Improving Trend Reversal Estimation in Forex Markets Under a Directional Changes Paradigm with Classification Algorithms

Adesola Adegboye[1] | Michael Kampouridis[2] | Fernando Otero[1]

[1]School of Computing, University of Kent, UK

[2]School of Computer Science and Electronic Engineering, University of Essex, UK

Correspondence
M. Kampouridis, School of Computer Science and Electronic Engineering, University of Essex, Wivenhoe Park, CO4 3SQ, UK
Email: mkampo@essex.ac.uk

The majority of forecasting methods use a physical time scale for studying price fluctuations of financial markets. Using physical time scales can make companies oblivious to significant activities in the market as the flow of time is discontinuous, which could translate to missed profitable opportunities or risk exposure. Directional Changes (DC) has gained attention in the recent years by translating physical time series to event-based series. Under this framework, trend reversals can be predicted by using the length of events. Having this knowledge allows traders to take an action before such reversals happen and thus increase their profitability. In this paper, we investigate how classification algorithms can be incorporated in the process of predicting trend reversals to create DC-based trading strategies. The effect of the proposed trend reversal estimation is measured on 20 foreign exchange markets over a 10-month period in a total of 1,000 datasets. We compare our results across 16 algorithms, both DC and non-DC based, such as technical analysis and buy-and-hold. Our findings show that the introduction of classification leads to return higher profit and statistically outperform all other trading strategies.

KEYWORDS
Directional changes, Classification model, Genetic programming, Forex data

1 | INTRODUCTION

Forecasting future behaviour of financial instruments is a major activity in financial markets [12]. A typical way of studying price series is by sampling data on a physical time scale, e.g. daily closing prices, where prices are sampled on daily intervals. However, this approach has a disadvantage: it disregards potentially important price movements that happen in between those fixed points in time. For instance, if we were using daily closing prices, we wouldn't have been able to record the flash crash which occurred across US stock indexes on the 6th of May 2010 from 2:32 pm EDT until 3:08 pm EDT, as prices rebounded shortly afterwards.

An alternative to sampling data in fixed time intervals is the so-called intrinsic time data sampling. In this approach, data is sampled by the observance of significant events in the market [26, 14, 33]. The rationale is to record key events in the market representing significant price movements (e.g. price changes by 2%), which would normally be missed by traditional physical time methods. There have been documented different intrinsic time sampling techniques, e.g. perceptual important points [16, 17], turning point [34], zigzag [9, 27], and more recently directional changes (DC) [19, 32].

Under the Directional Changes (DC) paradigm, a threshold θ is defined, which is used to detect significant price changes. The market is then summarised into upward and downward trends. Each of these trends consists of a directional change (DC) event, which is usually followed by an overshoot (OS) event. Using different threshold values allows the detection of different events and, as a consequence, the creation of different trend summaries. Therefore, the DC framework focuses on the size of a price change as time varies, while under physical time, the time interval is fixed (e.g. daily closing prices).

While the theory behind DC tells us that a DC event is usually followed by an OS event, we have found from preliminary experimental results that there can exist a high number of cases where a DC event is not followed by an OS event; instead it is followed by a DC event in the opposite direction [2]. As a result, we can have datasets with extremely low number of DC events with a corresponsing OS event, even as low as 14.77%—this is, of course, threshold-dependent. Therefore, it cannot be assumed that every DC event will be followed by a corresponding OS event.

In this work, we are interested in investigating this further by introducing a classification step, which predicts whether a DC event is be followed by an OS event. Having this knowledge will allow us to have a better understanding of market dynamics and increase trading profitability. To validate this, we apply a classification step to three different DC-based trading algorithms and report how the trading performance is affected. Our goal is to demonstrate that the introduction of the classification step can significantly improve the profitability of DC-based trading strategies and also outperform other non-DC-based trading strategies, such as technical analysis and buy-and-hold. To achieve this, we undertake rigorous experiments using data from 20 Forex currency pairs over 1,000 different datasets; thus making our results robust and generalisable.

The remainder of this paper is organized as follows. Section 2 presents a summary of directional changes and relevant literature. Section 3 presents our methodology, detailing how we incorporated classification algorithms in the prediction of whether an OS event occurs or not. Section 4 presents the setup of our experiments, while Section 5 presents our results and analysis. Finally, Section 6 concludes this article and discusses future work.

2 | DIRECTIONAL CHANGES

2.1 | Overview

A DC event is identified by price changes defined by a user-specified threshold value. DC events are divided into upturn and downturn events. Once a DC event is confirmed, an overshoot (OS) event usually follows; an OS event finishes once a DC event in the opposite direction is confirmed. A DC trend, upward or downward, consists of the combination of a DC and an OS event. Different thresholds generate different event series. Smaller thresholds create higher number of DC events than larger thresholds, which produce fewer events.

Let us now look at Figure 1, where we present how we can summarise a physical-time price series into DC and OS events. In this example, we summarise price movements with two different thresholds, namely $\theta = 0.01\%$ (lines in red) and $\theta = 0.018\%$ (lines in blue). Price changes below θ are not considered a significant event. Price changes above θ are considered significant events, and divide the market into uptrends and downtrends. Solid lines represent DC events, and dashed lines represent OS events. For example, under $\theta = 0.01\%$, between Points A and B we have a downturn DC event followed by a downward OS event from Point B to C; when a trend reversal occurs, an upturn DC event starts from Point C to D. Lastly, between Point D and E it is an upward OS event. The price point where a DC trend begins or ends is called *DC extreme point (DCE)*; under $\theta = 0.01\%$, Points A, C, and E are DC extreme points.

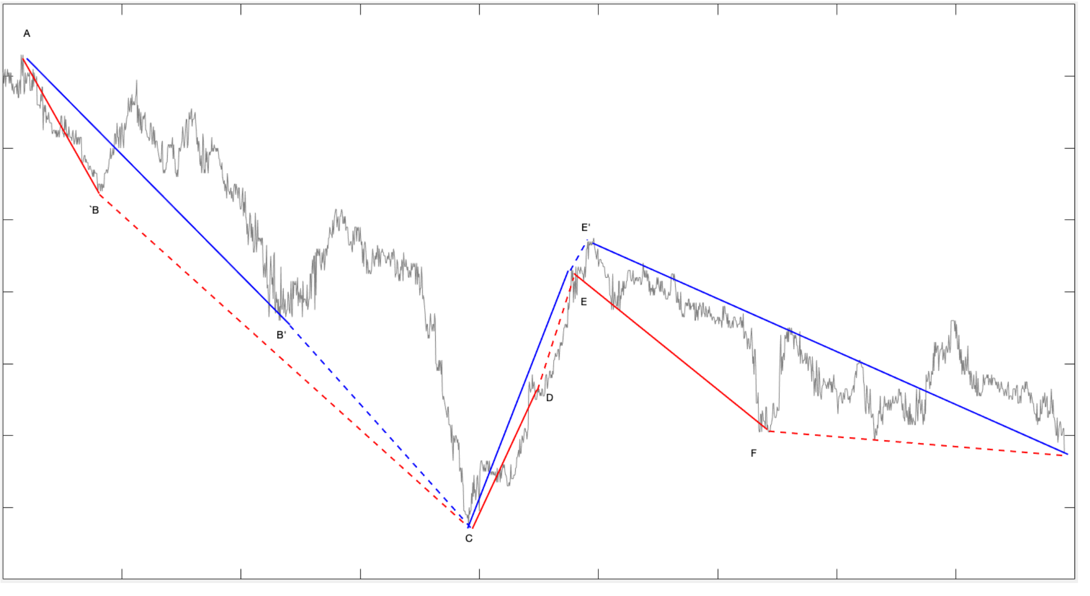


FIGURE 1 Directional changes for the GBP/JPY FX currency pair. The red lines represent events created by a threshold $\theta = 0.01\%$, and the blue lines events created by a threshold $\theta = 0.018\%$. DC events are denoted by solid lines, and OS events by dashed lines. Under $\theta = 0.01\%$, we summarise data as follows: Downturn DC event: Point $A \mapsto B$; Downward OS event: Point $B \mapsto C$; Upturn DC event: Point $C \mapsto D$; Upward OS event: Point $D \mapsto E$; Downturn DC event: Point $E \mapsto F$. Under $\theta = 0.018\%$, we summarise data as follows: Downturn DC event: Point $A \mapsto B'$; Downward OS event: Point $B' \mapsto C$; Upturn DC event: Point $C \mapsto E$; Upward OS event: Point $E \mapsto E'$. DC Extreme points (DCE): Points A, C, E, and E'. DC Confirmation points (DCC): Points B, B', D, E, and F.

Under $\theta = 0.018\%$ (lines in blue), we obtain a different set of events: from A to B′: a downturn event; from B′ to

C: a downward OS; from C to E: an upturn DC event; lastly, from Point E to E′ we have an upward OS trend.

Note that we can only confirm a DC event in hindsight, i.e., after there has been a price change of θ. For instance, under $\theta = 0.01\%$ we would not know we are in an upward trend until we have reached Point D. This point is called a *DC Confirmation point (DCC)*. Before Point D, one would consider that the market has been in a downward trend since Point A. Similarly, we would not know the trend has reversed from upward to downward until we have reached the DCC Point F. It is therefore crucial to be ablel to accurately predict when a trend reversal will take place.

Directional changes have offered traders new perspectives on price movements and have led to new research directions, which were not possible under the physical time price summaries. In the following section, we briefly present some of the recent DC works.

2.2 | Review of recent DC literature

According to [29], scaling laws refers to empirical findings accepted as truth due to their consistency. [20] was the first work deriving scaling laws using directional changes. Since then, several other works have discovered new scaling laws. Some examples are [18], which discovered 17 new scaling laws; [19], which uncovered 12 additional laws; [5], which added 4 more scaling laws; and [7], which further contributed 5 scaling laws. Furthermore, [32], presented 4 new DC indicators for profiling financial markets. Since then, more DC indicators have been introduced [31] and used for profiling [15, 8].

Another body of DC work has focused on estimating trend reversal and trading. For example, [6, 5] combined directional change with trend following and contrary trading technical indicators, and developed new trading strategies. Furthermore, [11, 10] created a DC-based trading strategy, named 'DBA', and reported mean returns of around 14%. Directional changes have also been combined with machine learning algorithms, e.g. a neuro-fuzzy system [22]; genetic programming [21]; and decision trees [3]; but also with econometric models, like an F-GARCH [4].

