parameters for all the systems studied, having refined its back-testing phase, rendering it less computationally intensive.

With this work, a weighted metric is introduced and examined that can be used for evaluating the trading systems, as well as a new factor, the Extended Profit Margin, which is used in the weighted metric mentioned above. Another contribution consists of the improvement of the back-testing phase of the d-Backtest PS method so that it becomes much less computationally demanding, without loss of the d-Backtest's edge, which provided us with data on time, for all 10 systems and 7 assets of financial markets over a trading period of 3.5 years. To the best of our knowledge, this is the first time that the d-Backtest PS method has been used

to effectively compare different automated trading systems, paired with a newly introduced metric for their evaluation.

2. METHODS

2.1. Trading systems assessment factors

When it comes to evaluating the performance of the automated trading systems, some qualities such as return, volatility, and risk are important (Ilić & Brtko, 2011; Kumiega & Vliet, 2012). Some metrics that are related to the net profit, the consistency, and the security of the trading strategy will be described in the following sections. Because the available data from the methods used are weekly results, to summarize the data of 183 weeks into a single value, some metrics had to be revised.

The total net *profit* is calculated by subtracting the gross loss of all losing trades from the gross profit of all winning trades.

While certainly a valuable metric, the metric alone can be deceptive as it cannot determine if a trading system is performing efficiently, nor can it normalize the results of a trading system based on the sustained amount of risk. Total net profit should be viewed in concert with other performance metrics. Henceforth, Total Net Profit will be referred to as *Profit*.

The *extended Profit Margin (xPM)* is a variant of profit margin. Profit margin indicates the profitability of a product, service, or business. It is expressed as a percentage; the higher the number, the more profitable the business.

In foreign exchange (FOREX) trading, profit margin is equal to the ratio of the final net profit divided by the gross profit during the examined period.

As discussed in Kim et al. (2010), although profit margin hedging is not the optimal rule for mean reversion, it can still be profitable if prices are mean-reverting. The need for a measure, which will represent both profit and loss, has led us to an extended equation. When the final net *Profit*

is negative, the ratio of the net profit divided by the Gross loss is calculated. When the net profit amounts to zero, then *xPM* is also zero. So the following equation can describe *xPM* :

$$xPM = \begin{cases} \frac{NetProfit}{GrossProfit}, & Profit > 0 \\ 0, & Profit = 0 \\ \frac{NetProfit}{GrossLoss}, & Profit < 0 \end{cases}, \quad (1)$$

extended Profit Margin (xPM) could express the weekly profit margin for both negative and positive values. It also has values within the range of $[-1, 1]$.

SharpeRatio (SR) is a popular tool that investors and fund managers use to calculate the risk-adjusted return in stocks, and it was first introduced by Sharpe (1975). Otherwise, *SR* is referred to as reward-to-variability ratio. Essentially, the ratio shows how much excess return one receives in return for the extra volatility endured as the 'price' for holding a riskier asset:

$$SharpeRatio = \frac{AHPR - (1 - RFR)}{Standard\ Deviation(HPRs)}, \quad (2)$$

where *RFR* is the risk-free rate, which is normally assumed as 0% in FOREX trading, *AHPR* is the average holding period return on investment or, simply put, an arithmetic mean of a relative gain per trade. Calculated as the sum of HPR divided by HPR total, *HPR* is calculated as a ratio between balance after out or in-out operation and previous balance (after balance/previous balance).

In general, the higher the *SR*, the more risk-efficient is the trading system and the smoother its return over time. A negative *SR* either means that the risk-free rate is greater than the portfolio's return, or that the expected return is likely to be negative.

To evaluate the performance of the systems, the *SR* of the weekly data was calculated using the following formula:

$$SR = \frac{\text{average}\left(\frac{\text{profit} + 10000}{10000}\right) - 1}{\text{standard deviation}\left(\frac{\text{profit} + 10000}{10000}\right)}, \quad (3)$$

where at the start of each week – the balance is \$10,000, so *Profit* + \$10,000 is the balance at the end of the week. As the data for each individual trade were not available, it was considered that one week is one trade to summarize the metric for the data.

A *maximum drawdown* (*MDD*) is the maximum observed loss from a peak to a trough of a portfolio before a new peak is attained. *MDD* is an indicator of downside risk over a specified time period. It is an important measure of risk for a trading strategy (Pardo, 2012). The formula for *MDD* is as follows:

$$MDD = \frac{\text{TroughValue} - \text{PeakValue}}{\text{PeakValue}}. \quad (4)$$

It is important to note that it only measures the size of the largest loss without considering the frequency of large losses.

MDD is an indicator used to assess the relative riskiness of one strategy versus another as it focuses on capital preservation, which is a key concern for most investors. A low maximum drawdown is preferable, as this indicates that losses from investment were minimal.

The expected payoff (*EP*) is a statistically calculated index representing the average profit/loss factor of a trade. It can also be considered when it comes to displaying the expected return of the next trade. It is calculated as total net profit divided by total trades.

All the measures above are useful, and each of them shows something different about the automated trading systems. To evaluate all factors simultaneously, it was decided to examine a weighted metric for five factors. A similar effort was made by Svoboda (2012) who integrated five indices into one index with equal weights. Their sub-indices represent yield, liquidity, success factor, stability, financial default. In our version of the weighted

metric, each factor used in the equation can have a different weight. The following basic assumptions were used as the starting point that can help minimize the space of the weights combinations to be examined:

- Profit and Extended Profit Margin are the most important because they represent the profitability of the system and its safety, so their weights should be greater than those of other factors.
- Drawdown is the next most important because it represents the risk of a strategy, but its weight must be negative as lower values are more desirable.
- Given the correlation of the Expected Payoff metric with the Profit metric, its weight has to be relatively small, smaller than the weights for the other metrics, and smaller than the Sharpe Ratio's weight, which is a measure of the consistency of a system's returns.
- Normalization must take place for higher quality results because of the wide range of values between the different metrics of each system, which allows for the sum of the Weights to be equal to 1.

For the normalization of each metric, each value was divided by the sum of the absolute values of the measure:

$$\text{normalized } x_i = \frac{x_i}{\sum_{i=1}^n \text{abs}(x_i)}, \quad (5)$$

where x_i is the value of a measure and n is the number of the trading systems being evaluated.

Using the normalized metrics, the value of the combined Weighted Metric was examined:

$$\begin{aligned} wm = & w_{\text{Profit}} \cdot \text{Profit} + w_{\text{xPM}} \cdot \text{xPM} + \\ & + w_{\text{SR}} \cdot \text{SR} + w_{\text{Drawdown}} \cdot \text{DD} + w_{\text{EP}} \cdot \text{EP}, \end{aligned} \quad (6)$$

where each measure is normalized as described above. To avoid arbitrary weight setting and to examine the robustness of the Weighted Metric as a method of systems' classification, the classification

of each system was examined and summarized using all the possible combinations of weights that fulfill the criteria mentioned in the above bullet points, using a value space of $[0,1]$ with a step of 0.02 for each of the weights.

2.2. Automated trading strategies

For this research, 10 popular automated trading strategies were implemented based on well-known indicators and examined their efficiency using the metrics described in sub-section 2.1. Each automated trading system (ATS) is enhanced with an ATR based stop-loss strategy, as it was introduced by Vezeris et al. (2019).

The Adaptive single Moving Average (*Ad1MA*) ATS employs a simple strategy of a single moving average.

When the close price passes over the MA, it is an indication to buy the asset. So, all opposite positions are being secured, and the system opens a long position. If the close price passes under the MA, it is an indication to sell the asset. So, all opposite positions are being secured, and the system opens a short position.

