Optimization of Trading Strategies Using a Genetic Algorithm under the Directional Changes Paradigm with Multiple Thresholds

Ozgur Salman, Themistoklis Melissourgos, Michael Kampouridis
School of Computer Science and Electronic Engineering
University of Essex
Wivenhoe Park, United Kingdom
{os20224, themistoklis.melissourgos, mkampo}@essex.ac.uk

***Abstract*—This paper explores the use of the Directional Changes (DC) paradigm for financial forecasting. DC is an event-based alternative to the traditional approach of time-series with fixed intervals. In the DC approach, price movements are recorded when specific events occur, rather than in fixed time intervals, while significant price changes are identified using a threshold. Here, we consider a more general model that allows multiple weighted thresholds, and propose three novel trading strategies built within the DC paradigm. To optimize the weights of the thresholds, we use a genetic algorithm and manage to find strategies that outperform previously known single-threshold strategies under the common efficiency metrics. Furthermore, our method manages to create profitable trading strategies that outperform some traditional ones, such as buy-and-hold, MACD, and RSI.**

***Index Terms*—directional changes, multiple thresholds, trading strategies, genetic algorithm, financial forecasting**

I. Introduction

The field of financial forecasting has advanced tremendously over the last few decades with a focus on the concepts of return and risk. The introduction of modern portfolio theory by the seminal work of Markowitz, [1], ignited a line of research that revolved around making profitable portfolios for investors while at the same time managing their risk. Although traditional forecasting models have been improved through the addition of new technical tools, the use of time-series has largely remained unchanged. However, relying solely on interval-led data can be risky as unexpected news events or price movements outside of the interval can be missed or detected too late, potentially resulting in losses by the selection of hourly, daily, or weekly intervals.

In order to prevent neglecting price and market movements, instead of using time-series data to sample price values in regular time intervals (e.g. daily closing prices, hourly data), we use an event-based system to capture significant points in price movements, namely, the (DC) paradigm. This uses a sampling method to take snapshots of historical data whenever there is a change in price that exceeds a specific threshold. The characteristic parameter of the DC paradigm is a threshold $\theta > 0$, whose value is determined by the trader based on their belief of what constitutes a significant price change. Note that this change can be either an increase or a decrease in value. The price data is then divided into uptrend and downtrend intervals, where each interval consisting of a directional change, called *directional change (DC) event* is usually followed by an *overshoot (OS) event*.

In finance, a *trading strategy* is a term for creating a plan for buying, selling, or holding assets, such as stocks, bonds, material, or intellectual property, with the goal of making a profit. In this work our aim is, at first, to use various indicators and scaling laws from the DC paradigm to create three novel trading strategies. The next step is to feed the strategies with a range of different thresholds to see what recommendations they produce. The goal of this step is to identify multiple concurrent events by different thresholds over a profiling period and provide more opportunities to trade profitably. Intuitively, such a method should yield at least as good trading strategies as those of the single-threshold model, since the latter is a special case of our multi-threshold model.

In the presented experiments, we test ten predefined threshold values for two trading strategies, while the remaining strategy is tested with five of those thresholds. As a result, there are multiple buy-sell-hold recommendations at each time-step. To resolve any conflicting recommendations, a genetic algorithm (GA) is used to optimize the weights of each DC threshold. Genetic algorithms are evolutionary algorithms that are commonly used for stochastic optimization [2] and have been successfully employed to address financial forecasting problems [3], [4].

We conduct our experiments on 18 New York Stock Exchange (NYSE) listed stocks. Our research compares the multi-threshold DC strategy to: (i) a collection of single-threshold strategies, (ii) the two technical analysis indicators in a time-series context, and (iii) a buy-and-hold strategy. Section II summarises the necessary DC background and Section III covers the related literature. Section IV and Section V present the methodology and experimental setup, respectively. Finally, we provide the results in Section VI and the conclusions from our experiments in Section VII.

979-8-3503-1458-8/23/31.00 ©2023 Crown

II. DIRECTIONAL CHANGES

In this section, we give an overview of our model, namely the DC paradigm. Particularly, in Section II-A, we outline the fundamental principles of DC. Section II-B describes the background on the indicators that are relevant to our proposed methodology.

A. DC Basics

The DC approach is a way of recording price changes that are caused by events. It involves only two types of events, namely, OS events and DC events. A DC event is confirmed when there is a price change that exceeds the pre-specified threshold in the opposite direction of the current trend. The time interval between two consecutive DC events is an OS event, which usually is of non-zero length. The threshold θ used in this model is specified by the user according to the application (asset) at hand. Fig. 1 shows an example of the formation of consecutive DC and OS events at a threshold $\theta = 5\%$. Any point A on the graph is a pair of time-step T_A and price P_A. Suppose we have a financial product whose price starts at 100$ at time-step 0 and decreases to 96$ at time-step T_{EXT_i}. Since the price change is smaller than the pre-specified value of θ, we do not consider the time interval $0-T_{EXT_i}$ as a DC event. However, from T_{EXT_i} to time-step T_{DCC_i}, the price experiences a significant change of 5%, so indeed, the time interval T_{EXT_i}-T_{DCC_i} is a DC event.