2.3 | DC-OS event length relationships

One of the most interesting scaling laws is the one that describes the relationship between the duration (length) of DC and OS events. More specifically, [19] found empirical evidence that the duration of an OS event is on average twice the duration of its corresponding DC event (Equation 1). Such laws can be leveraged to traders and take trading actions before the end of a trend is reached. Similar empirical observations were made in [24, 23]. However, they proposed a more generic formulation, where the DC and OS length relationship was expressed as a linear function with a M constant, where M is the average DC event length to OS event length ratio for the given datasets (Equation 2). Lastly, [2] presented a genetic programming (GP) algorithm, which undertook a symbolic regression task, and evolved linear and non-linear formulas (Equation 3). Using a GP allowed to search for the best function that describes the DC and OS length relationship, without the need to make any assumptions about the form of this relationship.

$$OS_l \approx 2 \times DC_l \tag{1}$$

$$OS_l = M \times DC_l; M > 0 \tag{2}$$

$$OS_l = f(DC_l) \tag{3}$$

As we mentioned earlier, it is possible that not all DC events have a corresponding OS event; instead another DC from the opposite direction could follow. Empirical observations in [2] showed that there could be datasets with as little as 14.77% of DC events with a corresponding OS event. This is an important observation, because it indicates that Equations 1–3 are meaningful only in those cases where a DC is followed by an OS. Let us use an example to clarify this issue. Let us assume that an upward DC event lasts for 2 days, and then it is directly followed by a downward DC event that lasts for 4 days. If we attempt to apply any of the above equations, say Equation 1, we would be predicting that the first DC event would be followed by an OS event, of the same direction, which would end 4 days later (since $OS_l \approx 2 \times DC_l$). A trader with the above expectation could thus wait until the end of the 4th day of the OS event to take a sell action, where the price is expected to be at its highest point. However, as in this case there has already been a directional change, the trader would end up selling at a much lower price, since we are in a downward trend. In other words, $\text{OS}_l = 2\times\text{DC}_l$ is a OS length approximation that does not hold for all DC events.

To avoid such issues, we are proposing to introduce a classification step before the prediction of trend reversal by Equations 1–3. During this step, we will be predicting whether a DC event will be followed by a corresponding OS event. Only when there is a corresponding OS event, we will be applying the above trend reversal equations.

3 | METHODOLOGY

Our aim is to be able to identify cases where a DC event is followed by an OS event. In order to achieve this, we will be introducing a binary classification step. The first class value will be that a DC is followed by an OS event (we call this αDC), and the second class value will be that the DC is not followed by an OS event (we call this βDC). When the classifier predicts αDC, we apply Equations 1–3 to estimate the relationship between DC and OS lengths in order to predict the trend reversal point (i.e., end of the OS event). On the other hand, when the classifier predicts βDC, the trend reversal point will be estimated as the end of the DC event. This process is summarised in Figure 2—to evaluate its profitability, we embed the classification step as part of a trading strategy.

The remainder of this section is organised as follows. Section 3.1 presents the length estimation task, where we use Equations 1–3 to estimate the relationship between DC and OS lengths; Section 3.2 describes the classification step used to predict whether there is an OS even or not; and finally, Section 3.3 presents the trading strategy.

3.1 | OS Length estimation technique

As we can observe in Figure 2, our framework consists of three main steps: classification, OS length estimation, and embedding this prediction into a trading strategy. However, if we were to implement the classification step first, the errors made in this step would be carried forward to the step of estimating OS length. For example, let us assume that (1) a dataset consists of 10 DC events, 8 DC events are followed by an OS event and 2 DC events are not; and (2), a classifier predicts that all 10 DC events have a corresponding OS event. In this case, when we move to the OS length estimation step and apply any of Equations 1–3, the information (data) from all 10 events will be used to construct the OS length estimation models, incorrectly including the information from the 2 events that we already know that are not followed by OS events.

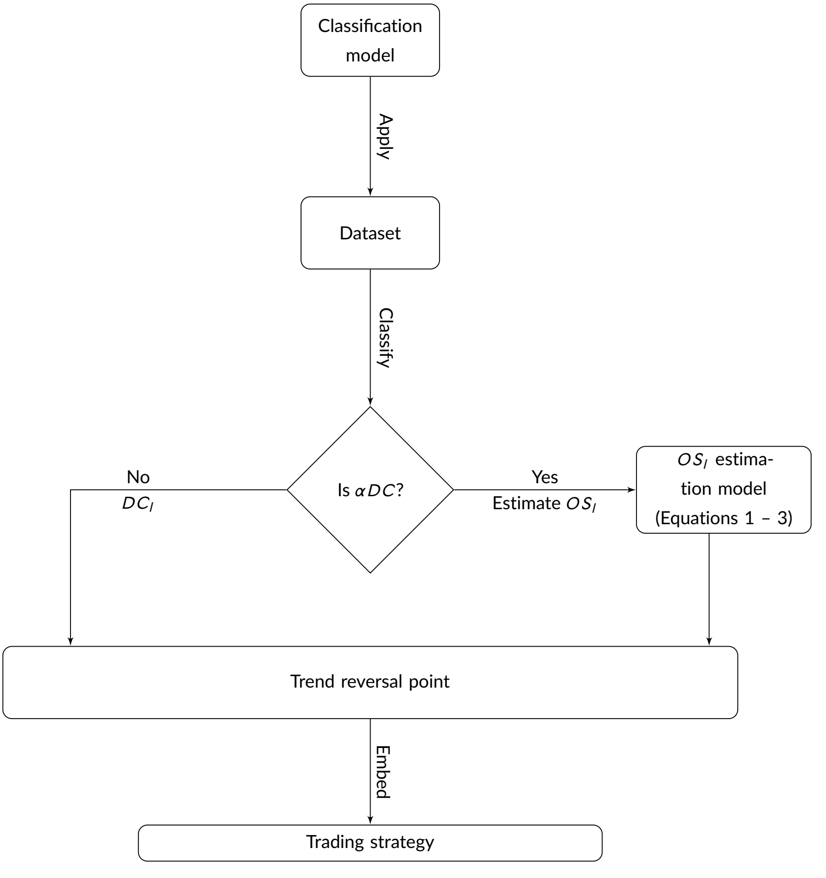


FIGURE 2 Proposed trend reversal estimation under directional changes. When a DC trend is predicted to be followed by an OS event, then it will reverse at DCE, which is calculated by the sum of the DC and OS event lengths; the OS length is estimated by one of Equations 1-3. When a DC trend is predicted not to have a corresponding OS event, then the trend reversal point is going to be the DCC.

To avoid having this classification error being carried forward, we have performed the length estimation task first, before the classification step, and under perfect foresight on the training data. Having perfect foresight allows us to identify the DC events that are followed by an OS event and apply Equations 1–3 to that data only. As a result, we eliminate the problem of classification errors affecting the length estimation step. Equations 1–3 thus now exclude noise (i.e., DC events without a corresponding OS event), and can focus on data that matters. Lastly, we should note that because we perform the above task only on the training set, we avoid introducing bias when we eventually apply the selected length estimation model to the (unseen) test data.

3.1.1 | $OS_l \approx 2 \times DC_l$ (Factor-2)

This approach is based on empirical observations in the Forex market, which led to the creation of several scaling laws [19]. According to [19], the OS length is on average twice the length of its corresponding DC event. Therefore, whenever we want to predict the OS_l, we measure the relevant DC_l and multiply by 2. Due to the fact that the DC-OS length relationship is dependent on a factor of 2, we call this approach *Factor-2*.

We would like to once again note that the confirmation of a trend change—upward to downward DC and vice-versa—is only detected in hindsight, i.e., only when the corresponding DC event is confirmed. Therefore, we use the above formula to predict the length of the OS event, and thus anticipate when the current trend will end, only when a DC event is confirmed and classified to have a corresponding OS event.

3.1.2 | $OS_l = m \times DC_l$ (Factor-M)

This equation expresses a linear relationship between DC and OS lengths, where M is a constant denoting the average DC/OS event length ratio for a given dataset. Because $OS_l \approx 2 \times DC_l$ is an approximation and does not take into account the underlying dataset, this new formula offers a tailored estimation of the DC:OS event length ratio on a given dataset. Since the factor that defines the relationship of DC-OS length is equal to an m constant, we call this approach *Factor-M*. In this approach, the average length of each OS event in the training set is calculated. In addition, there is a distinction between the average length of upwards and downwards OS events. This is because the DC:OS ratio could be different for the two types of trend.

For a detailed presentation of this approach, we refer the reader to [24].

3.1.3 | $OS_l = f(DC_l)$ (Reg-GP)

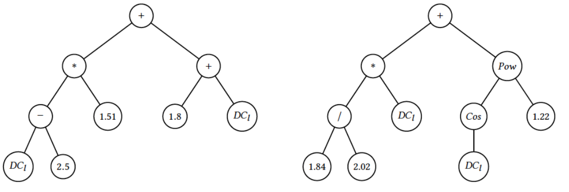


FIGURE 3 Sample GP trees

This approach uses a GP algorithm to evolve a (non-)linear symbolic regression model to represent the relationship between DC and OS lengths. We call it *Reg-GP*. Contrary to standard linear regression techniques (e.g. linear regression), where we fit coeffients under a given model, symbolic regression allows us to determine both the functional form of the model, as well as its coefficients. This has the advantage that we do not need to make assumptions about the relationship of DC and OS length. GP will thus create mathematical equations describing the relationship between DC and OS length, without making any assumptions of their form. It is based on the Darwinian principle of evolution, where it creates a population of unfit (usually random) programs (equations describing the DC-OS length relationship in our case) and searches the space of mathematical expressions to find linear and non-linear models that best fit the dataset. GP is considered one of the state-of-the-art methods for symbolic regression [28].

Table 1 presents the configuration of the GP. Figure 3 presents two sample trees from the GP we have created. The first tree represents the equation that calculates OS length as $((DC_l - 2.5) \times 1.51) + (1.8 + DC_l)$ and the second trees represents the equation $((\frac{1.84}{2.02}) \times DC_l) + (\cos(DC_l)^{1.22})$, where DC_l in both equations is the length of DC event.

TABLE 1 GP algorithm configuration

Configuration	Value
Function set	+, -, /, *, sin, cos, pow, log, exp.
Terminal set	DC_l, ephemeral random constant.
Genetic operation	elitism, subtree crossover, subtree mutation.

The fitness function of the GP is the RMSE between actual OS length (OS_l) and estimated OS length ($\hat{OS_l}$).

$$\varepsilon = \sqrt{\frac{\sum_{i=1}^{N}(OS_l - \hat{OS_l})^2}{n}} \tag{4}$$

where n is the sample size.

For a detailed presentation of the GP algorithm, we refer the reader to [2].

3.1.4 | OS length estimation outputs

As we have previously mentioned, different DC thresholds generate different DC summaries. We select the optimal thresholds per dataset from a pool of thresholds. To do this, we first use each threshold to generate its own DC summary. Then, we apply each one of Equations 1 –3 to each DC summary, and thus obtain different OS length estimation models, one for each equation and DC summary. Finally, we rank each model by RMSE. Note that the process for each equation is independent from each other; thus each equation returns a different OS length estimation model and DC summary.