The Adaptive double Moving Average (*Ad2MA*) ATS uses two moving averages to enter a position: the fast-MA with a short period and the slow-MA with a longer period.

When the fast-MA line crosses above the slow-MA line, it is an indication to buy the asset. So, all opposite positions are being secured, and the system opens a long position. When the fast-MA line crosses below the slow-MA line, it is an indication to sell the asset. So, all opposite positions are being secured, and the system opens a short position.

The Adaptive triple Moving Average (*Ad3MA*) ATS uses three exponential moving averages (EMA) to enter a position. The moving average with the smallest period is called the fast-EMA, the one with the medium period is called the medium-EMA and the other with the biggest period is called the slow-EMA.

When the fast-EMA line is above the medium-EMA line and the medium-EMA line above

the slow-EMA line, this constitutes a buy signal, and any short positions are closed, and a long position is opened. When the fast-EMA line is below the medium-EMA line and the medium-EMA line below the slow-EMA line, this constitutes a sell signal, and any long positions are closed, and a short position is opened. When the hierarchies described above cease to exist, any open positions are closed.

The Adaptive Average Directional Index (*AdADX*) ATS uses the well-known ADX indicator, created by Wilder (1978). ADX shows the strength of a trend. The direction of the trend is shown by two accompanying indicators, the Negative Directional Indicator $-DI$ and the positive Directional Indicator $+DI$.

When the $+DI$ line is above the $-DI$ line and the ADX line crosses above a certain threshold, this constitutes a buy signal, and a long position is opened. When the $-DI$ line is above the $+DI$ line and the ADX line crosses above a certain threshold, this constitutes a sell signal, and a short position is opened.

As the trend is ongoing and a position is open, it would be valuable to calculate the decline of the ADX line as a drawdown percentage of the ADX value to grasp the weakening of the trend early on and exit the position.

So, when the ADX line crosses below an exit threshold, or its value drops below the ADX drawdown value, exit signals are triggered, and any open positions are closed.

The Adaptive Ichimoku (*AdIchimoku*) ATS uses the well-known Ichimoku indicator (Elliott, 2007). The indicator comprises five lines called the *tenkan-sen*, *kijun-sen*, *chikou span* and lastly *senkou span A* and *senkou span B* which – when combined – form the “kumo cloud”.

When the *tenkan-sen* line is above the *kijun-sen* line, the price is above the kumo cloud, the kumo cloud's width is not zero and the *chikou span* is above the equivalent price of the past, these conditions constitute a buy signal, and a long position is opened.

When the *tenkan-sen* line is below the *kijun-sen* line, the price is below the kumo cloud, the kumo

cloud's width is not zero and the chikou span is below the equivalent price of the past, these conditions constitute a sell signal, and a short position is opened.

When the price breaks above the kumo cloud, this constitutes an exit signal and any short positions are closed. Similarly, when the price breaks below the kumo cloud, this constitutes another exit signal, and any long positions are closed.

The Adaptive Moving Average Convergence Divergence (*AdMACD*) ATS uses the well-known MACD indicator, created by Appel (2005). The indicator comprises two lines: the MACD and a moving average of the MACD called the Signal or Trigger line.

If the MACD passes over the Trigger line, it is an indication to buy the asset. So, all opposite positions are being closed, and the system opens a long position. If the MACD passes under the Trigger line, it is an indication to sell the asset. So, all opposite positions are being closed, and the system opens a short position.

The Adaptive Parabolic Stop and Reverse (*AdParSAR*) ATS uses the Parabolic SAR indicator, created by Wilder (1978).

When the close price crosses above the SAR value, this constitutes a buy signal, and any short positions are closed, and a long position is opened. When the close price crosses below the SAR value, this constitutes a sell signal, and any long positions are closed, and a short position is opened.

To confirm that a trend has really reached its end and a reversal is underway, the system can delay the exit of a position before entering a new one.

The Adaptive Pivot (*AdPivot*) ATS uses as an indicator, Support and Resistance levels created around Pivot Points.

When the close price crosses above a chosen Resistance level, this constitutes a buy signal, and a long position is opened. When the closing price crosses below a chosen Support level, this constitutes a sell signal, and a short position is opened.

The Take Profit can be set either on the next Support level for short positions or the level after that. Similarly, for the long positions, it can be either set on the next Resistance level or the level after that. There is also an option, where instead of closing the whole position on the Take Profit level, only half is closed. Like this, one manages to secure a portion of the profits, and the rest of the volume that continues to ride on the potential trend, closing only by the trailing Stop Loss.

The Adaptive Turtle (*AdTurtle*) ATS uses various Donchian Channels and is formulated based on the strategy described in Curtis (2007). This version is a revision of the one implemented in the research of Vezeris et al. (2019). There are four (4) different periods for each Donchian Channel:

- a period x for opening positions;
- a period $x_c = x/n$ for closing positions;
- a period y for opening delayed positions;
- a period $y_c = y/m$ for closing delayed positions.

Typically, $x < y$ and $n = m = 2$.

When the price crosses above the x_{period} upper band, this constitutes a buy signal, and a long position is opened, and when the price crosses below the x_{c_period} lower band, this constitutes an exit long signal, and any long positions are closed. When the price crosses below the x_{period} lower band, this constitutes a sell signal, and a short position is opened, and when the price crosses above the x_{c_period} upper band, this constitutes an exit short signal, and any short positions are closed.

In times of volatile periods, the y_{period} upper and lower bands are used to re-open a position when the previous position was not closed by a stop loss signal and it was profitable. To close these positions, the y_{c_period} upper and lower bands are used.

In addition to these rules, the *AdTurtle* ATS keeps on adding volume to the initial open position as the price moves to more profitable levels. The maximum permitted additions to the initial po-

sition are five, and each time it invests less than it previously did.

The Adaptive Bollinger Bands Anti-trend (*AdBBAntiTrend*) ATS, uses the well-known Bollinger Bands indicator, created by Bollinger (2001). The indicator consists of 3 bands, the middle band is a simple moving average, while the upper and lower bands are typically two standard deviations.

When the close price first breakouts the upper band and then closes below it, this constitutes a sell signal, and any long positions are closed, and a short position is opened. When the close price first breakouts the lower band and then closes above it, this constitutes a buy signal, and any short positions are closed, and a buy position is opened.

An example of each strategy is shown in Appendix A. Each figure shows the EUR/USD time series – from Metatrader 5 trading terminal – along with the respective indicator of each ATS. The horizontal axis represents time, while the vertical axis the price of EUR/USD. Finally, the red downward arrows indicate sell signals, while the blue upward ones, buy signals. A green dot indicates an exit signal.

2.3. Back-testing

The Metatrader 5 trading platform by MetaQuotes was utilized to carry out the backtests to examine the automated trading systems. Besides, the Microsoft SQL Server was used to collect and handle the findings.

Backtests were carried out on 7 assets of financial markets, specifically on EUR/USD, GBP/USD, USD/CHF, USD/JPY, XAU/USD, WTI, and BTC/USD, with weekly tests over a three and a half-year period, from 24/01/2016 to 27/07/2019 with data from the ForexTime, FxPro, and Alpari brokers. The authors refrained from choosing assets from categories such as equities or rates because their trading sessions are brief in comparison.

The initial capital for each test is set at \$10,000. Each ATS risks losing 20% of the equity with every trade over a 1% movement of price on every asset apart from WTI and BTC/USD. These two assets

are highly more volatile than the rest. Therefore, the percentage was adjusted to 3% and 20% price movement for WTI and BTCUSD, respectively.