As a result, with relation to T_{DCC_i}, we need to define two important notions of points: the *extreme point* EXT_i and the directional change *confirmation point* DCC_i. Throughout the paper, without loss of generality, we will consider discrete time, in the form of time-steps $0, 1, 2, \dots$ which models the particular points in time where we “sense” the price of the underlying asset (e.g., here it is the daily closing price of a stock). The extreme point is the leftmost limit of the DC interval, while the confirmation point is the earliest time-step in which we have a DC. The definitions of these points will be useful in our strategies' description in Section IV-A. To detect another DC event, the threshold must be reached in the opposite direction of the previous DC event's direction, as shown at point DCC_{i+1} in Fig. 1. The price change of a DC event can be longer than the minimum required price change (determined by θ) that would qualify it as a DC event. To account for this, the concept of a *theoretical confirmation point*, $DCC*$, is introduced. This can be seen in Fig. 1, where a price change of 4.8$ (recall that $\theta = 5\%$ in our example) between points EXT_i and $DCC*$ is sufficient to confirm a DC event. T_{DCC*} indicates that the time-step $\lceil T_{DCC*} \rceil = T_{DCC_i}$ (time-step 4) is the upper limit of the DC event, in other words, the time-step of confirmation point DCC_i. In the DC paradigm, the individual event's direction is taken into account, and note that there are only two directions: by *uptrend (UT)* (resp. *downtrend (DT)*) we refer to the series of points between a low and a high (resp. a high and a low) extreme point. In Fig. 1 the interval EXT_i-EXT_{i+1} is an uptrend.

By presenting the data in this way, the DC paradigm offers users a new perspective on observing price changes.

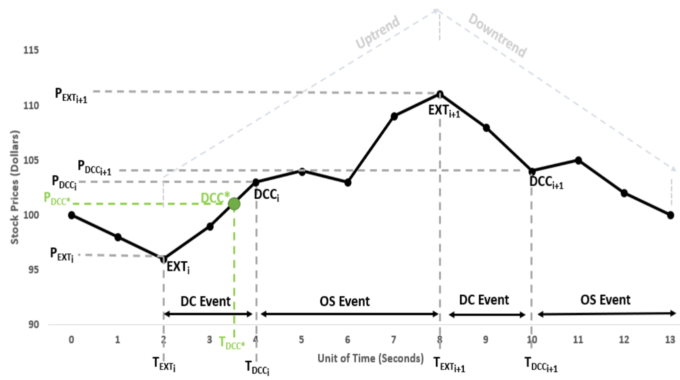

Fig. 1. Transformation of time-series data into the DC paradigm. There are two DC event confirmation points, at time 4 and 10. An uptrend takes place between the two extreme points, EXT_i and EXT_{i+1}, which are confirmed retrospectively at their subsequent confirmation points, DCC_i and DCC_{i+1}.

A significant benefit of this approach is that users can focus on key points rather than potentially missing important events due to their preferred time intervals in the physical time-series. In the rest of this section, we discuss auxiliary results from the DC literature, which we will also use in our model.

B. Indicators

Here, we focus on the model's parameters, called *indicators*, most of which have been defined in prior works. The use of indicators in the literature of technical analysis serves the purpose of uncovering hidden patterns in financial data, which can be utilized by decision-making tools such as DC in order to make trading even more profitable. We now give a thorough description of the indicators that we applied to our strategies' creation.

- Number of DC events (N_{DC}): The total number of DC events throughout the investigated period.
- Number of Overshoot Events (N_{OS}): The total number of OS events in the profiled data.
- Theoretical Confirmation Point ($DCC*$): The earliest time at which a price change equals θ.
  At the uptrend:
$$P_{DCC*} = P_{EXT_i} \cdot (1 + \theta) \tag{1}$$
- Overshoot Values at Current Points (OSV_{CUR}): The main goal of this indicator is to measure the magnitude of an OS event. It can be calculated as follows:
$$OSV_{CUR} = \frac{P_{CUR} - P_{DCC*}}{\theta \cdot P_{DCC*}}, \tag{2}$$
  where P_{CUR} is the current price of the asset.

III. RELATED WORK

We can explain the two main sources of literature on which this research is based as follows: First, the discoveries made in evolutionary algorithms trading strategy optimization; and second, the findings about DC and their usage on trading strategies via evolutionary algorithms.

A. *Evolutionary Algorithms and Trading*

Evolutionary algorithms (EA) have emerged as a popular tool for solving complex financial optimization problems by mimicking the process of natural selection to identify optimal solutions in large search spaces. In [5], although the authors comprehensively studied 51 journal articles, they were unable to make a definitive conclusion about which EA performs better in different research areas of finance. In the meantime, optimization of trading strategies based on GA has gained attention [6], [7]. However, both the mentioned articles and the attention in the literature generally continued to use physical time, other than [8], where authors tried to optimize different strategies under the DC paradigm. It should be noted here that profitable results have been achieved not only through GA but also through genetic programming under the umbrella of DC [9], or even without directional change [10], as highlighted in recent literature. In this paper, we propose using GA as an optimizer to experiment with different trading recommendations based on various θ values derived from DC. The aim is to investigate whether the use of multiple threshold values can improve the quality of trading decisions by providing more diverse and informative data profiles.

B. *Scaling Laws & Indicators*

Scaling laws, conceptually, describe the functional relationship between two physical quantities that scale with each other over a significant interval. In DC, these relationships focus on to establish mathematical connections among price moves, duration and frequency. Here, we will cover the literature regarding to these laws in DC with the addition of indicators that are derived from paradigm itself.

At scaling laws, early findings provided a deeper understanding of the behavior of foreign exchange markets, among 13 pairs, 17 scaling laws were demonstrated to the research community [11]. With newly added 12 more, the DC helped to understand the patterns by profiling time-series into event-based system [12]. Among them, one which is related to our work. Authors showcased an empirical consistency between numerous foreign currency data that the duration of an OS event, over the profiled data, is approximately twice the duration DC event on average. Following the notation of [12], let us denote by $\langle T_{OS} \rangle$ and $\langle T_{DC} \rangle$ the average time of an OS and DC event, respectively. Then the aforementioned scaling law can be written as:

$$\langle T_{OS} \rangle \approx 2 \cdot \langle T_{DC} \rangle, \tag{3}$$

where "$\approx$" indicates empirical equality. This is closely tied to the mathematical relationships between DC and OS, which will be used in Section IV-A in creating a trading strategy. After the initial discoveries, 4 and 1 more were added, respectively [13], [14], where the financial product differed from the previous researches and the usage of DC under equity products added to the field. These findings are already being used to develop trading strategies in the field [8], and it is likely that there will be further progress in this area.