The above process returns 3 outputs: (i) the best OS length estimation model, (ii) the best threshold, and (iii) the respective DC summary of that particular threshold. Figure 4 illustrates this process.

3.2 | Classification Step

The next step is to predict whether a DC event is followed by an OS event (αDC) or not (βDC). For this step, we use Auto-Weka [30], which is an automated machine learning tool that searches in the space of 39 classification algorithms and their hyperparameters. We run Auto-Weka 10 times for each dataset and we select the algorithm with the highest f-measure. To avoid bias, we apply ten-fold cross-validationto the training set; therefore we do not expose the testing set to Auto-Weka. The above process allows us to obtain a tailored classifier with tailored hyperparameters for each dataset.

Classifiers are trained on 6 DC-based attributes. These are: difference between the price of upturn/downturn and DCC (DC confirmation) points (A1); difference between the time of upturn/downturn and DCC points (A2); speed of the prices change from the start of a trend and DCC point (A3); price at previous confirmation points (A4); (boolean) information on whether the immediate previous DC event has a corresponding OS event or not (A5); (boolean) indication on whether the start and end time of a DC event are equal (A6). Attributes A1 and A2 were introduced in [19], while A3-A6 were first introduced in [1]. Table 2 summarises these attributes.

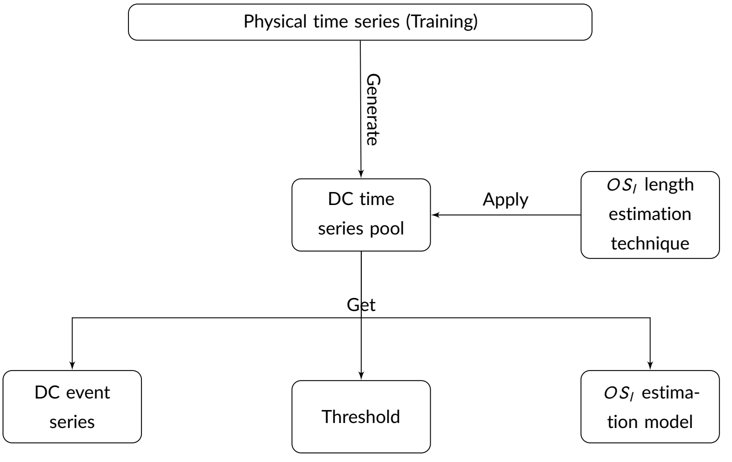


FIGURE 4 Proposed framework for obtaining an OS length estimation model, and selecting threshold and DC events with high DC:OS event ratio.

Figure 5 summarises the classification process. After obtaining the output of the OS length estimation process presented in Figure 4, namely 'Best DC event series', we use it to calculate the values of attributes A1-A6. Then, we use Auto-Weka on the training set to obtain the best classification model per dataset.

Each obtained classification model will then make predictions in the (unseen) test set on whether a DC event is followed by an OS event. If this is the case, we will then use an OS length estimation model (1–3), which we discussed in Section 3.1, to predict the trend reversal point (i.e., the sum of DC and OS length). If the classification model predicts that there is no corresponding OS event, the end of the current DC event becomes the end of trend. This process was also summarised in Figure 2.

3.3 | Trading Strategy

So far, we have used the classification and OS trend estimation steps to predict the end of a trend. The last step of our framework is to embed the trend reversal prediction into a trading strategy, so that we can test its effectiveness by reporting financial metrics, such as return and risk perfomance. Next, we provide information of our trading strategy.

3.3.1 | Overview

Before we present the details of our prososed trading strategy, we should define two important notions: opening and closing a position. We open (close) a position when we sell (buy) the base currency (e.g., GBP) and buy (sell) the quoted currency (e.g., JPY).

In order to open a position, there are two requirements: (i) there is not an already open position, and (ii) the return from opening the position would be positive, after accounting for transaction costs. If the above requirements hold, we open a position at the extreme point of an upward DC trend. Similarly, to close a position, there are two requirements that need to hold: (i) there is an existing open position, and (ii) the return from closing the position would be positive after accounting for transaction costs. If these conditions hold, we close the position at the extreme point of a downward DC trend.

Attributes	Name	Description
A1	DCprice	Price difference between the upturn/downturn point and the directional change confirmation point.
A2	DCtime	Time difference between the upturn/downturn point and the directional change confirmation point.
A3	$Sigma'$	Speed at which prices change from the start of a trend to directional changes confirmation point.
A4	DC_{t-1}price	Price at previous confirmation points.
A5	DC_{t-1}OS	Indicates whether the immediate previous DC event has a corresponding OS event or not.
A6	FlashEvent	Indicates whether the DC event start time and end time are equal or not.

TABLE 2 Classification attributes

The extreme point in both of the above cases can be either at αDC or βDC, depending on the prediction of the classification model. When the above requirements are not met, no trading takes place. All transactions take place by using our entire capital. Transaction cost value is 0.025% for each transaction. Algorithms 1 and 2 provide a detailed overview of the opening and closing strategies.

Algorithm 1 Trading rule for selling base currency

Require: Sell rule
  if DC is in upward trend && There is no open position **then**
    if Is βDC && Return is not negative **then** Open a position at DCC point
    else if Is αDC && DC trend does not reverse before estimated DCE point && Return is not negative **then** Open a position at estimated DCE point
    else Hold
    end if
  end if

3.3.2 | Trading strategy evaluation

Our trading strategy is evaluated by three metrics: return, maximum drawdown (MDD) and sharpe ratio. Return (Equation 5) is calculated after accounting for transaction costs. MDD (Equation 6) measures the downside risk by calculating the maximum observed loss from a peak price to a trough before a new peak is reached. Finally, the

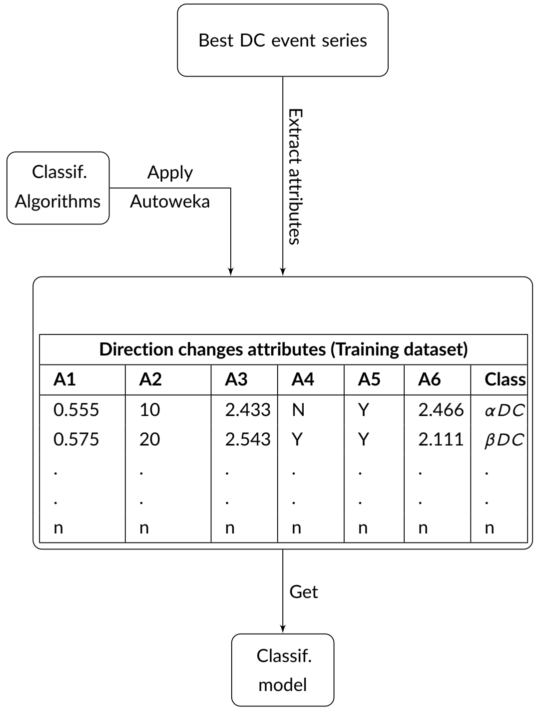


FIGURE 5 Proposed framework for creating a classification model to predict whether a DC event is followed by an OS event (αDC) or not (βDC)

sharpe ratio (Equation 7) reports the risk-adjusted return, after accounting for the return by a risk-free asset (e.g. a government bond at 0.25% rate).

$$R = (Q - TC) * FXrate \tag{5}$$

where Q is the trading quantity, TC is the transaction cost, and $FXrate$ is the price that the trade takes place.

$$MDD = \frac{P_{trough} - P_{peak}}{P_{peak}} \tag{6}$$

where P_{trough} is the price at the trough point, and P_{peak} is the price at the peak point.

$$SharpeRatio = \frac{R - i}{\sigma_{return}} \tag{7}$$

where R is the return, i is the value of the risk free asset, and σ_{return} is the standard deviation of the return.

Algorithm 2 Trading rule for buying base currency

Require: Buy rule
if DC trend is downward && There is an open position **then**
  if Is βDC && Return is not negative **then** Close position at DCC point
  else if Is αDC && DC trend does not reverse before estimated DCE point && Return is not negative **then** Close position at estimated DCE point
  else Hold
  end if
end if

4 | EXPERIMENTAL SETUP

The datasets used in our experiments consist of high-frequency (10-minute interval) data for 16 FX currency pairs from March 2016 to February 2017 and an additional 4 currency pairs from June 2013 to May 2014. These pairs are presented in Table 3. Each month is considered as a separate physical-time set. We used the first two months for each currency pair for parameter tuning, which resulted in 200 datasets (5 DC thresholds $\times$ 20 currency pairs $\times$ 2 months). The remaining 10 months for each dataset were used for the main experiments, which resulted to 1000 datasets (5 DC thresholds $\times$ 20 currency pairs $\times$ 10 months). All datasets were using a 70:30 split between training and test sets.

Currency pairs		**Currency pairs**	
March 2016 to February 2017			
AUD/JPY	Australian \$ / Japan. Yen	EUR/NOK	Euro / Norwegian Krona
AUD/NZD	Australian \$ / N. Zeal. \$	GBP/AUD	British Pound / Australian \$
AUD/USD	Australian \$ / US \$	NZD/USD	New Zealand \$ / US \$
CAD/JPY	Canadian \$ / Japan. Yen	USD/CAD	US \$ / Canadian \$
EUR/AUD	Euro / Australian \$	USD/NOK	US \$ / Norwegian Krona
EUR/GBP	Euro / British Pound	USD/JPY	US \$ / Japan. Yen
EUR/CAD	Euro / Canadian \$	USD/SGD	US \$ / Singaporean Dollar
EUR/CSK	Euro / Czech Krona	USD/ZAR	US \$ / South African Rand
June 2013 to May 2014			
EUR/USD	Euro / US \$	GBP/CHF	British Pound / Swiss Franc
EUR/JPY	Euro / Japan. Yen	GBP/ USD	British Pound / US \$

TABLE 3 FX currency pairs used in our experiments.

As each threshold produces a different DC summary, we will be evaluating 5 different DC thresholds for all datasets (both tuning and non-tuning). These thresholds are the best thresholds dynamically selected during the OS length estimation step (see Section 3.1). When reporting results in Section 5, we will be presenting the mean performance of each algorithm, over the 5 DC thresholds.

4.1 | Parameter tuning

There are two tasks that require parameter tuning: the classification task, and the OS length estimation task. As we have already mentioned, the classification task uses Auto-Weka to search in the space of algorithms and their hyperparameters. The only parameter that we needed to tune was Auto-Weka's execution time. Higher execution times allow Auto-Weka to consider more algorithms and more hyperparameter settings. We considered five different execution times: 15, 30, 45, 60, and 75 minutes. After experiments in a training and validation set, we chose the 60 minute duration based on average f-measure results.