Backtesting processes are widely used today in forecasting experiment tests. The d-Backtest PS method dynamically finds the best back-testing period that will be considered for the following week.

The d-Backtest PS method requires several backtests, typically 30; it has to choose parameters for each week. This means that many backtests have to be run in to have enough input for the d-Backtest PS method over a long period of trading time, which makes this first phase the most computationally expensive one. In this research, the authors have devised a new way to shorten the backtests' phase by running only the backtests for the one-week periods instead of all the back periods until 30 weeks. Using the data from the one-week tests, a backtest of any number of consecutive weeks can then be extracted by combining the resulted metrics from each combination of parameters throughout the consecutive one-week backtests, using the techniques described in sub-section 2.1. The tradeoffs of this technique are those not considered the compounding effects in the back-testing periods but can examine larger combinations of parameters for each system over a much larger period. The comparison of this new version of the d-Backtest PS method with the best BTs and the 6-month back-testing methods described next offers an assessment of the new methodology's performance.

An artifact of the d-Backtest PS method is the fact that one had to run tests and collect enough data for a few weeks before the examined period, in order for the d-Backtest PS method to initialize the different methods it uses for classifying the back-testing periods.

The best back-testing periods, which will be mentioned as best BTs, are the ideal version of the d-Backtest PS method. The best BTs are determined using ex post-computed historical data to select the optimal back-testing period and presented in this work to show the absolute limit that the d-Backtest PS method could theoretically achieve.

To compare the d-Backtest PS method with more traditional and used methods, a simple back-testing method was also examined, which uses a constant back-testing period of 6 months (26 weeks). By that simple method, each system selects the best parameters of the previous period of 26 continuous weeks and uses them in the next adjacent week.

3. RESULTS

The results for each system can be presented now that the automated trading systems and the measures used to evaluate them are described. Apart from the data resulting from the d-Backtest PS, the results of the best back-testing periods that the d-Backtest PS method could have chosen are also presented. Besides, the data of a simple 6-month back-testing method are presented.

If excluding the results from best BTs method, which was expected to outperform the other two methods, the results from the d-Backtest PS method far exceed the results from the simple 6-month back-testing method, which suffers losses in all systems. The detailed results can be found in Appendix.

3.1. Weighted metric classification

The WM metric was calculated for all the combinations of weights with values from 0 to 1 with a

step of 0.02 fulfilling the requirements in paragraph 2.6, and the ranking of each system in all of these combinations (a total of 224,956 combinations) was examined.

The following tables summarize the classification occurrences of each system for every different weight combination examined. The d-Backtest PS method's results can be seen in Table 1, while the results for the best BTs and the 6-month constant BT can be seen in Tables 2 and 3, respectively. For the d-Backtest PS method, the Ad2MA stands out clearly as the best system in any of the WM used, with the AdPivot and AdADX coming next. For the Best BTs, the first 3 systems remain the same, but this time AdPivot comes first with Ad2MA and AdADX following. For the 6-month constant BT, there is a different ranking with the AdBBAntiTrend coming first, followed by the AdTurtle, with the AdPivot in the third place.

3.2. Compounding capital diagrams

One key technique for stock trading is also the portfolio management (Petroopoulos et al., 2017; Yao et al., 2007). In the testing round, each system started each week with \$10.000 for each asset. But there could be value in examining what a system would have done when investing its profits from the previous week to the next, compounding its capital. To examine this, the authors started with

Table 1. Classification occurrences for all systems based on the WM values for d-Backtest PS method

ATS	1st place	2nd place	3rd place	4th place	5th place	6th place	7th place	8th place	9th place	10th place
Ad2MA	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
AdPivot	0.0%	61.6%	40.4%	0.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
AdADX	0.0%	38.1%	47.8%	10.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
AdTurtle	0.0%	0.2%	11.7%	89.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
AdBBAntiTrend	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%
AdParSAR	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
AdIchimoku	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%
AdMACD	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	99.7%	0.6%	0.0%
Ad1MA	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.3%	99.4%	0.0%
Ad3MA	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%

Table 2. Classification occurrences for all systems based on the WM values for best BTs

ATS	1st place	2nd place	3rd place	4th place	5th place	6th place	7th place	8th place	9th place	10th place
AdPivot	72.6%	32.0%	4.3%	1.2%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%
Ad2MA	18.5%	56.0%	19.0%	1.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
AdADX	0.2%	8.5%	65.5%	10.3%	7.5%	3.5%	2.7%	1.5%	0.0%	0.0%
AdBBAntiTrend	0.0%	0.0%	1.1%	80.1%	14.5%	1.9%	0.4%	0.0%	0.0%	0.0%
AdParSAR	8.7%	1.0%	9.7%	2.6%	71.3%	4.5%	0.0%	0.0%	0.0%	0.0%
AdIchimoku	0.0%	0.0%	0.1%	0.0%	5.1%	78.0%	13.1%	0.0%	0.0%	0.0%
AdMACD	0.0%	2.5%	0.5%	4.4%	1.5%	11.6%	45.5%	30.4%	8.5%	0.0%
Ad1MA	0.0%	0.0%	0.0%	0.0%	0.0%	0.5%	38.3%	57.9%	0.0%	0.0%
AdTurtle	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	10.1%	91.5%	0.0%
Ad3MA	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%

Table 3. Classification occurrences for all systems based on the WM values for 6-month constant BT

ATS	1st place	2nd place	3rd place	4th place	5th place	6th place	7th place	8th place	9th place	10th place
AdBBAntiTrend	73.7%	25.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
AdTurtle	26.3%	73.4%	2.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
AdPivot	0.0%	1.6%	97.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
AdADX	0.0%	0.0%	0.0%	80.6%	4.3%	14.9%	0.0%	0.2%	0.0%	0.0%
AdIchimoku	0.0%	0.0%	0.0%	19.4%	79.6%	0.3%	0.0%	0.0%	0.0%	0.0%
Ad3MA	0.0%	0.0%	0.0%	0.0%	15.7%	81.0%	4.4%	0.0%	0.0%	0.0%
Ad2MA	0.0%	0.0%	0.0%	0.0%	0.3%	3.6%	92.2%	0.7%	2.3%	0.5%
Ad1MA	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	54.8%	21.7%	24.2%
AdParSAR	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	2.6%	41.8%	52.5%	0.0%
AdMACD	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.8%	2.5%	23.5%	75.3%

\$90,000 initial capital. Each week, the starting balance for every asset amounts to one ninth of the total capital and the remaining two ninths of the capital are reserved in case of a shortage of capital. The net profit was estimated for each week for every asset if trading with the final balance of the previous week instead of the \$10,000 starting balance was continued:

$$\text{EstimatedProfit} = \frac{\text{StartingBalance} \cdot \text{ActualProfit}}{10000} \quad (7)$$

Thus, a full picture of the 3.5-year testing period using parameters generated by the d-Backtest PS method can be observed in Figure 1.