DC has also been enhanced by the addition of indicators, which helps new users to gain a better understanding of the paradigm and use them as technical-analysis-like tools in the field. The study in [13] was one of the first to explore the use of four indicators. Conceptually, [15] can be thought of as a dictionary of DC-based indicators, providing information on how to extract pattern-based data from the paradigm itself. In this paper, we adapted our indicators from the original versions, and they contribute to the field as new indicators as well. Their construction and implementation are discussed comprehensively in Section II-B and IV.

C. *Trading Strategies*

An important aspect of DC-based trading strategies is the use of classification tasks, which have been shown to outperform technical analysis techniques [16]. In that work, Forex instrument pairs were investigated, and the use of classification tasks provided a new option for users to observe trend reversals under DC. Very recently, researches showed that with 20 different Forex pairs under DC, the proposed algorithm by DC trend reversion projection were able to outperform majority of the DC and non-DC benchmarks (i.e., exponential moving average) in terms of both return and risk [17]. Overall, these works show that DC framework is flexible and open for improvements in trading strategies.

Based on our review of this literature, a first observation is that DC has been used to a limited extent in the creation of trading strategies. To increase the application of indicators and scaling laws in trading strategies, our first motivation is to create strategies for traders to act on. As another limitation, the trading strategies that were tested have only used a single threshold for determining buy, hold, or sell recommendations most of the time, while we aim to implement multi-thresholds for a richer recommendation through the addition of different thresholds. With the addition of GA optimization, we target to improve these recommendations and provide a more comprehensive decision-making tool for traders.

IV. METHODOLOGY

We implement the scaling laws and indicators derived by DC, by constructing simple strategies, since this seems sufficient in improving the effectiveness of the current state-of-the-art DC strategies.

In what follows, we first introduce the strategies in Section IV-A, then we discuss the GA methodology and how it was used to optimize the aforementioned trading strategies in Section IV-B.

A. *Individual Strategy Creation Process*

Table I summarizes the three individual DC-based trading strategies discussed in the paper. These strategies all follow four key rules: (i) a new position (i.e., executing a buy, or sell on a stock) cannot be opened if a position is already open; therefore, a position must be closed before a new one can be opened, (ii) short selling is not permitted, so all opening positions must involve going long on a financial product, (iii)

TABLE I
INDIVIDUAL STRATEGIES

Strategy	Buying	Selling
St1	Twice the duration of DC from P_{DCC} in DT	Twice the duration of DC from P_{DCC} in UT
St2	$\mid OSV_{CUR} \mid \geq \mid OSVBest \mid$ in DT	$\mid OSV_{CUR} \mid \geq \mid OSVBest \mid$ in UT
St3	3rd consecutive OS in UT	P_{DCC} in DT

a transaction cost of 0.25% is applied to each trade, (iv) execution of buying orders should be in downtrend (DT), selling in uptrend (UT), for St1 and St2 only. The first trading strategy is based on the scaling law presented in Equation (3). The two remaining strategies are based on indicators. It's worth noting that the definition of OSV_{CUR} in this paper is a modified version of the original indicator from [15]. Instead of using extreme values, this indicator uses current values. This allows us to use these indicators at every data point, rather than only during extreme points. St3 is again based on the indicators defined in Section II-B. The following paragraphs describe the strategies rigorously.

Strategy St1 uses the first scaling law, Equation (3), which dictates that the duration of an OS event is approximately twice the DC event. In this paper, we aim to test the scaling law as performing it to each trend rather than using on the whole profiled data. As an execution signal, we check the time duration that lasts for DC, at the moment we confirm the DC at P_{DCC} at DT, and then wait for the double of that duration to buy the stock. In order to sell the stock, we wait for the P_{DCC} at UT, and then act on its double duration again. The motivation behind this strategy is to make an informed decision based on the scaling law would hold on each trend.

Strategy St2 is based on Overshoot Values at Current Points indicator. In the creation, we check if $|OSV_{CUR}| \geq |OSVBest|$, where $|OSV_{CUR}|$ is the absolute value of indicator where we defined at Equation (2). In order to estimate the $|OSVBest|$, we do the following: first we obtain the distribution of all OSV_{CUR} values for DC profiled data; then we divide these values into quartile and then select one median OSV_{CUR} for each quartile (i.e., we test four indicator value). In the end, we identify best OSV_{CUR} value with the highest Sharpe ratio (in a validation set) among four indicator value, and attain the best one as $OSVBest$. Finally, whenever we see the $|OSV_{CUR}| \geq |OSVBest|$ rule holds, we check the trend direction as signal, if it is DT we buy the stock, and wait for the opposite to hold to sell the stock.

In Strategy 3 (St3), we check the consecutive OS events in UT, for example, in order to buy a position, we need to see subsequent three OS events without seeing any OS in DT prior to 3rd OS. Then, wait for the next DT confirmation point to close the position. The motivation behind St3 is to make a profit, by capturing a resistance of uptrend.