With regards to the OS length estimation task, the only necessary tuning was for the GP algorithm (Equation 3). We used the I/F-Race package [25], which implements a racing method to select the best configuration of an optimisation algorithm. We present the selected GP parameters in Table 4.

Parameter	
Population	500
Generation	37
Tournament size	3
Crossover probability	0.98
Mutation probability	0.02
Maximum depth	3
Elitism	0.10

TABLE 4 GP parameters

4.2 | Trading experimental setup

As we have already mentioned, after completing the classification and OS length estimation tasks, we can use Equations 1–3 to predict when a DC trend is going to end. We embed these equations into the trading strategy we presented in Section 3.3.

Our goal is to investigate if the introduction of the classification step is beneficial to the trading performance. Thus, we will be comparing each DC algorithm to a version with and without classification. In addition, we will be also experimenting with non-DC algorithms, to investigate how the DC performance compares to other state-of-the-art trading algorithms.

We present below all algorithms used in our experiments.

4.2.1 | DC-related benchmarks

(C+) Factor-2

Originally presented in [19], this DC trend reversal approach says that on average the OS event length is twice the DC event length. In this strategy, Equation 1 replaces the classification and regression steps of our methodology. Thus, the trend reversal point is the point where the OS event length is twice the DC length. *Factor-2* is the approach without classification, and *C+Factor-2* is the same approach with classification.

(C+) Factor-M

This is a DC trend reversal approach originally presented in [24], where the DC-OS relationship is expressed by a linear function. The constant M is tailored to each dataset and describes the above linear relationship. In this strategy,we use Equation 2 instead of the classification and regression steps of our methodology. Thus, the trend reversal point is tailored to each dataset. *Factor-M* is the approach without classification, and *C+Factor-M* is the same approach with classification.

(C+) Reg-GP

This is a DC trend reversal approach presented in [2], where equations are created by a symbolic regression GP. In this trading strategy, we embed Equation 3. *Reg-GP* is the approach without classification, and *C+Reg-GP* is the same approach with classification.

p-trading

We obtain the training set probability p of a DC event being followed by an OS event. At the directional change confirmation (DCC) point, we decide with this p probability whether a DC trend has a corresponding OS event; if true, we apply any of the above 3 OS length estimation models to estimate OS event length (i.e. predict trend reversal point which is our estimated directional change extreme point). If false, the DCC point is the estimated extreme point and we apply the trading strategy and see if a trade at that point can be profitable. The motivation behind this scenario is to demonstrate that the introduction of the classification step is advantageous and outperforms the informed decision of having or not having an OS event, based on a probability that is tailored to the training set. As there are three different OS length estimation models, there are as a result three variations of the tailored trading benchmark:

- *p+Factor-2*
- *p+Factor-M*
- *p+Reg-GP*

Trade at DCC point

In this scenario, we will always trade as soon as a directional change has been confirmed, i.e., at the DC confirmation point (DCC). The motivation behind this scenario is to investigate the trading profitability if we were not to take into account the OS events at all and instead only focus on the DC events. Provided that our proposed classification algorithms outperform this scenario, it would again demonstrate that the introduction of the classification step is advantageous and better than not having classification and the knowledge of the OS length.

As there are three different OS length estimation models, there are as a result three variations of the random trading benchmark:

- *DCC+Factor-2*
- *DCC+Factor-M*
- *DCC+Reg-GP*

4.2.2 | Non-DC benchmarks

Technical analysis trading strategy

We also use three popular technical indicators to make trading decisions. These indicators are:

- Relative Strength Index (RSI)
- Exponential Movement Average indicator (EMA)
- Moving Average Convergence Divergence (MACD)

Buy-and-hold

Buy and hold is a well-known benchmark for trading algorithms. In the first month of the non-tuning dataset, we buy the quoted currency; after the end of the 10-month period, we sell in exchange for the base currency.

5 | RESULT AND ANALYSIS

In this section we present the results of our experimental work. We first present the regression results in Section 5.1, and then we present the trading results in Section 5.2.

We would like to remind the reader that the goal of our work is twofold: (i) Demonstrate that the introduction of the classification step can significantly improve the profitability of DC-based strategies, and (ii) Demonstrate that DC-based strategies, which use a classification step to predict whether a DC event is followed by an OS event, are able to be profitable and outperform other DC and non-DC-based trading strategies.

5.1 | Regression result

Table 5 presents the average RMSE result of the OS length estimation step over the 5 DC thresholds and the 10 months of data per currency pair for GP-Reg, Factor-M, and Factor-2. For each of these three algorithms, we present the RMSE for two variations: (1) with the classification step and (2) without the classification step. For the variation that includes the classification step (denoted with prefix $C+$), we estimate the OS event length only in DC trends that have been classified as αDC. RMSE is calculated regardless of the prediction of the classification (i.e., being correct or incorrect). In cases where the classification algorithm incorrectly predicts that there is an OS event, the length returned by the OS estimation algorithm is compared to length zero (0). When the classification algorithm incorrectly predicts that there is not an OS event, the length of the OS event is compared to length zero (0). In other words, in both cases, the incorrect predictions (FP, FN) from the classification step are accounted for the RMSE.

From Table 5 we can observe that C+Reg-GP has the lowest average RMSE (18.6170) across all six algorithms. C+Reg-GP also has 11 cases (out of the 20 currency pairs) that returned the lowest RMSE per currency pair, Reg-GP had 4 such cases, C+Factor-M 3 cases, and C+Factor-2 2 cases. More importantly, we can observe that the average RMSE for each algorithm with the classification step has returned a lower average RMSE when compared to its respective variation without classification: C+Reg-GP (18.6550) vs Reg-GP (20.3216), C+Factor-M (20.5592) vs Factor-M (34.4474), and C+Factor-2 (21.3247) vs Factor-2 (25.7951). It is also worth noting that the classification accuracy (presented in brackets) is quite high, usually ranging between 70% and 85%. As we have hypothesised, the high accuracy appears to have played an important role in reducing the average RMSE for all three algorithms (Reg-GP, Factor-M, and Factor-2).

An interesting observation is that while the EUR/CSK pair has relatively low classification accuracy across all variants (55-58%), their average RMSE still outperforms the variants without the classification step. We further looked into this and found that the EUR/CSK currency pair has the lowest number of DC events (55 in the training set, 18 in the test set), while the average number of DC events for all other currency pairs is 194 in training and 60 in test. In addition, the length of DC events (i.e. number of physical time data points making up a single event) is the highest for EUR/CSK (46 in training, 32 in test), compared to an average of 12 in both training and test for all other currency pairs. This means that EUR/CSK experiences a smaller number of DC events, followed by long OS events. When the algorithms are run without the classification step, every DC event is assumed to be followed by an OS event. Given that the DC events are long for EUR/CSK, as a consequence, the OS events are long, as well. Therefore, every time that an OS length estimation algorithm (Reg-GP, Factor-M, Factor-2) makes a prediction when there is no OS event, this results in a larger RMSE. The classification step, although it has a lower accuracy (55-58%), significantly reduces the RMSE simply by reducing the number of times that the OS estimation algorithm is used when there is no OS event. In this case, the classification step should not be evaluated only on the merits of classification accuracy, but on the actual effect it has on the regression error.

To further understand the above results, we have performed the Friedman non-parametric statistical test, under the null hypothesis that all algorithms come from the same continuous distribution. We present these results in Table 6. The first column presents the average rank of each algorithm. The second column presents the adjusted p-value of the test when that algorithms average rank is compared to the average rank of the control algorithm (i.e. algorithm with the best rank). The adjusted p-value is calculated with the Hommel post-hoc test. As we can observe, C+Reg-GP ranks first and statistically outperforms all other algorithms at the $\alpha = 0.05$ level. More importantly, C+Reg-GP outranks Reg-GP, C+Factor-2 outranks Factor-2, and C+Factor-M outranks Factor-M.

To sum up: (i) introducing the the classification step (C+Reg-GP, C+Factor-M, C+Factor-2) to existing DC-based algorithms (Reg-GP, Factor-M, Factor-2) has reduced the average predictive error, and (ii) the DC algorithms that are using the classification step outrank their respective algorithm without classification. Our interest now shifts to the trading step in order to investigate whether the introduction of classification step also leads to an increase in trading profit margins (in addition to reduced OS length estimation error, as we have just seen in Section 5.1).

5.2 | Trading result

5.2.1 | Comparison against DC-based and technical analysis algorithms

Table 7 presents the average returns per currency pair.. When value 0.00 is reported, it means that the trading strategy took a hold action throughout the 10-month period. The best value for each row (currency pair) is shown in boldface. Best value among the different variants of the same algorithm (Reg-GP, Factor-M, Factor-2) is underlined, i.e. for AUD/NZD C+Reg-GP has the best return across all algorithms, as well as among the Reg-GP algorithm variants, namely C+Reg-GP, Reg-GP, *p*+Reg-GP, DCC+Reg-GP.

In terms of average results, C+Reg-GP has the highest return (0.2247) across all algorithms. Furthermore, all versions that have introduced the classification step have the highest average return for their respective group (C+Reg-GP: 0.2247, C+Factor-M: 0.0684, C+Factor-2: 0.1186). It is also important to note that these three versions show a strong positive average return, whereas all other algorithms experience negative average returns, with the exception only of Factor-2, which shows marginally positive average returns of 0.0260. It is also worth highlighting that Forex market is open 24 hours a day in different parts of the globe and at any point in time, there is at least one market open and has overlapping open markets for few hours. Thus, an average return of the scale of 0.2247% (C+Reg-GP) after discounting' transaction cost will have a significant cumulative effect in the long run.

Having a closer look at the individual currency pairs, we can observe that C+Reg-GP outranks Reg-GP, *p*+Reg-GP, and DCC+Reg-GP in 14 out of the 17 pairs (the remaining 3 pairs have 0.00 return, thus no trading took place). Similarly, for the second group of DC algorithms, C+Factor-M ranks first with 13 cases. Lastly, C+Factor-2 had the highest return in 10 currency pairs.

TABLE 5 Average RMSE values for each OS length estimator algorithm measured over 1000 datasets consisting of 5 different dynamically generated thresholds tailored to each DC dataset, 20 currency pairs, and 10 months of 10-minute interval data for each currency pair. In brackets we also report the classification accuracy, for reference (for C+Reg-GP, C+Factor-M, C+Factor-2). The Best value for each row (currency pair) is shown in boldface.