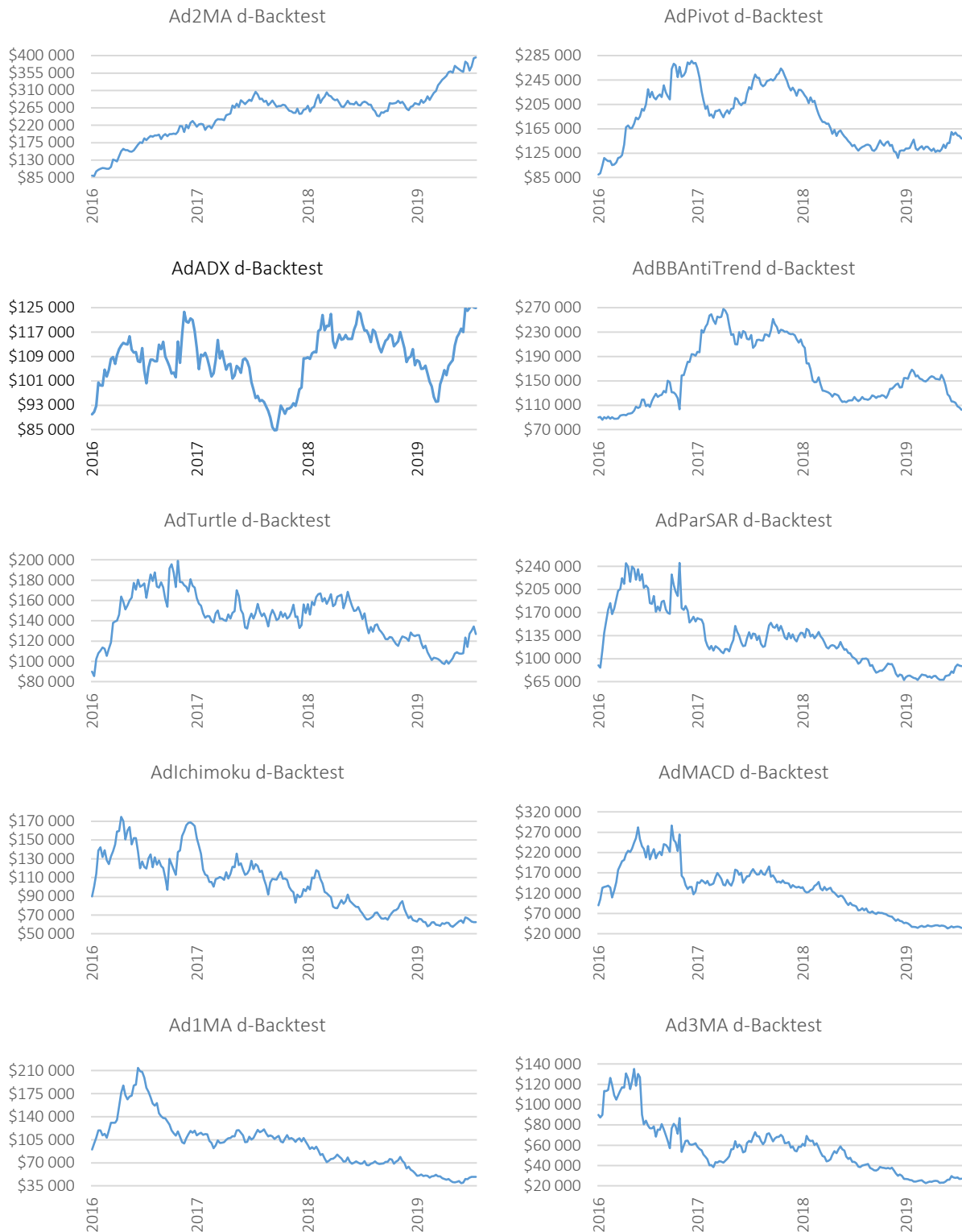


Figure 1. Systems' 3.5-year balance graph (January 2016 – July 2019) with compounding capital using the d-Backtest PS method for parameter selection

CONCLUSION

With the current research, the study regarding the comparison of high-frequency algorithmic trading systems in financial markets is broadened by implementing ten such automated trading systems, based on a widely used technical indicator. To evaluate these systems, the d-Backtest PS method was used at weekly intervals, the best BTs periods, which provide the ideal outcome of the d-Backtest PS method and a 6-month constant back-testing strategy. To assess the results, consistency and low risk were also taken into consideration. Typical metrics provided by Metatrader 5 were used. Instead of profit factor, the more robust xPM metric was devised, which is an extended version of Profit Margin, and its values range is $[-1, 1]$. Because of the importance of every metric, a weighted metric aggregation was created to evaluate the systems, and the evaluation results of different weights for every metric were examined.

Eventually, from all the automated trading systems, Ad2MA, along with AdPivot, excel both in the d-Backtest PS method and the best BT periods. Especially in the d-Backtest PS method, the difference between Ad2MA and the rest of the systems was obvious. The simple 6-month constant back-testing strategy does not yield profitable results for any of the systems. Moreover, an attempt of primary portfolio management was made, using seven selected symbols, and Ad2MA was the only one who had an overall upward tendency.

In the future, we plan to study more closely the two automated trading systems that stood out in the current research – Ad2MA and AdPivot – and examine whether variations of them could perform profitably with less dependence on their parameters' values.

AUTHOR CONTRIBUTIONS

Conceptualization: Dimitrios Vezeris.

Investigation: Dimitrios Vezeris.

Methodology: Dimitrios Vezeris, Ioannis Karkanis, Vasiliki Bizergianidou.

Software: Ioannis Karkanis, Themistoklis Kyrgos.

Supervision: Dimitrios Vezeris.

Visualization: Themistoklis Kyrgos, Ioannis Karkanis, Vasiliki Bizergianidou.

ACKNOWLEDGMENTS

We would like to thank Dr. Christos Schinas for his time and invaluable guidance towards the methodology of the weighted metric. We would also like to thank Michalis Foulos for the hardware setup and support and Nektarios Mitakidis for his contribution to the representation of the results.

This research has been co-financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (project code: T1EDK-02342).