B. Genetic Algorithm

GA starts by creating a population of potential solutions, called candidates, or chromosomes. These solutions are evaluated using a fitness function to determine their *quality* (i.e., the Sharpe Ratio, which will be covered in Section IV-B2). At each iteration, a new population is generated by selecting the more fit individuals from the current population probabilistically. Some of these selected candidates will undergo crossover and mutation, which are processes that introduce changes to the chromosome in order to explore different areas of the search space. The rest of the selected chromosomes are carried forward without any changes. In our research, we assigned weight to each candidate's gene, and in terms of the genes, they represent *"what would be the performance under this particular threshold by DC"*. Eventually, the Sharpe Ratio in the resulting model depends on these weights. It is important to note here that in the upcoming sections, we will dive into the models representation and optimization phases more. The new population then replaces the old one, and the process repeats until we reach the pre-specified number of generations.

In one of the recent works [8], instead of thresholds researchers use different DC-based strategies as genes in chromosome. However, they implemented these strategies under one threshold $\theta = 2.5\%$, which limits research to analyze the events by only one DC profiled data. Therefore, we aim to answer the question of how the GA can improve its performance by utilizing multiple thresholds. In the remainder of this section, we will show how we implement the GA, which eventually constructs a recommendation on buy-hold-sell any unit of time.

It is clear that any chosen value of θ can only generate a single set of DC and OS events. For instance, while smaller thresholds result in more frequent events and the opportunity to take prompt actions, larger thresholds detect fewer events but allow for the possibility of taking action in response to larger price changes. Thus, in this work, we aim to capture the spectrum of events by optimization of multiple thresholds. Each one of the following values of θ creates its own set of buy-hold-sell strategies. The chosen thresholds are as follows: 0.098%, 0.22%, 0.48%, 0.72%, 0.98%, 1.22%, 1.55%, 1.70%, 2%, 2.55%. The reason we choose these thresholds is changes in the price of the stock that fall approximately within 0.1 to 2.5 percent range would encapsulate the events that we see important. Furthermore, this range also encompasses a wide spectrum of price fluctuations, covering both smaller and larger changes. It is important to note here, due to application of St3 is heavily depending on rapid price changes, we used only first five thresholds for it.

Consequently, it is possible that at a given time, one threshold value may recommend a "buy" action while another recommends a "sell", or "hold" action. One benefit of having multiple θ's is that it can provide a greater depth of information and more recommendations for each data point. However, if a trader wants to consider the recommendations from multiple θ's, they may encounter conflicting actions, such as one strategy suggesting to buy a position while another suggesting to sell it. To address these conflicting recommendations, we assign a weight to each θ and adjust these weights over time. When it comes to making a decision, we can then follow the recommendation that has the highest overall weight, based on the sum of the weights of all the θ's. The following details describe the process using a GA.

1) Representation of Individuals

An individual in the GA consists of 10 genes, with each gene representing a threshold. The value of each gene is a weight ranging from 0 to 1. An example of an individual is shown in Table II. There are total of 10 thresholds. The first row in the example individual is a label for reference. The remaining rows show that each of the 10 θ has been assigned a weight. At each point in time, each θ will make a recommendation on whether to buy a position, hold the current position, or close a position. This means that different recommendations may be made at any given time based on the weights and recommendations of the individual strategies. Let us assume that $\theta1$, $\theta2$, $\theta3$, $\theta4$, $\theta5$ recommend a buy at that time of the data, $\theta6$, $\theta7$ hold, and $\theta8$, $\theta9$, $\theta10$ to sell the position. We then sum up the weights for each recommendation, i.e. the sum of buying a position is $0.10 + 0.20 + 0.05 + 0.15 + 0.10 = 0.60$ ($\theta1$, $\theta2$, $\theta3$, $\theta4$, $\theta5$); the sum of holding is $0.05+0.10 = 0.15$ ($\theta6$, $\theta7$); and the sum of selling a position is $0.08+0.07+0.10 = 0.25$ ($\theta8$, $\theta9$, $\theta10$). In the example, buying the position has the highest weight sum (0.60). Therefore, at that specific time, the decision would be to buy the position. Eventually, GA evolves the weights of the individual θ's in order to maximize the fitness function, which represents the overall performance of the thresholds. Below, we define the operators and fitness function of the GA that we used in order to optimization of a better performance metric.

TABLE II
INDIVIDUAL CHROMOSOME REPRESENTATION. FIRST ROW IS FOR THE REPRESENTATION OF CHROMOSOME THAT IS CONSTRUCTED BY 10 THRESHOLD WEIGHT.

θ	$\theta1$	$\theta2$	$\theta3$	$\theta4$	$\theta5$	$\theta6$	$\theta7$	$\theta8$	$\theta9$	$\theta10$
W	0.1	0.2	0.05	0.15	0.1	0.05	0.1	0.08	0.07	0.1

2) Genetic Operators & Fitness Function

We use one-point crossover with a probability p and one-point mutation with a probability $1-p$. We also use elitism, to ensure that the best individual is copied to the next generation.

We use the Sharpe Ratio (SR) as our fitness function to take into account risk-adjusted returns. The equation for SR is as follows:

$$SR = \frac{R_p - R_f}{\sigma_p} \tag{4}$$

where, R_p is total rate of return for a given GA individual, R_f is risk-free asset, which is selected as 2.5% for a two-year data-set to preserve the resemblance of USA government bonds, and σ_p is the standard deviation of returns, i.e. the risk of the trading strategy. In what follows, we will be calculating the empirical SR by using the empirical Rate of Return (RoR) and the well-known empirical standard deviation derived from the asset data.

V. EXPERIMENTAL SETUP

In this section, we first describe the data we are using in Section V-A. Then, in Section V-B, we outline the procedure for adjusting the parameters of the GA. Finally, we cover the benchmarks in Section V-C.