Algorithms	C+Reg-GP	Reg-GP	C+Factor-M	Factor-M	C+Factor-2	Factor-2
AUD/JPY	**15.5670** (0.851)	15.6270	17.1570 (0.778)	25.5270	18.4720 (0.782)	22.2690
AUD/NZD	27.368 (0.805)	**24.3320**	27.4110 (0.806)	51.2420	31.9260 (0.761)	41.5920
AUD/USD	**11.5800** (0.829)	12.8140	12.7200 (0.768)	16.0950	11.8270 (0.745)	14.0600
CAD/JPY	**11.8430** (0.820)	18.7850	14.6860 (0.779)	39.9700	16.8800 (0.764)	27.2510
EUR/AUD	21.1710 (0.821)	20.5690	20.2010 (0.799)	25.7280	**14.7520** (0.751)	19.7490
EUR/CAD	**16.2050** (0.839)	17.7190	21.0950 (0.784)	23.1830	22.6420 (0.750)	24.8670
EUR/CSK	**41.9900** (0.557)	52.9490	46.0270 (0.581)	188.6080	63.0420 (0.565)	83.8450
EUR/GBP	24.1730 (0.825)	22.6350	25.5870 (0.766)	31.4300	**17.2120** (0.752)	18.7900
EUR/JPY	**19.9650** (0.821)	21.1170	23.4540 (0.758)	28.1620	23.1640 (0.748)	25.2040
EUR/NOK	**13.7170** (0.818)	13.7620	20.4120 (0.727)	27.2010	19.5710 (0.728)	22.4990
EUR/USD	28.2600 (0.806)	31.0610	**26.8990** (0.786)	38.5320	27.6690 (0.762)	30.0380
GBP/AUD	15.1380 (0.837)	**14.7190**	19.2820 (0.832)	21.6700	14.8810 (0.780)	17.9100
GBP/CHF	**15.9610** (0.831)	17.2040	17.5260 (0.784)	19.3580	21.4210 (0.769)	23.6690
GBP/USD	19.2040 (0.851)	24.8890	**17.8250** (0.790)	21.2230	25.3210 (0.746)	27.7780
NZD/USD	**10.2300** (0.848)	10.5880	11.0920 (0.772)	14.7310	13.1350 (0.773)	15.8960
USD/CAD	26.9340 (0.797)	**26.8180**	27.1330 (0.766)	34.6540	27.5190 (0.739)	29.3150
USD/JPY	**13.7040** (0.850)	14.5430	15.9860 (0.774)	17.9980	16.0310 (0.777)	18.3260
USD/NOK	7.7180 (0.887)	**7.3570**	9.96900 (0.813)	14.1280	8.1830 (0.792)	10.7640
USD/SGD	**26.9320** (0.780)	34.1480	31.9440 (0.799)	41.7120	27.4980 (0.720)	34.3600
USD/ZAR	5.4400 (0.877)	4.7960	**4.7770** (0.807)	7.7960	5.3470 (0.813)	7.7200
Average RMSE	**18.8175** (0.816)	20.5687	20.7382(0.773)	34.9169	21.4748 (0.749)	25.9807

To support our findings, we applied Friedman's non-parametric statistical test. The null hypothesis is that the algorithms come from the same continuous distribution. The result of the statistical test presented in Table 8 shows that all three DC versions with the classification step (i.e. C+Reg-GP, C+Factor-M, C+Factor-2) rank the highest, and outperform all other variants without the classification step. In addition, C+Reg-GP ranks first and statistically outperforms all algorithms, apart from C+Factor-M and C+Factor-2, at the 5% significance level.

Even though C+Reg-GP recorded higher returns than other trading strategies, it is important to also take into account the risk taken to achieve it. For this reason, we also present results of MDD (maximum drawdown) and Sharpe ratio. We did not record risk measures for currency pair AUD/JPY, CAD/JPY and USD/JPY, as no trading took place in these markets.

Table 9 presents the MDD result. In terms of overall average MDD, C+Reg-GP has the lowest (best) average value of 0.1259. In terms of individual currency pairs and performance of each group of algorithm, we can observe that C+Reg-GP returned the lowest MDD value in 13 out of the 17 cases, when compared to Reg-GP, *p*+Reg-GP, and DCC+Reg-GP. So once again, the introduction of the classification step has been beneficial. What is interesting, however, is the results for the other two sets of algorithms, Factor-M and Factor-2. It appears there's a trade-off here between higher return and risk, as in both cases, it is actually the variant without the classification step that ranks first in their respective group (both Factor-M and Factor-2 rank first in 15 out of the 17 currency pairs). The classification

TABLE 6 Statistical test results of OS length estimation according to the non-parametric Friedman test with the Hommel post-hoc test. Significant differences at the α = 0.05 level are shown in boldface.

Algorithm	Average Rank	$Adjust_{pHomm}$
C+Reg-GP (c)	1.90	-
Reg-GP	**2.50**	0.3105
C+Factor-M	**2.95**	0.1518
C+Factor-2	**3.15**	0.1038
Factor-2	**4.90**	1.5835E-6
Factor-M	**5.60**	1.9985E-9

variants ranked second in both groups.

In addition, Table 10 presents the Friedman test for Table 9. The null hypothesis is again that the algorithms come from the same continuous distribution. As we can observe, the best ranking algorithm is one that includes a classification step, C+Reg-GP, and it statistically outranks 9 other algorithms at the 5% significance level.

Figure 6 illustrates the Sharpe ratio of the trading strategies. The time period is presented in the x-axis, and the sharpe ratio in the y-axis. The pairs where there are no values are the ones where no trading took place. As we can observe, C+Reg-GP consistently reports positive sharpe ratio, whereas the other strategies have a mixture of both positive and negative values. There are 34 risk-adjusted return summaries, excluding the 6 periods where no trading took place. Out of the 34, C+Reg-GP had positive Sharpe ratio in 28; Meanwhile, Reg-GP, p+Reg-GP, DCC+Reg-GP, C+Factor-M, Factor-M, p+Factor-M, DCC+Factor-M, C+Factor-2, Factor-2, p+Factor-2, DCC+Factor-2, EMA, MACD and RSI had 11, 15, 8, 22, 7, 19, 13, 21, 17, 14, 13, 8, 4 and 11 respectively. Out the 28 positive Sharpe ratio results, 6 where above 0.5, 18 were above 0.2 and less that 0.5; the rest were below 0.2.[1] Table 11, which presents the Friedman test for the result presented in Figure 6, confirms our findings, as C+Reg-GP ranks first and statistically outperforms all other trading strategies at the 5% level. In addition, C+Factor-M and C+Factor-2 rank second and third, respectively, which again demonstrates that the introduction of the classification step is beneficial to the DC algorithms.

[1] A negative value represents a negative return; a value between 0.2–0.3 is in line with the general market; a value between 0.5–1 represents a market-beating performance; a value greater than 1 represents superb performance.

TABLE 7 Average return of trading strategies under 10-minute interval out-of-sample data, 20 different currency pairs, and 10 calendar months. 5 DC datasets were generated using 5 dynamically generated thresholds tailored to each DC dataset. The best value for each row (currency pair) is shown in boldface. Best value among the different variants of the same algorithm (Reg-GP, Factor-M, Factor-2) is underlined.

Dataset	C+Reg-GP	Reg-GP	*p*+Reg-GP	DCC+Reg-GP	C+Factor-M	Factor-M	*p*+Factor-M	DCC+Factor-M	C+Factor-2	Factor-2	*p*+Factor-2	DCC+Factor-2	RSI	EMA	MACD
AUD/JPY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AUD/NZD	**0.2600**	-0.0890	0.0709	0.0110	0.0747	-0.0626	0.1084	0.0223	0.0328	-0.0122	0.0616	0.0248	0.0558	0.0017	0.0047
AUD/USD	**0.2727**	-0.4636	-0.2061	-0.2223	-0.0037	-0.3270	-0.1749	-0.2933	-0.0223	-0.1321	-0.2752	-0.4222	0.0464	-0.1452	-0.1473
CAD/JPY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
EUR/AUD	**0.1861**	-0.0391	-0.0868	-0.1626	0.0139	-0.1244	-0.0974	-0.0197	0.1434	0.0547	0.0926	-0.1400	-0.0596	0.0566	-0.0916
EUR/CAD	**0.1922**	-0.2428	-0.2218	-0.1332	0.0784	-0.0194	0.0621	-0.2208	0.0810	-0.1512	-0.1225	-0.0871	-0.0127	-0.2260	-0.3459
EUR/CSK	0.0336	0.0102	0.0191	0.0455	0.0381	0.0355	0.0264	**0.0643**	0.0139	0.0046	0.0146	0.0548	-0.1382	-0.2327	-0.2812
EUR/GBP	0.1040	-0.0865	-0.0350	0.0218	0.0682	-0.0609	0.0625	-0.1136	**0.1105**	0.0317	0.0792	-0.0595	-0.0275	-0.1347	-0.2398
EUR/JPY	0.0202	-0.0623	-0.0036	-0.0486	0.0112	-0.0197	**0.0218**	0.0007	-0.0287	-0.0156	0.0161	-0.0145	-0.0222	0.0154	0.0135
EUR/NOK	**0.3509**	-0.0428	-0.1281	0.0048	0.1703	-0.1475	-0.0895	0.0955	-0.0300	-0.0691	0.1651	0.0693	-0.0429	-0.1181	-0.2333
EUR/USD	-0.0006	0.0202	-0.0688	-0.2548	-0.0894	-0.1049	-0.1939	-0.1396	-0.0779	**0.0318**	-0.0416	-0.1367	-0.1057	-0.4923	-0.4094
GBP/AUD	0.3542	0.2956	-0.1312	-0.0526	0.1012	-0.2473	0.0719	0.0035	0.0990	**0.4061**	-0.2356	0.0387	-0.1595	-0.3021	-0.0606
GBP/CHF	**0.2022**	-0.1160	-0.0536	-0.1384	-0.0209	-0.0866	-0.1372	-0.1590	0.0514	-0.0283	-0.1698	-0.2131	0.0355	-0.2677	-0.3305
GBP/USD	-0.0590	-0.0478	-0.1415	-0.4172	-0.1226	-0.2035	-0.2336	-0.3485	**-0.0050**	-0.0466	-0.1524	-0.3188	0.0080	-0.0755	-0.3612
NZD/USD	**0.2803**	-0.4779	-0.0115	0.0738	-0.1234	-0.1586	-0.0155	-0.0886	0.1294	-0.1738	-0.1421	-0.2045	0.1238	-0.2339	-0.3662
USD/CAD	**0.0443**	0.0109	-0.3405	-0.3064	-0.0293	-0.2238	-0.2230	-0.2475	-0.0465	0.0224	-0.3631	-0.1372	-0.2991	-0.3056	-0.5711
USD/JPY	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
USD/NOK	**0.4612**	-0.0208	-0.2210	-0.0662	0.1419	-0.4332	-0.1482	0.0011	0.3929	-0.0188	-0.1846	0.0995	-0.1440	-0.0754	-0.1541
USD/SGD	0.0303	0.0272	-0.0516	-0.1478	**0.1108**	-0.0233	0.0229	-0.0028	0.0712	0.1299	-0.0222	0.0348	-0.0574	-0.0436	-0.2949
USD/ZAR	**1.7625**	0.8403	-0.0913	0.6432	0.9516	0.6954	0.3467	0.6417	1.4571	0.4870	0.7265	0.8492	0.0439	0.3436	0.1100
Average	**0.2247**	-0.0242	-0.0851	-0.0575	0.0686	-0.0756	-0.0295	-0.0402	0.1186	0.0260	-0.0277	-0.0281	-0.0378	-0.1118	-0.1879

TABLE 8 Statistical test results of average returns according to the non-parametric Friedman test with the Hommel post-hoc test. 10-minute interval out-of-sample date. Significant differences between the control algorithm (denoted with (c) and the algorithms represented by a row at the α = 5% level are shown in boldface indicating that the adjusted p value is lower than 0.05.