Co-financed by Greece and the European Union

REFERENCES

1. Appel, G. (2005). *Technical Analysis: Power Tools for Active Investors*. Upper Saddle River: FT Press.
2. Azhikodan, A. R., Bhat, A. G. K., & Jadhav, M. V. (2019) Stock Trading Bot Using Deep Reinforcement Learning. In H. Saini, R. Sayal, A. Govardhan & R. Buyya (Eds.), *Innovations in Computer Science and Engineering* (pp. 41-29). Springer, Singapore. https://doi.org/10.1007/978-981-10-8201-6_5
3. Bollinger, J. (2001). *Bollinger on Bollinger Bands*. New York City: McGraw-Hill.
4. Booth, A., Gerding, E., & McGroarty, F. (2014). Automated trading with performance weighted random forests and seasonality. *Expert Systems with Applications*, 41(8), 3651-3661. <https://doi.org/10.1016/j.eswa.2013.12.009>
5. Chong, T. T.-L., & Ng, W.-K. (2008). Technical analysis and the London stock exchange: testing the MACD and RSI rules using the FT30. *Applied Economics Letters*, 15(14), 1111-1114. <https://doi.org/10.1080/13504850600993598>
6. Chou, Y.-H., Kuo, S.-Y., Chen, C.-Y., & Chao, H.-C. (2014). A Rule-Based Dynamic Decision-Making Stock Trading System Based on Quantum-Inspired Tabu Search Algorithm. *IEEE Access*, 2, 883-896. <https://doi.org/10.1109/ACCESS.2014.2352261>
7. Curtis, F. (2007). *Way of the Turtle: The Secret Methods that Turned Ordinary People into Legendary Traders*. New York City: McGraw-Hill.
8. Elliott, N. (2007). *Ichimoku Charts: An introduction to Ichimoku Kinko Clouds*. Hampshire, UK: Harriman House.
9. Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2018). Stock price prediction using support vector regression on daily and up to the minute prices. *The Journal of Finance and Data Science*, 4(3), 183-201. <https://doi.org/10.1016/j.jfds.2018.04.003>
10. Ilić, V., & Brtko, V. (2011). Evaluation of algorithmic strategies for trading on foreign exchange market. *Proceedings from Information and Communication Technologies for Small and Medium Enterprises*. Arandjelovac, Serbia. <https://doi.org/10.13140/RG.2.1.2605.9924%20>
11. Kim, H. S., Brorsen, B. W., & Anderson, K. B. (2010). Profit Margin Hedging. *American Journal of Agricultural Economics*, 92(3), 638-653. <https://doi.org/10.1093/ajae/aap038>
12. Kumiega, A., & Vliet, B. E. (2012). Automated Finance: The Assumptions and Behavioral Aspects of Algorithmic Trading. *The Journal of Behavioral Finance*, 51-55. <https://doi.org/10.1080/15427560.2012.654924>
13. Liu, Y., Yu, X., & Han, J. (2002). Sharpe Ratio-Oriented Active Trading: A Learning Approach. In *Proceedings from Mexican International Conference on Artificial Intelligence* (pp. 331-339). Mexico. https://doi.org/10.1007/3-540-46016-0_35
14. Marcus, D. (2013). Portfolio Theory Forward Testing. *Advances in Management & Applied Economics*, 3(3), 225-244. Retrieved from https://scienpress.com/journal_focus.asp?main_id=55&Sub_id=IV&Issue=644
15. Osunbor, V., & Egwali, A. (2016). Development of OSEG: A FOREX Expert Advisor. *The Pacific Journal of Science and Technology*, 17(2), 206-212. Retrieved from http://akamaiuniversity.us/PJST17_2_206.pdf%20
16. Pardo, R. (2012). *The Evaluation and Optimization of Trading Strategies, Second Edition*. Hoboken, New Jersey: John Wiley & Sons, Inc.
17. Petropoulos, A., Chatzis, S., Sialkoulis, V., & Vlachogiannakis, N. (2017). A stacked generalization system for automated FOREX portfolio trading. *Expert Systems with Applications*, 90, 290-302. <https://doi.org/10.1016/j.eswa.2017.08.011>
18. Sharpe, W. (1975). Adjusting for Risk in Portfolio Performance Measurement. *Journal of Portfolio Management*, 1(2), 29-34. <https://doi.org/10.3905/jpm.1975.408513>
19. Silva, E., Castilho, D., Pereira, A. C., & Brandao, H. (2014). A neural network based approach to support the Market Making strategies in High-Frequency Trading. In *Proceedings from International Joint Conference on Neural Networks* (pp. 845-852). Beijing, China. <https://doi.org/10.1109/IJCNN.2014.6889835>
20. Svoboda, J. (2012). 5 F performance indicator: A robust metric for trading systems evaluation? In *Proceedings from: 1st WSEAS International Conference on Finance, Accounting and Auditing (FAA '12)* (p. 92). Zlin, Czech Republic: WSEAS Press. Retrieved from <http://wseas.us/e-library/conferences/2012/Zlin/FAA/FAA-14.pdf>
21. Trippi, R. R., & Desieno, D. (1991). Trading Equity Index Futures With A Neural Network. *The Journal of Portfolio Management*, 19(1), 27-33. <https://doi.org/10.3905/jpm.1992.409432>
22. Vezeris, D. T., & Schinas, C. J. (2018). Performance Comparison of Three Automated Trading Systems (MACD, PIVOT and SMA) by Means of the d-Backtest PS Implementation. *International Journal of Trade, Economics and Finance*, 9, 170-173. <https://doi.org/10.18178/ijtef.2018.9.4.609>
23. Vezeris, D. T., Kyrgos, T. S., & Schinas, A. C. (2018a). Hedging and non-hedging trading strategies on commodities using the d-Backtest PS method. *Investment Management and Financial Innovations*, 15, 351-369. [https://doi.org/10.21511/imfi.15\(3\).2018.29](https://doi.org/10.21511/imfi.15(3).2018.29)
24. Vezeris, D. T., Schinas, C. J., & Papaschinopoulos, G. (2016). Profitability Edge by Dynamic Back Testing Optimal Period Selection for Technical Parameters Optimization, in *Trading Systems with Forecasting. Computational*

- Economics*, 51, 761-807. <https://doi.org/10.1007/s10614-016-9640-x>
25. Vezeris, D., Karkanis, I., & Kyrgos, T. (2019). AdTurtle: An Advanced Turtle Trading System. *Journal of Risk and Financial Management*, 12, 96. <https://doi.org/10.3390/jrfm12020096>
26. Vezeris, D., Kyrgos, T., & Schinas, C. (2018b). Take Profit and Stop Loss Trading Strategies Comparison in Combination with an MACD Trading System. *Journal of Risk and Financial Management*, 11, 56. <https://doi.org/10.3390/jrfm11030056>
27. Wang, Y., Lee, H., Xiang, Y., Liu, Y., Lei, Z. B., & Chau, K. Y. (2019). Trading Strategies Evaluation Platform with Extensive Simulations. *Proceedings from IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr)*. Shenzhen, China. <https://doi.org/10.1109/CIFEr.2019.8759059>
28. Wilder, W. (1978). *New Concepts in Technical Trading Systems*. Greensboro, N.C: Trend Research.
29. Wu, H.-C., Tseng, C.-M., Chan, P.-C., Huang, S.-F., Chu, W.-W., & Chen, Y.-F. (2012). Evaluation of stock trading performance of students using a web-based virtual stock trading system. *Computers & Mathematics with Applications*, 64(5), 1495-1505. <https://doi.org/10.1016/j.camwa.2012.03.097>
30. Yao, S., Pasquier, M., & Quek, C. (2007). A foreign exchange portfolio management mechanism based on fuzzy neural networks. In *Proceedings from IEEE Congress on Evolutionary Computation* (pp. 2576-2583). Singapore. <https://doi.org/10.1109/CEC.2007.4424795>
31. Yong, B. X., Rozaini, M., & Abdullah, A. S. (2017). A Stock Market Trading System Using Deep Neural Network. *Proceedings from Asian Simulation Conference 2017*. Melaka, Malaysia. https://doi.org/10.1007/978-981-10-6463-0_31