A. Data

For this study, we are using 18 publicly traded stocks' daily adjusted closing prices listed on the New York Stock Exchange. The tickers are: ALL, ASGN, CI, COP, CTXS[1], EME, EVR, GILD, GPK, ISRG, MKL, MOH, PEG, PXD, QCOM, UBSI, VFC, XEL. The time period is from November 27, 2009 to November 27, 2019, and it is sourced from YAHOO Finance [18], using the python library yfinance. The data is divided into three parts: 60% for training, 20% for validation, and 20% for testing. We chose this time period to exclude any potential distortion in the stock market data caused by the COVID-19 pandemic.

B. GA parameters tuning

We conducted a grid search to fine-tune the following parameters: population size, with values in $\{20, 50, 70, 100, 150, 200, 300\}$; generations number, with values in $\{15, 18, 25, 30, 35, 45\}$; and crossover probability p, with values in $\{0.75, 0.85, 0.95, 0.99\}$. It's worth noting that the mutation probability is equal to $1 - p$. In tuning, after removing the last two years of the entire 10-year data, we further split the remaining data into new 80% training and 20% validation sets, and performed grid search on this subset of data. The tuned parameters are presented in Table III. Each GA experiment was run for 50 individual runs.

TABLE III
GA PARAMETER TUNING

Population size	100
Generations	18
Crossover probability	0.95
Mutation probability	0.05
Tournament size	2

C. Benchmarks

1) DC Benchmarks

The goal of this study is to show that using a stochastic search technique, namely a genetic algorithm to optimize recommendations from multiple thresholds, we improved the trading performance, and surpassed that of single threshold-based strategies. It is worth noting that our method uses three strategies that we introduce. However, due to implementation

[1]Recently, the ticker changed into RKOS

of 10 thresholds for St1, St2, and 5 thresholds for St3, while the methodology of strategies stays the same, these 10 thresholds form 10 individual recommendations to St1 and St2; similarly 5 of those thresholds form 5 individual recommendations to St3. Thus, each threshold determines different recommendations which produced different results. In order to differentiate them in Section VI, we will denote these results as $\theta 1$, $\theta 2$, $\theta 3$, $\theta 4$, $\theta 5$, $\theta 6$, $\theta 7$, $\theta 8$, $\theta 9$ and $\theta 10$.

In order to test the performance of strategies (St1, St2, and St3) that use recommendations from a multiple-threshold DC setup, we compared them to the performance of single-threshold strategies. Single-threshold strategies are based on 10 different thresholds for St1 and St2, and 5 different thresholds for St3. Finally, the multiple-threshold DC strategies optimized using the GA will be called *GA-optimized strategies* and will be denoted by *GA1*, *GA2*, and *GA3*, respectively.

2) Financial Benchmarks

We use two benchmarks from technical analysis, the relative strength index (RSI), and the moving average convergence divergence (MACD). From the literature, it is worth noting that in the majority of related studies to ours, Buy-and-Hold, RSI, and MACD have been the primary benchmarks used for comparison. They are widely used approach in trading that relies on technical indicators to guide trading decisions. We utilize both MACD and RSI with default period lengths of 26 and 12 for MACD and 14 for RSI in this research.

As another benchmark, we use Buy and Hold (BandH), which is a well-known benchmark for trading. This is a passive investment strategy in which an investor purchases the product and holds them for a long period of time, regardless of market fluctuations. In our model, the trader buys the financial product on the first day of the test set and sells it on the last day.

VI. Results

Table IV shows the Sharpe Ratio (SR), Rate of Return (RoR), and Standard Deviation (Risk) results of 18 stocks for each of the three strategies in combination with the ten thresholds, while also highlighting the GA-optimized strategies' performance. We note that the metric we are using is the average (for each of the SR, RoR, and Risk) over the 18 stocks for each individual combination of threshold (or GA-optimized thresholds) and strategy. Furthermore, we use the test set result from a certain run among the 50, which is supplied by the best results among the training runs with a highest SR. While it is important to evaluate the performance by number of runs, it is also essential to identify the best trained model for use in real-world applications. By this approach, we reflect the practical considerations involved in implementing GA-based trading strategies in real-world contexts, where trader would be implementing the best returned GA-model during the training phase.

Table IV also indicates that many thresholds have poor performance across all three metrics. Among the thresholds, $\theta 4$, $\theta 10$ under St1, $\theta 3$, $\theta 10$ under St2, and $\theta 1$, $\theta 2$ under St3 have relatively good SR compared to the other thresholds. Meantime, the RoR for all thresholds is middling, with St2 under $\theta 1$ having slightly higher profit. Again, the performance for risk under each individual thresholds fluctuates between 0.05 to 0.07 for St1, 0.07 to 0.11 for St2, and 0.02 for St3. In short, with a certain threshold, trader would be able to profit 9%, 16%, and 9% on average with 18 stocks if the portfolio is equally weighted, under St1, St2, and St3, respectively.

TABLE IV
AVERAGE PERFORMANCE RESULTS, GA AND THE 10 INDIVIDUAL DC-THRESHOLDS. WE PRESENT THE BEST VALUE PER METRIC IN BOLDFACE.