Trading strategies	Average Rank	$Adjust_{pHomm}$
C+Reg-GP (c)	2.8571	-
C+Factor-M	4.7619	0.1675
C+Factor-2	4.8095	0.1675
Factor-2	**6.6190**	0.0192
p+Factor-M	**7.7142**	0.0017
p+Factor-2	**7.8095**	0.0017
RSI	**8.1904**	6.6827E-4
DCC+Factor-2	**8.1904**	6.6827E-4
DCC+Factor-M	**8.5238**	2.9701E-4
Reg-GP	**8.5714**	2.7736E-4
p+Reg-GP	**9.4762**	1.6190E-5
DCC+Reg-GP	**9.6190**	9.6100E-6
Factor-M	**10.0952**	1.8805E-6
EMA	**10.6190**	2.4251E-7
MACD	**12.1429**	2.4059E-10

TABLE 9 %Average maximum drawdown result for trading strategies compared. 10-minute interval out-of-sample data. 20 different currency pairs and 10 calendar months each representing the physical dataset. 5 DC dataset were generated using 5 dynamically generated thresholds tailored to each DC dataset. Best (lowest) value for each row (currency pair) is shown in boldface. Best value among the different variants of the same algorithm (GP, M, O) is underlined.

Dataset	C+Reg-GP	Reg-GP	p+Reg-GP	DCC+Reg-GP	C+Factor-M	Factor-M	p+Factor-M	DCC+Factor-M	C+Factor-2	Factor-2	p+Factor-2	DCC+Factor-2	RSI	EMA	MACD
AUD/NZD	0.1230	0.1506	0.1713	0.3185	0.2699	0.1588	0.2896	0.6500	0.3270	0.2059	0.3368	0.4763	**0.0712**	0.1704	0.1586
AUD/USD	0.1595	0.3123	0.5710	0.6917	0.5750	0.1484	0.6657	0.7631	0.4846	0.2754	0.7167	0.8327	0.1617	**0.1183**	0.1705
EUR/AUD	0.1058	0.1545	0.4086	0.6214	0.3006	**0.0768**	0.4332	0.9292	0.2599	0.2062	0.3246	0.5396	0.1036	0.1199	0.1732
EUR/CAD	0.1353	0.2577	0.4772	0.4494	0.2334	0.2033	0.3325	0.7941	0.2191	0.3079	0.3928	0.3773	**0.0797**	0.1230	0.1484
EUR/CSK	0.0057	0.0080	0.0218	0.0253	0.0133	0.0128	0.0189	0.1117	0.0171	**0.0024**	0.0254	0.0224	0.1863	0.0356	0.0386
EUR/GBP	0.1005	0.0778	0.1460	0.2789	0.1891	0.1375	0.2513	0.4845	0.2120	0.0864	0.2000	0.3884	**0.0716**	0.2415	0.1957
EUR/JPY	0.0106	0.0383	0.0112	0.0255	0.0091	0.0084	0.0157	0.0509	0.0372	0.0519	0.0100	0.0210	0.0144	0.0041	**0.0012**
EUR/NOK	0.1331	0.1476	0.2844	0.3871	0.2699	0.2560	0.4492	0.4974	0.4232	0.1516	0.2737	0.4209	**0.0897**	0.1140	0.1013
EUR/USD	0.1555	**0.0688**	0.2059	0.4034	0.2383	0.1262	0.3368	0.8276	0.2400	0.0828	0.2137	0.3925	0.1905	0.2240	0.2231
GBP/AUD	0.1912	0.2391	0.6403	0.6260	0.4910	**0.2009**	0.6226	0.8899	0.5337	0.1773	0.7777	0.9015	0.2832	0.2077	0.2158
GBP/CHF	0.0956	0.1064	0.1897	0.3278	0.3558	0.0493	0.4552	0.6282	0.2836	0.1908	0.3821	0.5696	**0.0165**	0.1629	0.2002
GBP/USD	0.1323	0.1797	0.2095	0.3867	0.3054	0.1222	0.3887	0.9857	0.2815	0.1617	0.3774	0.5990	**0.1125**	0.1224	0.2697
NZD/USD	0.2892	0.3242	0.4777	0.6830	0.5831	0.2664	0.6115	0.9989	0.4226	0.4036	0.7056	0.7510	**0.0000**	0.2701	0.2627
USD/CAD	0.1678	**0.1615**	0.5727	0.5389	0.2835	0.3403	0.6001	0.6407	0.3781	0.2393	0.5912	0.5521	0.3334	0.3230	0.1673
USD/NOK	0.1406	0.1747	0.7049	0.7367	0.4890	0.5694	0.6361	0.5926	0.3772	0.1716	0.7149	0.6544	0.1732	0.2351	**0.0994**
USD/SGD	0.0770	0.0741	0.1891	0.3058	0.1128	0.0689	0.1463	0.7689	0.1332	**0.0358**	0.2067	0.2644	0.1521	0.1269	0.0792
USD/ZAR	**0.1168**	0.1453	1.2417	1.2160	0.8950	0.2811	1.1893	0.8155	0.6761	0.3312	0.9046	1.0415	0.1494	0.3573	0.7089
Average MDD	**0.1259**	0.1542	0.3837	0.4719	0.3302	0.1780	0.4378	0.6723	0.3121	0.1813	0.4208	0.5179	0.1288	0.1739	0.1890

TABLE 10 Statistical test results of maximum drawdown of DC based trading strategies according to the non-parametric Friedman test with the Hommel post-hoc test. 10-minute interval out-of-sample data. Significant differences between the control algorithm (denoted with (c) and the algorithms represented by a row at the α = 5% level are shown in boldface indicating that the adjusted p value is lower than 0.05.

Trading strategies	Average Rank	$Adjust_{pHomm}$
C+Reg-GP (c)	3.1111	-
Factor-M	4.0556	0.5264
RSI	4.4111	0.5264
Reg-GP	4.6111	0.5264
EMA	5.2778	0.5264
Factor-2	5.3333	0.5264
MACD	4.3889	0.5061
C+Factor-M	**8.1111**	0.0056
C+Factor-2	**8.8889**	8.5002E-4
p+Reg-GP	**9.6111**	1.1688E-4
p+Factor-M	**11.2778**	9.8595E-7
p+Factor-2	**11.0556**	4.7220E-7
DCC+Reg-GP	**12.2222**	1.1813E-8
DCC+Factor-2	**12.7778**	1.1566E-9
DCC+Factor-M	**14.1667**	1.6862E-12

FIGURE 6 Average Sharpe ratio for all currency pairs.

TABLE 11 Statistical test results of risk-adjusted returns according to the non-parametric Friedman test with the Hommel post-hoc test. 10-minute interval out-of-sample date. Significant differences between the control algorithm (denoted with (c) and the algorithms represented by a row at the α = 5% level are shown in boldface indicating that the adjusted p-value is lower than 0.05.

Trading strategies	Average Rank	$Adjust_{pHomm}$
C+Reg-GP (c)	3.0857	-
C+Factor-2	**5.3143**	0.0371
C+Factor-M	**5.4286**	0.0371
p+Factor-M	**6.4857**	0.0044
Factor-2	**6.7714**	0.0023
p+Reg-GP	**7.2857**	4.2690E-4
p+Factor-2	**7.3143**	3.8193E-5
Reg-GP	**8.5143**	2.6705E-6
DCC+Factor-M	**8.6286**	1.5260E-6
DCC+Factor-2	**8.8571**	6.0421E-7
DCC+Reg-GP	**9.3143**	5.6671E-8
RSI	**9.6000**	1.2147E-8
Factor-M	**10.0571**	8.3711E-10
EMA	**10.9000**	3.4833E-12
MACD	**12.4429**	2.9136E-17

5.2.2 | Comparison with Buy-and-hold

In terms of returns and risk, since C+Reg-GP was found to be the best algorithm in comparison to other DC and technical analysis algorithms, we now shift our focus to comparing it to the well-known buy-and-hold (BandH) benchmark, and in the following section, we will look at a sample of best GP models. Under BandH, we buy on the first day of the first month and sell on the last day of the tenth month.

Table 12 compares the mean returns of C+Reg-GP and the BandH strategy. C+Reg-GP outperforms BandH in 12 currency pairs with an overal average return of 0.225% against a negative average return of -0.128% under BandH. In addition, C+Reg-GP's variance is 0.153 and BandH's is 0.515. This indicates that C+Reg-GP is not only more profitable, but also less risky than BandH. These results were also confirmed by a Komolorov-Smirnov statistical test with a p-value of 7.2529e-04.

5.2.3 | C+Reg-GP: distribution of returns

To further understand the performance of C+Reg-GP, we also look into its returns in more detail. Figure 7a presents the distribution of returns across 50 individual GP runs for all 20 currency pairs and the 5 different thresholds. Our goal is here to get insights into the overall performance of the algorithm across all datasets. We have also fitted the distribution of returns with a generalised Student-t distribution, to accommodate for extreme values that the normal

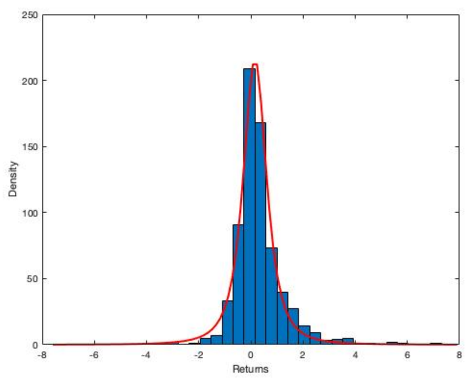
TABLE 12 % Comparison of C+REG-GP and buy-and-hold in terms of average return.