APPENDIX A

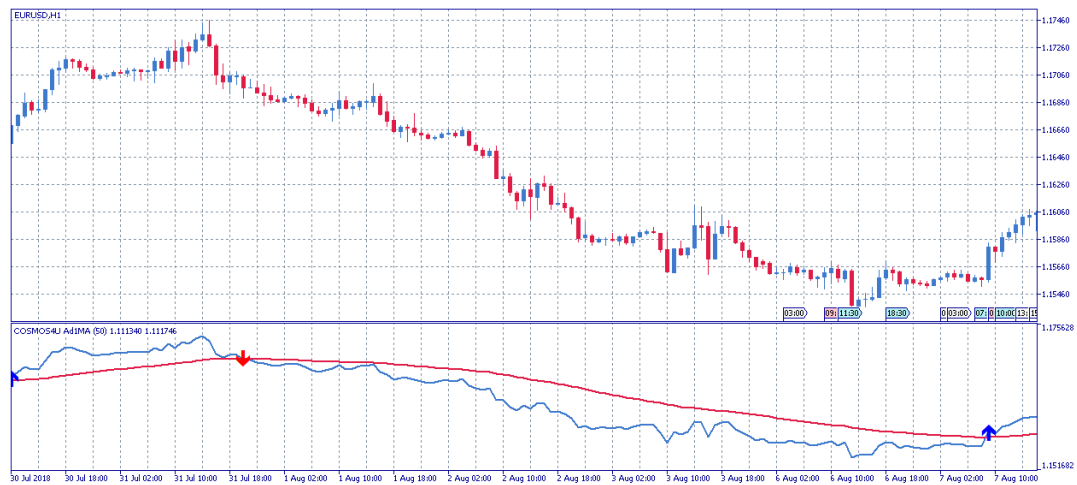


Figure A1. EUR/USD chart with signals from the Ad1MA strategy

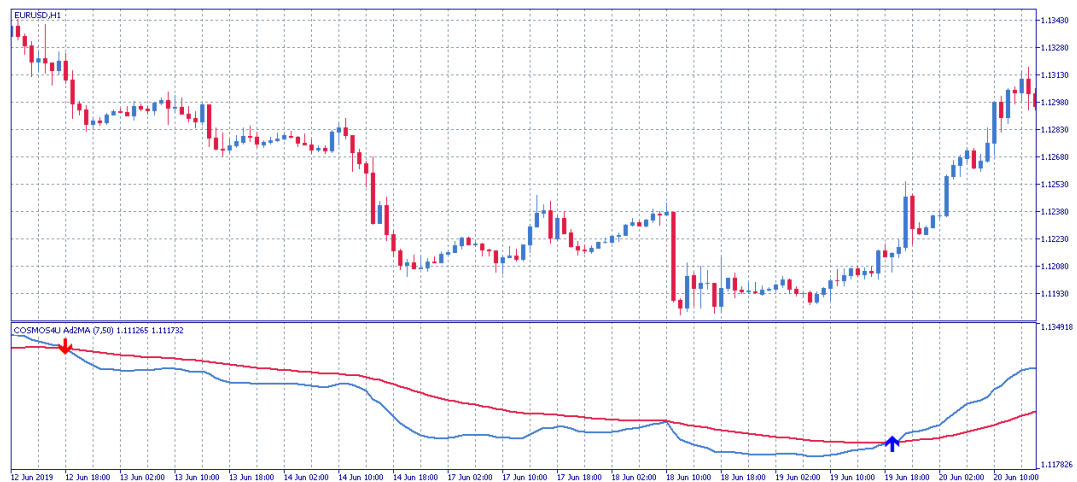


Figure A2. EUR/USD chart with signals from the Ad2MA strategy

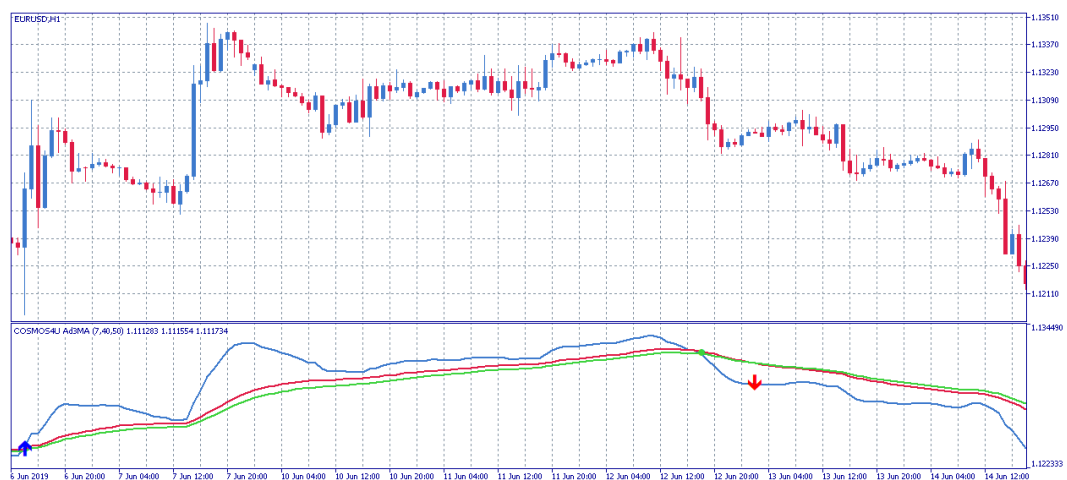


Figure A3. EUR/USD chart with signals from the Ad3MA strategy

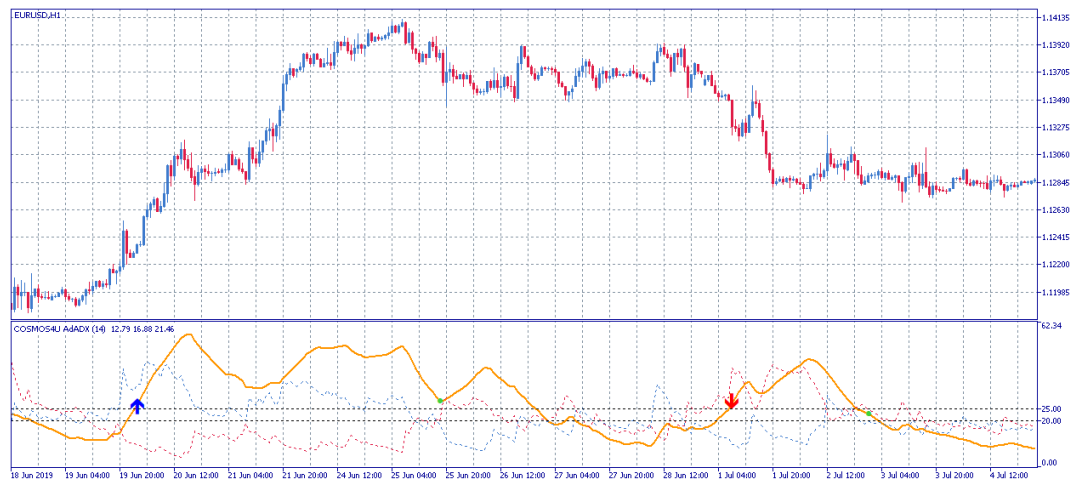


Figure A4. EUR/USD chart with signals from the AdADX strategy

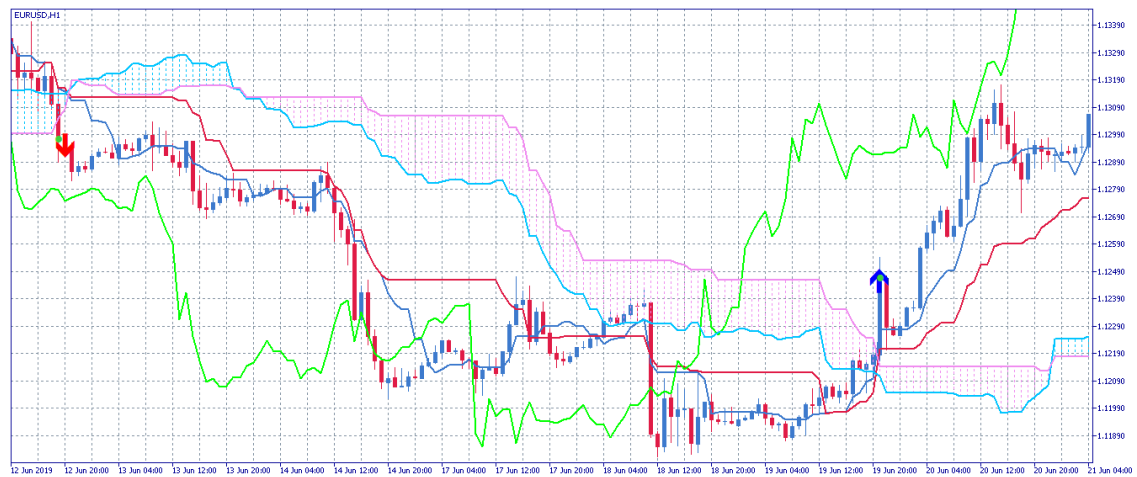


Figure A5. EUR/USD chart with signals from the AdIchimoku strategy

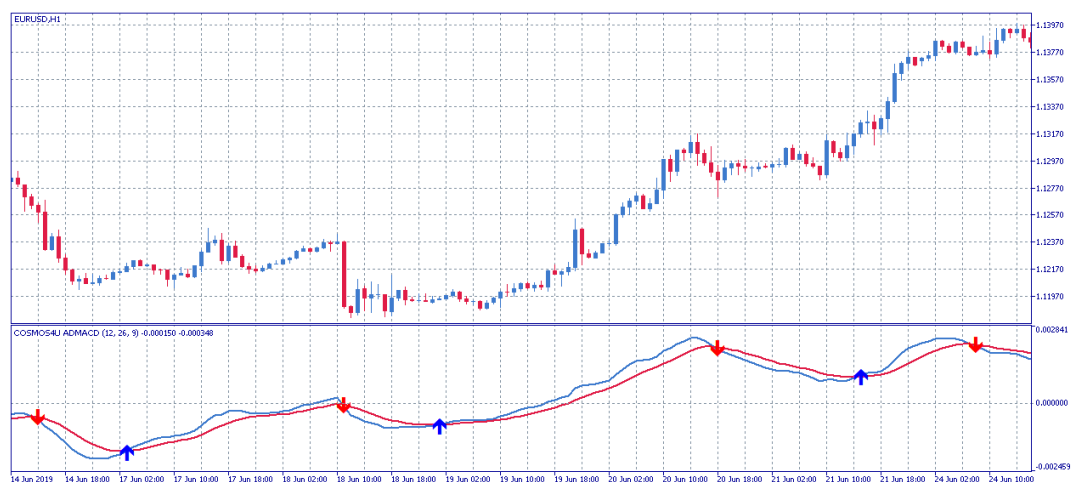


Figure A6. EUR/USD chart with signals from the AdMACD strategy

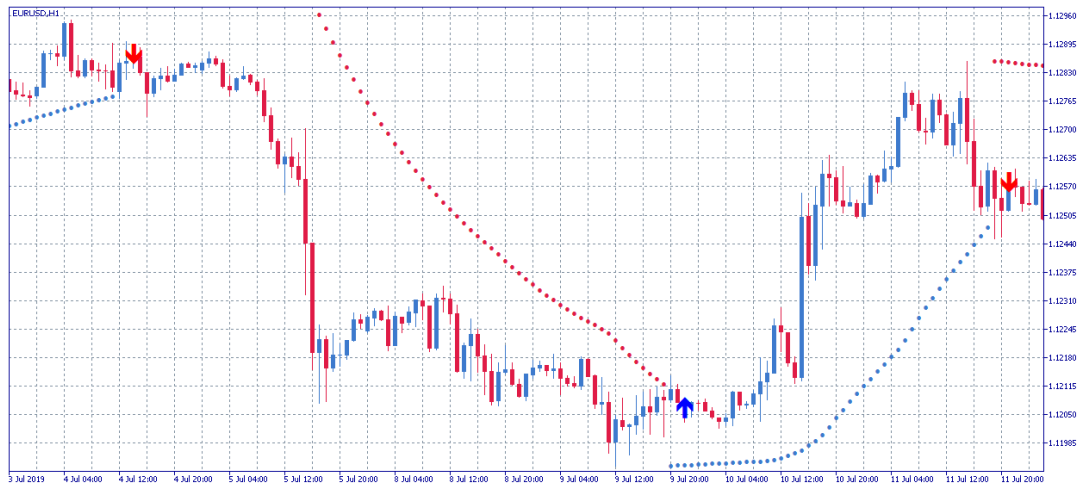


Figure A7. EUR/USD chart with signals from the AdParSAR strategy

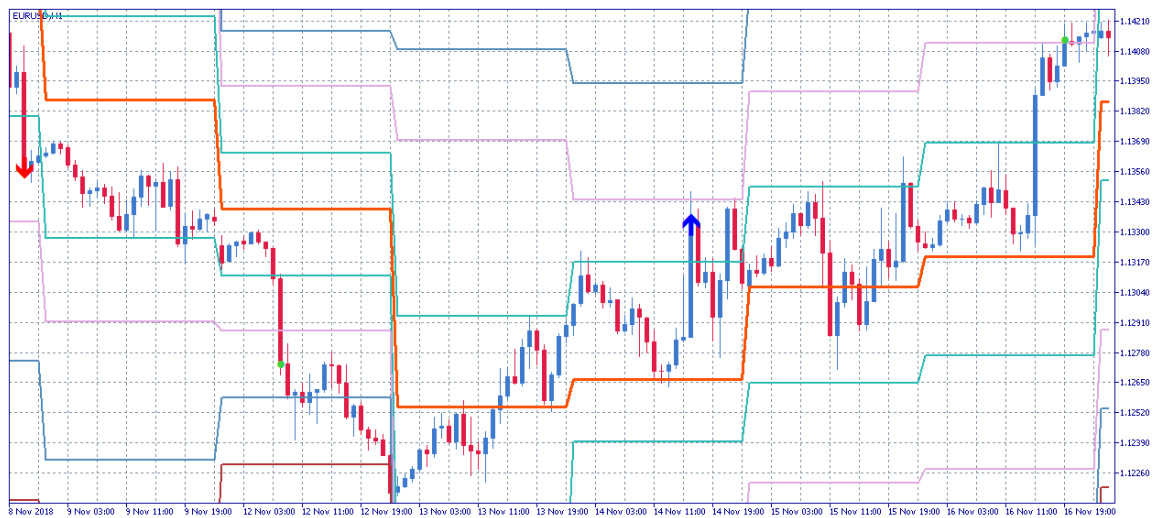


Figure A8. EUR/USD chart with signals from the AdPivot strategy

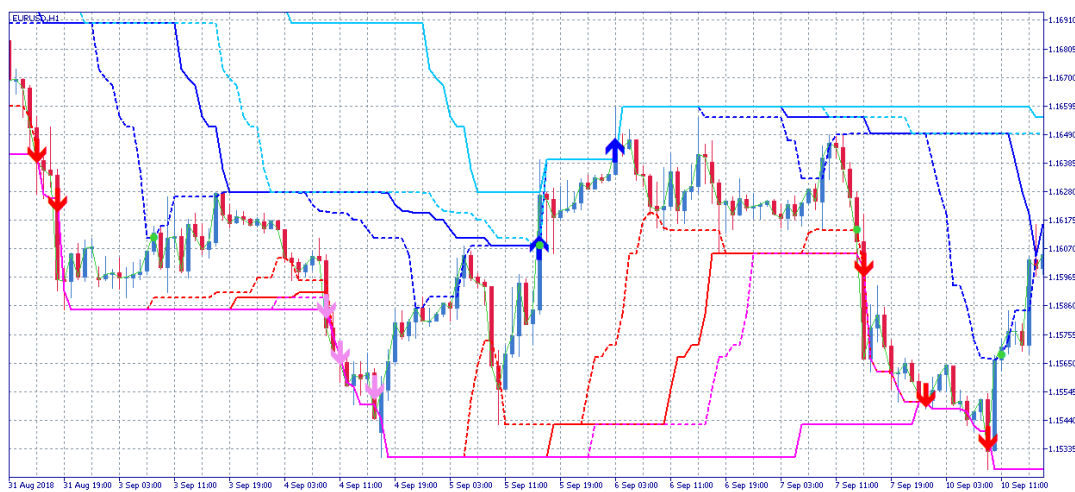


Figure A9. EUR/USD chart with signals from the AdTurtle strategy

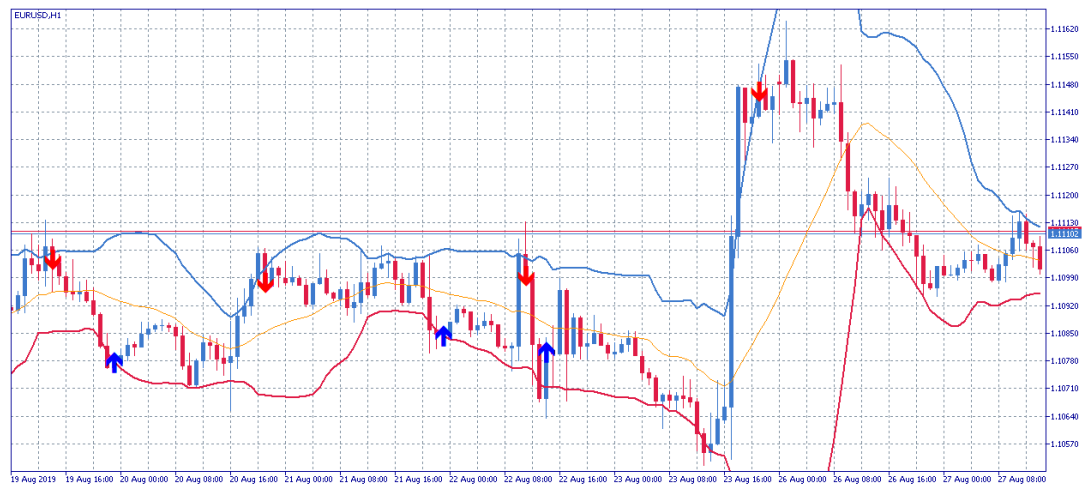


Figure A10. EUR/USD chart with signals from the AdBBAntiTrend strategy

APPENDIX B

Table B1. Total results for Extended Profit Margin (xPM)

ATS	d-Backtest xPM	ATS	Best BTs xPM	ATS	6-m BT xPM
Ad2MA	0.2646	AdPivot	0.9726	AdTurtle	-0.0024
AdADX	0.0926	AdADX	0.9636	AdBBAntiTrend	-0.0238
AdPivot	0.0872	Ad2MA	0.9422	AdPivot	-0.0431