	Sharpe Ratio			**RoR**			**Risk**		
Algo.	**St1**	**St2**	**St3**	**St1**	**St2**	**St3**	**St1**	**St2**	**St3**
GA	**4.14**	**2.50**	**5.63**	**0.24**	**0.19**	**0.16**	**0.05**	0.08	**0.02**
$\theta 1$	0.43	0.51	2.83	0.07	0.16	0.09	0.07	0.11	**0.02**
$\theta 2$	0.52	0.47	2.31	0.07	0.13	0.09	0.07	0.11	**0.02**
$\theta 3$	0.22	0.76	2.00	0.05	0.09	0.09	0.07	0.09	**0.02**
$\theta 4$	0.57	-2.16	0.06	0.07	0.05	0.07	0.07	0.08	**0.02**
$\theta 5$	0.54	0.13	-2.45	0.09	0.06	0.04	0.06	0.09	**0.02**
$\theta 6$	0.49	0.68	-	0.07	0.12	-	0.07	0.10	-
$\theta 7$	0.28	-0.72	-	0.09	0.08	-	0.06	0.08	-
$\theta 8$	0.41	-0.87	-	0.08	0.07	-	0.07	0.09	-
$\theta 9$	-0.35	-2.34	-	0.01	0.08	-	0.07	**0.07**	-
$\theta 10$	0.59	1.63	-	0.05	0.10	-	0.06	0.08	-

From Table IV it is also apparent that by the multiple-thresholds GA optimization, SR and RoR are improved drastically. From the Sharpe Ratio columns, single-threshold strategies $\theta 10$ in St1, $\theta 10$ in St2 and $\theta 1$ in St3 showed values of 0.59, 1.63, 2.83, respectively. Whereas, using GA we achieve nearly 8, 1.5, and 2 times the performance of the best single-threshold strategy (over the ten individual thresholds). RoR columns show us that all three GA-optimized strategies outperform their corresponding single-threshold strategies when compared to all ten thresholds. In St1, it yields more than twice the profit of its best performed single threshold strategy, whereas, in St3, it nearly doubles the profit. Additionally, in terms of risk, all strategies have performed similarly well with values ranging from 2% to 8%. Although, the St2 risk is 0.07%, and it falls behind the $\theta 9$ with a low margin, all of the GA-optimized strategies had relatively the same performance with their corresponding single-threshold strategies. In summary, the multiple-threshold optimization by the GA enhanced the trading performance in the SR and RoR metrics.

It should also be noted that the high SR and RoR values may be partially attributed to the strong bull market at the test set period. But more importantly, while the bull market may affect the SR and RoR performance, the DC paradigm also plays a huge role in the results. This will be evident when we show the results of our GA-optimized strategies to benchmarks such as BandH, RSI, and MACD, where we found similarly high performance.

For the performance metrics SR, RoR, and Risk, we also performed the Friedman non-parametric statistical test (Tables V, VI, and VII). The null hypothesis is that all strategies, including the GA-optimized strategy, come from the same continuous distribution. The second and fifth columns show the Friedman Ranking (FR) based on the SR and RoR metrics. The third and sixth columns show the adjusted p-value (p_{adj})

of the post-hoc Conover test [19].

From Table V, it is clear that the best ranking strategy is the GA-optimized strategy in St1. Also, we reject the null hypothesis at Sharpe Ratio, whereas, we only fail to reject it at RoR at 5% significance level for $\theta 10$. However, both results would be statistically significant if the significance level is 10%. In Table VI, the GA-optimized strategy is ranked first in both metrics. From the third column, we can see that the performance is statistically significant against every single threshold strategy for SR under St2, however, we fail to reject $\theta 10$, $\theta 8$, and $\theta 5$ strategies in RoR from the sixth column. From Table VII, it is indicative that the GA-optimized strategy is first in the rankings, and is also statistically significant against every single threshold. Therefore, it is noteworthy that GA-optimized strategies substantiated better performance in comparison to the great majority of single-threshold strategies.

Table VIII displays that rankings and p-values for strategies at the metric of Risk differs from the substantial progress that we achieved at SR and RoR. We find the GA-optimized ranked as 2nd, 2nd, and 5th, respectively to strategies, without any statistical significance. Thus, future work could focus on guiding the GA search towards solutions that are less risky.

TABLE V
NON-PARAMETRIC FRIEDMAN TEST WITH CONOVER POST-HOC TEST FOR ST1 BASED ON THE SHARPE RATIO AND RATE OF RETURN.

Sharpe Ratio			Rate of Return		
Algo.	FR	p_{adj}	Algo.	FR	p_{adj}
GA	2.28	-	GA	2.78	-
$\theta 10$	5.39	**0.0292**	$\theta 10$	5.56	**0.0714**
$\theta 5$	5.94	**0.0064**	$\theta 8$	5.72	**0.0528**
$\theta 4$	6.00	**0.0064**	$\theta 5$	5.78	**0.0517**
$\theta 8$	6.00	**0.0064**	$\theta 4$	6.06	**0.0286**
$\theta 3$	6.28	**0.0038**	$\theta 1$	6.33	**0.0154**
$\theta 7$	6.28	**0.0038**	$\theta 3$	6.44	**0.0134**
$\theta 1$	6.39	**0.0038**	$\theta 6$	6.50	**0.0134**
$\theta 6$	6.44	**0.0038**	$\theta 7$	6.56	**0.0134**
$\theta 2$	7.00	**0.0008**	$\theta 2$	6.61	**0.0134**
$\theta 9$	7.89	**7.3e-05**	$\theta 9$	7.67	**0.0012**

TABLE VI
NON-PARAMETRIC FRIEDMAN TEST WITH CONOVER POST-HOC TEST FOR ST2 BASED ON THE SHARPE RATIO AND RATE OF RETURN.