Trading strategies	C+Reg-GP	Buy-and-hold
AUD/JPY	**0.000**	-6.278
AUD/NZD	**0.260**	-0.516
AUD/USD	**0.273**	-5.728
CAD/JPY	**0.000**	-4.109
EUR/AUD	**0.186**	-2.672
EUR/CAD	0.192	**18.555**
EUR/CSK	0.034	**7.770**
EUR/GBP	**0.104**	-0.292
EUR/JPY	**0.020**	-6.211
EUR/NOK	0.351	**2.046**
EUR/USD	-0.001	**8.801**
GBP/AUD	0.354	**3.936**
GBP/CHF	**0.202**	-2.395
GBP/USD	-0.059	**8.464**
NZD/USD	**0.280**	-6.443
USD/CAD	0.044	**2.345**
USD/JPY	**0.000**	-9.430
USD/NOK	**0.461**	-6.102
USD/SGD	0.030	**0.207**
USD/ZAR	**1.762**	-4.505
Mean	**0.225**	-0.128

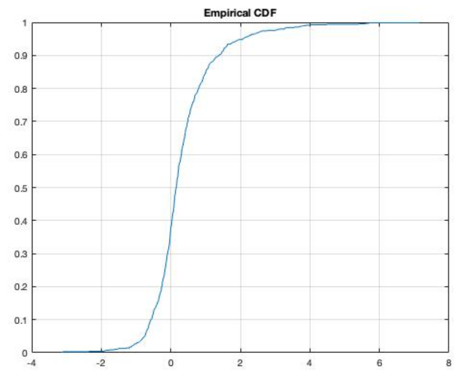
distribution cannot. It should also be noted that we did not plot the cases that the algorithm chose not to trade, as these cases did not yield any return. As we can observe from the figure, the mean of the returns is positive (0.3224); the median is also positive at 0.1606. The returns are positively skewed, with a value of 1.9249. There is also a significantly high kurtosis of 11.7598, which indicates that it produces more outliers than the normal distribution. The range of the returns is [-3.1413, 7.1294], and the standard deviation is 0.9628. In addition, from the distribution data, we have generated the empirical CDF of the returns in Figure 7b, and as we can observe, the algorithm achieves positive returns at a probability of about 65%.

To get more insights into the algorithm's performance for each currency pair, we also present the distribution of returns for each pair in Figure 8. The trading algorithm did not recommend any trading at all across all 50 GP runs for AUD/JPY, CAD/JPY, and USD/JPY, thus these plots are empty. This, in our opinion, is an important result, as C+Reg-GP is able to identify cases that are going to be extremely damaging in terms of profitability, and hence recommend to hold and not take any action at all. It is also very interesting that the algorithm was able to consistently do this for *all 50 GP runs* for the three aforementioned currency pairs. For the remaining datasets, C+Reg-GP is by and large yielding positive returns. We are presenting the mean, standard deviation, skewness and kurtosis for each currency pair in

FIGURE 7 (7a) Distribution of returns for the C+Reg-GP algorithm across 50 GP runs, 20 currency pairs, and 5 different DC thresholds. The distribution has been fit with a generalised Student-t distribution. (7b) Empirical CDF for the returns of C+Reg-GP algorithm across 50 GP runs, 20 currency pairs, and 5 different DC thresholds.



(a)



(b)

Table 13. With the exception of EUR/USD and GBP/USD, all other pairs are experiencing positive mean returns. Also, returns are mainly positively skewed, with the exception of EUR/GBP, EUR/USD, and USD/NOK. Lastly, the kurtosis for the majority of currency pairs is around 3, indicating that the tails of the distribution experiences extreme outliers on par with the normal distribution. Overall, we consider these as quite positive results, as they demonstrate that C+Reg-GP has a high potential of yielding positive returns across a variety of datasets.

TABLE 13 Summary statistics for all currency pair results for the C+Reg-GP algorithm. The trading algorithm did not recommend any trading for AUD/JPY, CAD/JPY, and USD/JPY, thus for these currency pairs a NaN value is shown in their respective statistics.

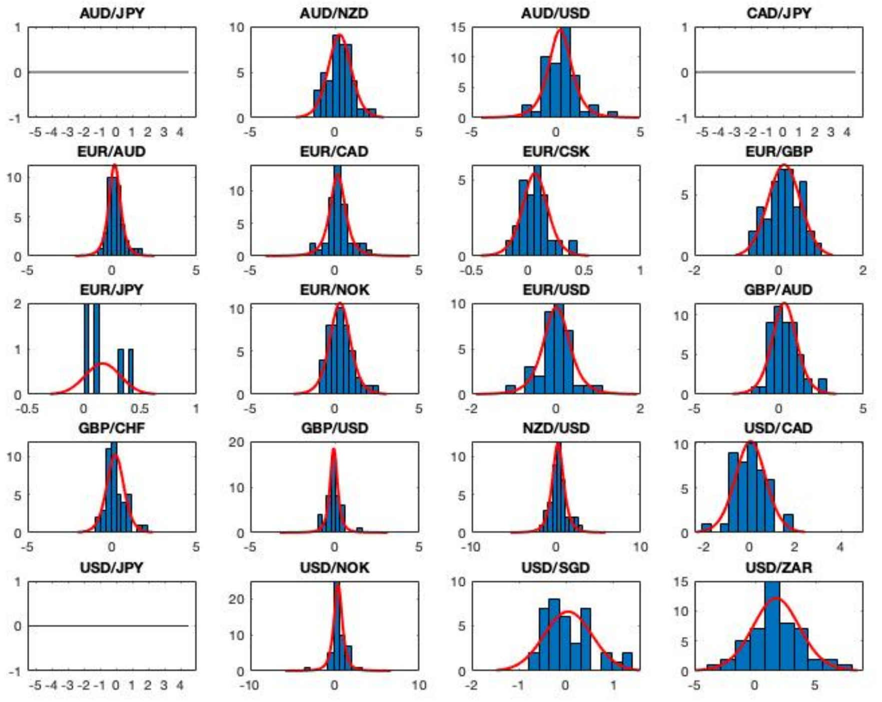
Currency pair	AUD/JPY	AUD/NZD	AUD/USD	CAD/JPY	EUR/AUD	EUR/CAD	EUR/CSK	EUR/GBP	EUR/JPY	EUR/NOK
Mean	NaN	0.3	0.29	NaN	0.23	0.23	0.07	0.13	0.17	0.37
Stand. Dev.	NaN	0.76	0.97	NaN	0.49	0.7	0.14	0.39	0.18	0.73
Skewness	NaN	0.25	0.75	NaN	1.14	0.36	0.6	-0.02	0.6	0.86
Kurtosis	NaN	3.33	4.93	NaN	4.71	3.7	3.67	2.28	1.88	3.84

Currency pair	EUR/USD	GBP/AUD	GBP/CHF	GBP/USD	NZD/USD	USD/CAD	USD/JPY	USD/NOK	USD/SGD	USD/ZAR
Mean	-0.0009	0.37	0.23	-0.06	0.33	0.06	NaN	0.47	0.04	1.77
Stand. Dev.	0.42	0.85	0.61	0.42	1.0	0.73	NaN	0.92	0.52	2.08
Skewness	-0.21	0.62	0.75	0.93	0.71	0.17	NaN	-0.45	0.79	0.17
Kurtosis	4.09	3.73	3.75	7.11	3.68	3.34	NaN	7.7	2.96	3.17

5.2.4 | A sample of best GP models

For completeness, we present a sample of the best trees (in terms of profitability) that C+Reg-GP evolved in their equation format. OS_l is the OS length and DC_l is DC length.

FIGURE 8 Distribution of returns for the C+Reg-GP algorithm. The returns are presented separately for each currency pair, and are over 50 GP runs and 5 DC thresholds. All distributions have been fit with a generalised Student-t distribution. The trading algorithm did not recommend any trading at all for AUD/JPY, CAD/JPY, and USD/JPY, thus their respective plots are empty.



$$OS_l = \log(a + DC_l^b) \tag{8}$$
where a= 1609.55 and b = 5.023.

$$OS_l = \log((DC_l \times a)^b) \tag{9}$$
where a= 4.117 and b = 5.764.

$$OS_l = \cos(a \times \cos(DC_l)) + \frac{b}{exp(\cos(DC_l))} \tag{10}$$
where a = 292.160 and b= 4.569

$$OS_l = \exp(\exp(\sin(\sin(DC_l)))) + (a \times (b + \log(DC_l))) \quad \text{where a = 1.750 and b = 1.957.} \tag{11}$$

Most equations have different structures. The first and second equations are using the logarithmic function, the third consists of the cosine and the exponential function, and the fourth equation has three components, namely the exponential, sine, and logarithm. What we can conclude from these equations for the DC and OS length relationship is that (i) it is non-linear, and (ii) it is dataset dependent. The latter is an important observation, as it indicates that we need to be evolving tailored equations for different datasets to understand the DC-OS relationship, and thus predict trend reversal.

TABLE 14 Average computational times per run for C+Reg-GP, Reg-GP , *p*+Reg-GP, DCC+Reg-GP, C+Factor-M, Factor-M, *p*+Factor-M, DCC+Factor-M, O+Reg-GP, Factor-2, *p*+Factor-2, DCC+Factor-2, RSI, EMA, MACD. BH takes less than 1 second to execute because we buy quoted currency at the start of trading period and sell quoted currency at the end of trading period.

Trading strategies	**C+Reg-GP**	**Reg-GP**	***p*+Reg-GP**	**DCC+Reg-GP**
Classification	~ 65 mins	–	–	–
Estimation	~ 5.45 mins	~ 6.20 mins	~ 5.25 mins	–
Trading	~ 3 sec	~ 3 sec	~ 3 sec	~ 3 sec
Trading strategies	**C+Factor-M**	**Factor-M**	***p*+Factor-M**	**DCC+Factor-M**
Classification	~ 65 mins	–	–	–
Estimation	~ 30 secs	~ 30 secs	~ 30 secs	–
Trading	~ 3 sec	~ 3 sec	~ 3 sec	~ 3 sec
Trading strategies	**C+Factor-2**	**Factor-2**	***p*+Factor-2**	**DCC+Factor-2**
Classification	~ 65 mins	–	–	–
Estimation	~ 20 secs	~ 20 secs	~ 20 secs	–
Trading	~ 3 sec	~ 3 sec	~ 3 sec	~ 3 sec
Trading strategies	**EMA**	**MACD**	**RSI**	**–**
Classification	–	–	–	–
Estimation	–	–	–	–
Trading	~ 3 sec	~ 3 sec	~ 3 sec	~ 3 sec

5.2.5 | Computational times

Table 14 presents the average computational times for all algorithms. We can observe that different algorithms can have significantly different computational times, which is not surprising. An algorithm such as C+Reg-GP includes the classification step, which consisted of Auto-WEKA running for 60 minutes in order to find the best classification model per dataset, and also optimise its hyperparameters;[2] it also includes a GP, which requires some time to evolve a good solution, since multiple individuals and generations are involved.