AdBBAntiTrend	0.0520	AdBBAntiTrend	0.9051	AdIchimoku	-0.0497
AdTurtle	0.0493	AdIchimoku	0.7536	Ad3MA	-0.0547
AdParSAR	0.0247	AdParSAR	0.7208	AdADX	-0.0960
AdIchimoku	0.0015	Ad1MA	0.6872	AdParSAR	-0.0990
AdMACD	-0.0161	AdMACD	0.5987	AdMACD	-0.1149
Ad1MA	-0.0271	AdTurtle	0.5581	Ad2MA	-0.1284
Ad3MA	-0.0417	Ad3MA	0.4922	Ad1MA	-0.1598

Table B2. Total results for Sharpe ratio

ATS	d-Backtest Sharpe ratio	ATS	Best BTs Sharpe ratio	ATS	6-m BT Sharpe ratio
Ad2MA	0.0852	AdADX	1.1163	AdTurtle	-0.0024
AdADX	0.0298	AdPivot	1.0597	AdBBAntiTrend	-0.0238
AdTurtle	0.0254	Ad2MA	1.0403	AdPivot	-0.0431
AdPivot	0.0056	AdBBAntiTrend	0.9207	AdIchimoku	-0.0497
AdBBAntiTrend	-0.0179	AdParSAR	0.7232	Ad3MA	-0.0547
AdParSAR	-0.0310	Ad1MA	0.6881	AdADX	-0.0960
AdIchimoku	-0.0331	AdMACD	0.6314	AdParSAR	-0.0990
AdMACD	-0.0461	AdIchimoku	0.5845	AdMACD	-0.1149
Ad3MA	-0.0565	AdTurtle	0.4749	Ad2MA	-0.1284
Ad1MA	-0.0704	Ad3MA	0.3089	Ad1MA	-0.1598

Table B3. Total results for Drawdown

ATS	d-Backtest Drawdown	ATS	Best BTs Drawdown	ATS	6-m BT Drawdown
AdTurtle	36.29	AdADX	26.36	AdTurtle	33.78
AdBBAntiTrend	42.93	AdTurtle	31.31	AdADX	35.04
AdADX	44.63	AdPivot	32.22	AdBBAntiTrend	36.35
Ad2MA	47.07	AdBBAntiTrend	33.08	AdPivot	49.31
AdIchimoku	51.46	Ad2MA	34.77	AdIchimoku	49.39
AdPivot	52.54	AdIchimoku	44.14	Ad1MA	55.62
Ad1MA	59.51	Ad1MA	47.64	Ad2MA	56.21