Sharpe Ratio			Rate of Return		
Algo.	FR	p_{adj}	Algo.	FR	p_{adj}
GA	2.33	-	GA	3.06	-
$\theta 6$	5.33	**0.0309**	$\theta 10$	5.11	0.2164
$\theta 10$	5.56	**0.0176**	$\theta 8$	5.28	0.2069
$\theta 2$	5.78	**0.0071**	$\theta 5$	5.67	0.1099
$\theta 3$	6.06	**0.0058**	$\theta 4$	5.83	**0.0963**
$\theta 1$	6.11	**0.0041**	$\theta 1$	5.89	**0.0856**
$\theta 4$	6.11	**0.0032**	$\theta 3$	6.11	**0.0626**
$\theta 7$	6.50	**0.0021**	$\theta 6$	6.22	**0.0626**
$\theta 9$	6.50	**0.0018**	$\theta 7$	7.22	**0.0030**
$\theta 5$	6.78	**0.0011**	$\theta 2$	7.33	**0.0030**
$\theta 8$	6.89	**0.0011**	$\theta 9$	7.39	**0.0030**

Additionally, we compare our GA optimized results with other financial benchmarks, namely, RSI, MACD, and the widely known buy-and-hold (BandH). The reason we chose

TABLE VII
NON-PARAMETRIC FRIEDMAN TEST WITH CONOVER POST-HOC TEST FOR ST3 BASED ON THE SHARPE RATIO AND RATE OF RETURN.

Sharpe Ratio			Rate of Return		
Algo.	FR	p_{adj}	Algo.	FR	p_{adj}
GA	1.67	-	GA	1.39	-
$\theta 2$	3.00	**0.0564**	$\theta 2$	3.44	**0.0033**
$\theta 1$	3.61	**0.0078**	$\theta 3$	3.72	**0.0010**
$\theta 3$	3.83	**0.0030**	$\theta 1$	3.83	**0.0007**
$\theta 4$	4.17	**0.0006**	$\theta 4$	4.06	**0.0003**
$\theta 5$	4.56	**0.0001**	$\theta 5$	4.56	**2.98e-05**

TABLE VIII
NON-PARAMETRIC FRIEDMAN TEST WITH CONOVER POST-HOC ON THREE STRATEGIES' RISK INDIVIDUALLY.

St1			St2			St3		
Algo.	FR	$p_{adj.}$	Algo.	FR	$p_{adj.}$	Algo.	FR	$p_{adj.}$
$\theta 10$	**5.00**	**-**	$\theta 9$	**4.27**	**-**	$\theta 5$	**2.8**	**-**
GA	5.17	0.880	GA	5.44	0.436	$\theta 3$	3.25	0.476
$\theta 5$	5.72	0.880	$\theta 7$	5.13	0.436	$\theta 2$	3.44	0.476
$\theta 4$	5.78	0.880	$\theta 8$	5.49	0.436	$\theta 4$	3.5	0.476
$\theta 9$	5.89	0.843	$\theta 10$	5.72	0.436	GA	3.83	0.397
$\theta 3$	5.94	0.786	$\theta 4$	6.27	0.352	$\theta 1$	4.16	0.145
$\theta 1$	6.06	0.770	$\theta 3$	6.52	0.211	-	-	
$\theta 8$	6.39	0.685	$\theta 5$	6.69	0.173	-	-	
$\theta 7$	6.5	0.685	$\theta 1$	6.75	0.152	-	-	
$\theta 2$	6.61	0.642	$\theta 2$	6.75	0.152	-	-	
$\theta 6$	6.94	0.587	$\theta 6$	6.91	0.119	-	-	

them is the fact that their construction follows the idea of trend chasing. That is why comparison of a trend based paradigm, namely DC, corresponds to these benchmarks. Table IX demonstrates that strategies St1, St2, and St3 yield on average 24.5%, 18.2%, and 16.1% at RoR, respectively, while the medians are 19.9%, 14.8%, and 11.7% for that metric, respectively. Also, RSI, MACD, and BandH were not capable of outperforming any of the GA-optimized strategies result on averages or medians. The metric values for these benchmark are: 11.17%, 1.3%, 14.9% on average and 0.5%, 2.93%, 11.6% on median, orderly by RSI, MACD, and BandH. Furthermore, while GA1 were able to outrank the other GA-optimized strategies and financial benchmarks, it only outperformed the RSI and MACD statistically at the 10% significance level.

TABLE IX
NON-PARAMETRIC FRIEDMAN TEST WITH CONOVER POST-HOC ON GA-OPTIMIZED AND FINANCIAL BENCHMARK STRATEGIES BASED ON ROR.

Algo.	Average	Median	FR	p_{adj}
GA1	**0.244**	**0.199**	**2.33**	-
GA2	0.186	0.148	2.78	0.5247
GA3	0.161	0.117	3.61	0.1211
BandH	0.149	0.116	3.67	0.1211
RSI	0.117	0.005	3.94	**0.0549**
MACD	0.013	0.0293	4.67	**0.0051**

From the Table X, it is indicative that each GA-optimized strategy outperformed the financial benchmarks at the risk adjusted return (SR) metric. Especially, GA3 achieved nearly

3.5 and 25 times the performance of the RSI, and MACD strategies, respectively. At 10% significance level, it statistically backed the performance over GA2, RSI, and MACD. It also cannot be overlooked that other GA-optimized strategies came as second and third in the rankings. Let us also reiterate that BandH acts on only one trade, where it buys the stock at first day and sells it on the last day. Therefore, there is no point in measuring SR and Risk.

TABLE X
NON-PARAMETRIC FRIEDMAN TEST WITH CONOVER POST-HOC ON GA-OPTIMIZED AND FINANCIAL BENCHMARK STRATEGIES BASED ON SR.

Algo.	Average	Median	FR	p_{adj}
GA3	**5.628**	**6.580**	**2.00**	-
GA1	4.142	3.762	2.44	0.3169
GA2	2.504	2.546	3.00	**0.0910**
RSI	1.607	-0.295	3.67	**0.0091**
MACD	0.172	0.403	3.89	**0.0050**

We also performed the Friedman statistical test with Conover post-hoc test on Risk (Table XI). Even though GA3 ranked first overall, the performance was not statistically significant against any of the other GA-optimized strategies and MACD, except RSI. Nevertheless, the risk gets slightly reduced when we trade on GA3.