To make things clearer, we are presenting the computational times for each task in our framework: classification, (OS length) estimation, and trading. Not all algorithms use the classification step, but the ones that do use it need about 65 minutes to complete this task. The estimation task is 5-6 minutes for algorithms that use a GP, and 20-30 seconds for the other algorithms. With regards to the trading step, all algorithms need around 3 seconds.

It should be noted that learning on the training set usually happens off-line, thus a 65-70 minute duration does not consist a problem. Once training is complete, the best model is applied in real time to the (unseen) test set, which only takes 3 seconds. We believe that the significant improvements we have observed in returns and risk justify the slower execution time. Lastly, the overhead of including a classification step can be reduced by parallelising the AutoWeka process. It has actually been shown in the literature (e.g, [13]) that parallelisation can reduce computational times significantly.

5.3 | Summary

Our findings can be summarised as follows:

Adding a classification step to a DC algorithm has a positive effect in predicting the trend reversal under directional changes. As we observed in Table 6, the very positive classification results have led to significantly reduced RMSE, ranking each algorithm that uses a classifier higher than its respective variant without classification. In addition, C+GP ranked first and statistically outperformed all other DC-based trend reversal algorithms.

Introducing a classification step to a DC algorithm leads to higher returns during trading. As we observed in Tables 7 and 8, all algorithms that used a classifier (C+Reg-GP, C+Factor-M, C+Factor-2) outperformed other variants without a classifier. Furthermore, C+Reg-GP ranked first among 15 trading algorithms and statistically outperformed 12 algorithms, with the only exception of the two other algorithms that were using a classifier.

Introducing a classification step to a DC algorithm leads to less risky strategies. As we saw in the Sharpe ratio results, all the variants with the classifier ranked in the first 3 places (Table 11). This could perhaps be attributed to the fact that the Sharpe ratio is a metric that includes both returns and risk. On the other hand, the MDD results presented a mixed picture, with C+Reg-GP ranking first across all algorithms, but Factor-M and Factor-2 ranking higher than their variants with a classifier.

There is no generalised formula for predicting trend reversal under directional changes. Each dataset has its own characteristics and requires tailored trend reversal equations.

C+Reg-GP is an effective trading algorithm. It not only outperformed other DC-based algorithms but also 3 different technical indicators, as well as buy-and-hold. This was the case not only for average returns but also for MDD and Sharpe ratio (risk metrics).

6 | CONCLUSION

To conclude, this paper presented an extensive investigation over of total 1,000 datasets from 20 different Forex currency pairs to demonstrate that the introduction of a classification step in DC-based price summaries, where we

[2] The time taken in the classification phase of C+Reg-GP, C+Factor-M, and C+Factor-2, went above the allotted time of 60 minutes due to CPU time slicing, as other processes were running on the hardware simultaneously. With the availability of dedicated hardware with sufficient CPU cores, a large speed-up could be obtained by switching the classification phase from serial mode to parallel mode and also reducing the execution time. For example, using 60 core hardware and reducing executing time to around 2 minute.

predict whether a DC event is going to be followed by an OS event, leads to improved returns and risk results in trading. We compared our results across three DC algorithms, where we run experiments for two versions per algorithm, one with a classifier, and one without a classifier. We also benchmarked our results to technical analysis and buy and hold. Our results confirmed that the use of classification leads to improved trend reversal prediction and thus profitable and low-risk trading strategies. We also found that one of the DC algorithms (C+GP+TS) consistently outperforms all other algorithms in all metrics that took place in our investigation.

As future work, we would like to create multi-threshold DC trading strategies. In those strategies, all thresholds will contribute towards the decision-making process of taking trading actions. This would have the advantage of combining the 'knowledge' of multiple thresholds, which could outperform the performance of individual thresholds' trading strategies. Additional research could take place in creating tailored classification algorithms, rather than using Auto-Weka.

references

[1] Adegboye, A. and Kampouridis, M. (2021). Machine learning classification and regression models for predicting directional changes trend reversal in FX markets. *Expert Systems with Applications*. In Press.

[2] Adegboye, A., Kampouridis, M., and Johnson, C. G. (2017). Regression genetic programming for estimating trend end in foreign exchange market. In *Computational Intelligence (SSCI), 2017 IEEE Symposium Series on*, pages 1–8. IEEE.

[3] Alkhamees, N. and Fasli, M. (2017). A directional change based trading strategy with dynamic thresholds. In *2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, pages 283–292. IEEE.

[4] Almeida, R. J., Baştürk, N., and Golan, R. (2017). Intraday value-at-risk estimation for directional change events and investment strategies. In *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–8. IEEE.

[5] Aloud, M. (2016a). Directional-change event trading strategy: Profit-maximizing learning strategy. *The Seventh International Conference on Advanced Cognitive Technology and Applications*.

[6] Aloud, M., Tsang, E., and Olsen, R. (2014). Modeling the fx market traders' behavior: An agent-based approach. *Banking, Finance, and Accounting: Concepts, Methodologies, Tools, and Applications: Concepts, Methodologies, Tools, and Applications*, page 350.

[7] Aloud, M. E. (2016b). Time series analysis indicators under directional changes: The case of saudi stock market. *International Journal of Economics and Financial Issues*, 6(1).

[8] Ao, H. (2018). *A Directional Changes based study on stock market*. PhD thesis, University of Essex.

[9] Azzini, A., da Costa Pereira, C., and Tettamanzi, A. G. (2010). Modeling turning points in financial markets with soft computing techniques. In *Natural Computing in Computational Finance*, pages 147–167. Springer.

[10] Bakhach, A., Chinthalapati, V., Tsang, E., and El Sayed, A. (2018). Intelligent dynamic backlash agent: A trading strategy based on the directional change framework. *Algorithms*, 11(11):171.

[11] Bakhach, A., Tsang, E., Ng, W. L., and Chinthalapati, V. R. (2016). Backlash agent: A trading strategy based on directional change. In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–9. IEEE.

[12] Brabazon, A., Kampouridis, M., and O'Neill, M. (2020). Applications of genetic programming to finance and economics: past, present, future. *Genetic Programming and Evolvable Machines.*, 21:33–53.

[13] Brookhouse, J., Otero, F. E., and Kampouridis, M. (2014). Working with opencl to speed up a genetic programming financial forecasting algorithm: initial results. In *Proceedings of the Companion Publication of the 2014 Annual Conference on Genetic and Evolutionary Computation*, pages 1117–1124.

[14] Cavalcante, R. C., Brasileiro, R. C., Souza, V. L., Nobrega, J. P., and Oliveira, A. L. (2016). Computational intelligence and financial markets: A survey and future directions. *Expert Systems with Applications*, 55:194–211.

[15] Chen, J. and Tsang, E. (2018). Classification of normal and abnormal regimes in financial markets. *Algorithms*, 11(12):202.

[16] Chen, T.-l. and Chen, F.-y. (2016). An intelligent pattern recognition model for supporting investment decisions in stock market. *Information Sciences*, 346:261–274.

[17] Chung, F.-L., Fu, T. C., Luk, R., and Ng, V. (2001). Flexible time series pattern matching based on perceptually important points. *International Joint Conference on Artificial Intelligence Workshop on Learning from Temporal and Spatial Data*, pages 1–7.

[18] Glattfelder, J., Dupuis, A., and Olsen, R. (2008). An extensive set of scaling laws and the fx coastline. *Centre for Computational Finance and Economics Agents, Working Paper Series WP025-08*.

[19] Glattfelder, J., Dupuis, A., and Olsen, R. (2011). Patterns in high-frequency fx data: discovery of 12 empirical scaling laws. *Quantitative Finance*, 11(4):599–614.

[20] Guillaume, D. M., Dacorogna, M. M., Davé, R. R., Müller, U. A., Olsen, R. B., and Pictet, O. V. (1997). From the bird's eye to the microscope: A survey of new stylized facts of the intra-daily foreign exchange markets. *Finance and stochastics*, 1(2):95–129.

[21] Gypteau, J., Otero, F. E., and Kampouridis, M. (2015). Generating directional change based trading strategies with genetic programming. In *European Conference on the Applications of Evolutionary Computation*, pages 267–278. Springer.

[22] Kablan, A. and Ng, W. L. (2011). Intraday high-frequency fx trading with adaptive neuro-fuzzy inference systems. *International Journal of Financial Markets and Derivatives*, 2(1-2):68–87.

[23] Kampouridis, M., Adegboye, A., and Johnson, C. (2017). Evolving directional changes trading strategies with a new event-based indicator. In *Asia-Pacific Conference on Simulated Evolution and Learning*, pages 727–738. Springer.

[24] Kampouridis, M. and Otero, F. E. (2017). Evolving trading strategies using directional changes. *Expert Systems with Applications*, 73:145–160.

[25] López-Ibánez, M., Dubois-Lacoste, J., Stützle, T., and Birattari, M. (2011). The irace package, iterated race for automatic algorithm configuration. Technical report, Citeseer.

[26] Mandelbrot, B. and Taylor, H. M. (1967). On the distribution of stock price differences. *Operations research*, 15(6):1057–1062.

[27] Özorhan, M. O., Toroslu, İ. H., and Şehitoğlu, O. T. (2018). Short-term trend prediction in financial time series data. *Knowledge and Information Systems*, pages 1–33.

[28] Poli, R., Langdon, W. B., McPhee, N. F., and Koza, J. R. (2008). *A field guide to genetic programming*. Lulu. com.

[29] Sewell, M. (2011). Characterization of financial time series. In *Research Note RN/11/01*. UCL Department of Computer Science- London.

[30] Thornton, C., Hutter, F., Hoos, H. H., and Leyton-Brown, K. (2013). Auto-weka: Combined selection and hyperparameter optimization of classification algorithms. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 847–855. ACM.

[31] Tsang, E. and Chen, J. (2018). Regime change detection using directional change indicators in the foreign exchange market to chart brexit. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2(3):185–193.

[32] Tsang, E. P., Tao, R., Serguieva, A., and Ma, S. (2017). Profiling high-frequency equity price movements in directional changes. *Quantitative finance*, 17(2):217–225.

[33] Wan, Y., Gong, X., and Si, Y.-W. (2016). Effect of segmentation on financial time series pattern matching. *Applied Soft Computing*, 38:346–359.

[34] Yin, J., Si, Y.-W., and Gong, Z. (2011). Financial time series segmentation based on turning points. In *Proceedings 2011 International Conference on System Science and Engineering*, pages 394–399. IEEE.