AdParSAR	63.85	AdParSAR	53.20	AdParSAR	62.28
Ad3MA	70.15	AdMACD	58.86	Ad3MA	68.11
AdMACD	72.05	Ad3MA	62.94	AdMACD	70.66

Table B4. Total results for Expected Payoff

ATS	d-Backtest Expected Payoff	ATS	Best BTs Expected Payoff	ATS	6-m BT Expected Payoff
Ad2MA	143.42	Ad2MA	1099.16	AdTurtle	-0.86
AdPivot	42.56	AdBBAntiTrend	991.14	AdBBAntiTrend	-39.15
AdADX	37.25	AdPivot	897.17	Ad1MA	-39.44
AdBBAntiTrend	28.68	AdADX	873.00	Ad3MA	-49.75
AdTurtle	9.19	AdIchimoku	604.78	AdIchimoku	-50.60
AdParSAR	8.13	AdParSAR	424.19	AdParSAR	-55.13
AdIchimoku	0.63	AdMACD	294.13	AdMACD	-56.11
AdMACD	-4.93	Ad3MA	287.14	AdPivot	-57.30
Ad1MA	-5.15	Ad1MA	277.31	AdADX	-123.02
Ad3MA	-17.09	AdTurtle	139.52	Ad2MA	-152.72

Table B5. Total results for Profit

ATS	d-Backtest Profit	ATS	Best BTs Profit	ATS	6-m BT Profit
Ad2MA	144,137.97	AdParSAR	1,798,997.18	AdTurtle	-3816.48
AdPivot	63,931.84	AdMACD	1,701,561.30	AdBBAntiTrend	-7320.40
AdTurtle	54,510.14	AdIchimoku	1,239,807.25	AdPivot	-18506.61
AdParSAR	41,681.97	Ad1MA	1,167,757.11	AdADX	-19805.80
AdBBAntiTrend	37,883.67	Ad2MA	1,114,543.18	Ad2MA	-86743.47
AdADX	36,695.09	AdPivot	1,099,933.99	AdIchimoku	-103218.44
AdIchimoku	1,556.98	AdBBAntiTrend	958,428.21	Ad3MA	-116127.92
AdMACD	-32,947.64	AdTurtle	868,538.09	Ad1MA	-266743.73
Ad1MA	-33,606.62	Ad3MA	800,542.72	AdParSAR	-274752.42
Ad3MA	-55,532.44	AdADX	733,323.78	AdMACD	-299852.51