TABLE XI
NON-PARAMETRIC FRIEDMAN TEST WITH CONOVER POST-HOC ON GA-OPTIMIZED AND FINANCIAL BENCHMARK STRATEGIES BASED ON RISK

Algo.	Average	Median	FR	p_{adj}
GA3	**0.021**	**0.018**	**1.06**	**-**
MACD	0.045	0.041	2.94	0.3624
GA1	0.052	0.052	3.22	0.2637
GA2	0.073	0.061	3.72	0.1227
RSI	0.080	0.077	4.06	**0.0002**

VII. CONCLUSION

In this paper, we suggest new trading strategies applied on an extension of the DC paradigm, where we use multiple thresholds instead of a single one. We demonstrate a new technique for achieving more profitable results in the financial trading field by using strategies emerging from the DC paradigm and optimizing them using a GA. We conjecture that this happens due to the fact that (a) the strategy space is now enhanced with a richer set of options for the traders, and (b) stochastic search via GA in the multi-threshold model is sufficiently strong to pinpoint strategies that turn out to perform better than the single-threshold ones. We conducted experiments where 18 stocks were tested under 10 different DC thresholds for strategy St1 and St2, and 5 thresholds for strategy St3, and used a GA to find a good mixture of thresholds. From our results we can draw the following conclusions: (i) The multi-threshold DC paradigm is capable of generating trading strategies that are profitable, (ii) using GA as an optimizer produces the highest SR and RoR among the great majority of the individual thresholds, and (iii) the GA-optimized strategy statistically outperforms the MACD and RSI benchmarks.

In the future, we plan to use the GA optimization approach to test trading strategies with multiple thresholds, and expand our research on a risk-oriented work. We believe that extending the chromosomes to contain multiple thresholds and multiple strategies will lead to even better performance.

REFERENCES

[1] H. Markowitz, "Portfolio selection*," *The Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.1952.tb01525.x

[2] S. Katoch, S. S. Chauhan, and V. Kumar, "A review on genetic algorithm: past, present, and future," *Multimedia Tools and Applications*, vol. 80, no. 5, pp. 8091–8126, 2021.

[3] S. Mahfoud and G. Mani, "Financial forecasting using genetic algorithms," *Applied artificial intelligence*, vol. 10, no. 6, pp. 543–566, 1996.

[4] A. Ponsich, A. L. Jaimes, and C. A. C. Coello, "A survey on multiobjective evolutionary algorithms for the solution of the portfolio optimization problem and other finance and economics applications," *IEEE Transactions on Evolutionary Computation*, vol. 17, no. 3, pp. 321–344, 2013.

[5] Y. Hu, K. Liu, X. Zhang, L. Su, E. Ngai, and M. Liu, "Application of evolutionary computation for rule discovery in stock algorithmic trading: A literature review," *Applied Soft Computing*, vol. 36, pp. 534–551, 2015.

[6] C.-F. Huang, C.-J. Hsu, C.-C. Chen, B. R. Chang, and C.-A. Li, "An intelligent model for pairs trading using genetic algorithms," *Computational Intelligence and Neuroscience*, vol. 2015, pp. 16–16, 2015.

[7] R. Kumar, P. Kumar, and Y. Kumar, "Multi-step time series analysis and forecasting strategy using arima and evolutionary algorithms," *International Journal of Information Technology*, vol. 14, no. 1, pp. 359–373, 2022.

[8] O. Salman, M. Kampouridis, and D. Jarchi, "Trading strategies optimization by genetic algorithm under the directional changes paradigm," in *2022 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2022, pp. 1–8.

[9] X. Long, M. Kampouridis, and P. Kanellopoulos, "Genetic programming for combining directional changes indicators in international stock markets," in *Parallel Problem Solving from Nature–PPSN XVII: 17th International Conference, PPSN 2022, Dortmund, Germany, September 10–14, 2022, Proceedings, Part II*. Springer, 2022, pp. 33–47.

[10] E. Christodoulaki and M. Kampouridis, "U sing strongly typed genetic programming to combine technical and sentiment analysis for algorithmic trading," in *2022 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2022, pp. 1–8.

[11] J. B. Glattfelder, A. Dupuis, and R. Olsen, "An extensive set of scaling laws and the fx coastline," *Unpublished manuscript*, 2008.

[12] J. B. Glattfelder, A. Dupuis, and R. B. Olsen, "Patterns in high-frequency fx data: discovery of 12 empirical scaling laws," *Quantitative Finance*, vol. 11, no. 4, pp. 599–614, 2011.

[13] E. P. Tsang, R. Tao, A. Serguieva, and S. Ma, "Profiling high-frequency equity price movements in directional changes," *Quantitative finance*, vol. 17, no. 2, pp. 217–225, 2017.

[14] E. Tsang and J. Chen, "Regime change detection using directional change indicators in the foreign exchange market to chart brexit," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 3, pp. 185–193, 2018.

[15] R. Tao, "Using directional change for information extraction in financial market data," Ph.D. dissertation, University of Essex, 2018.

[16] A. Adegboye, M. Kampouridis, and F. Otero, "Improving trend reversal estimation in forex markets under a directional changes paradigm with classification algorithms," *International Journal of Intelligent Systems*, 2021.

[17] ——, "Algorithmic trading with directional changes," *Artificial Intelligence Review*, pp. 1–26, 2022.

[18] R. Aroussi, "yfinance," https://pypi.org/project/yfinance/.

[19] W. J. Conover and R. L. Iman, "Rank transformations as a bridge between parametric and nonparametric statistics," *The American Statistician*, vol. 35, no. 3, pp. 124–129, 1981.