Chen, J. & Tsang, E.P.K., Tracking Regime Changes, IEEE Conference on Computational Intelligence for Financial Engineering and Economics (CIFEr) 2019 Conference, Shenzhen, China, 4th-5th May 2019

Tacking Regime Changes

1st Jun Chen
Centre for Computational Finance and Economic Agents (CCFEA)
University of Essex
Colchester, UK
jchenaj@essex.ac.uk

2nd Edward P K Tsang
Centre for Computational Finance and Economic Agents (CCFEA)
University of Essex
Colchester, UK
edward@essex.ac.uk

***Abstract*—In our previous work, we showed that regime changes in the market are retrospectively detectable using historic data in directional changes (DC). In this paper, we build on such results and show that DC indicators can be used for market tracking - using data up to the present - to understand what is going on in the market. In particular, we wanted to track the market to see whether the market is entering an abnormally volatile regime.**

The proposed approach used DC indicator values observed in the past to model the normal regime of a market (in which volatility is normal) or an abnormal regime (in which volatility is abnormally high). Given a particular value observed in the current market, we used a naive Bayes model to calculate independently two probabilities: one for the market being in the normal regime and one for it being in the abnormal regime. These two probabilities were combined to decide which regime the market was in, two decision rules were examined: a Simple Rule and a Stricter Rule.

We used DJIA, FTSE 100 and S&P 500 data from 2007 to 2010 to build the Bayes model. The model was used to track the S&P 500 market from 2010 to 2012, which saw two spells of abnormal regimes, as identified by our previous work, with the benefit of hindsight. The tracking method presented in this paper, with either decision rule, managed to pick up both spells of regime changes accurately. The tracking signals could be useful to market participants. This study potentially lays the foundation of a practical financial early warning system.

Keywords*—*regime change, directional change, financial market monitoring

I. Introduction

In our previous work, we proposed an approach to detect regime changes. The empirical results showed that the detected regime changes coincided with a significant market event, the UK's referendum in June 2016 on Brexit. Our work proved that the fluctuation of the financial market could be detected and summarised as regime changes, in other words moving from one regime to another. However, such regime changes are detected in hindsight. The question in this paper is whether one could use the information up to the present time to track regime changes, as they occur in real time?

In [1], we classified markets into two regimes. We named the regime with higher volatility the "abnormal regime", as it emerged after a significant event (namely, the Brexit referendum) had taken place. In a subsequent work, we applied the same method to different financial markets [2]. We found that the two regimes have unique characteristics; in other words, we can characterise normal and abnormal regimes across different financial markets.

The authors would like to thank Sara Colquhoun for proofreading the article.

However, in both [1] and [2], regimes were detected retrospectively. In this paper, we explain how one could use data up to the present to track the market, with the aim to recognise regime changes, preferably without too much delay.

The tracking method proposed can be divided into two steps: firstly, we use a naive Bayes model to compute the probability of the market being in the normal or abnormal regime, based on (i) the market data observed up to the present; and (ii) characteristics of the past regimes observed across markets (in [2]). Secondly, the two probabilities computed are combined to form a final classification on which regime the market is currently in.

The method proposed in this paper monitored the market as prices changed. Thus, it could be employed as a warning system, alerting market participants of likely regime changes. How traders and regulators may act upon such information is beyond the scope of this paper. It is also important to clarify that in this paper, we purely focus on what the data up to now tells us about the market: i.e. it is purely data-led. No forecasting is attempted.

II. Backgrounds

The price and time of each transaction are recorded in markets. Such data transactions are normally summarised using time series, e.g. the final transaction price of each day is recorded to form the end-of-day time series.

However, Directional Change (DC) is an alternative way to record data [3]. Instead of recording the transaction prices at fixed intervals, as it is done in time series, DC lets the data alone dictate when to record a transaction. A data point is recorded when the price has risen or dropped against the current trend by a significant percentage (threshold), and where the significance is determined by the observer. Different observers may use different percentage thresholds. Under this sampling scheme, markets are partitioned into up and down trends, or bear and bull markets, which traders are familiar with. The idea of DC has been known as "zigzag" by technical traders [4].

The definition of DC can be found in [5]. Figure 1 gives an example of price movement over time. In this example, we assume that the threshold is θ. Point A (time 4) is the lowest price before Point B (time 5), which rose θ above A. At time 5, we retrospectively confirm A as an extreme point, which started an uptrend. Similarly, Point C (time 9) is an extreme point, which ends the uptrend, and starts a downtrend. Point C is confirmed to be an extreme point at Point E (time 11). It is not confirmed at Point D (at time 10) because the price at D is not θ below C.

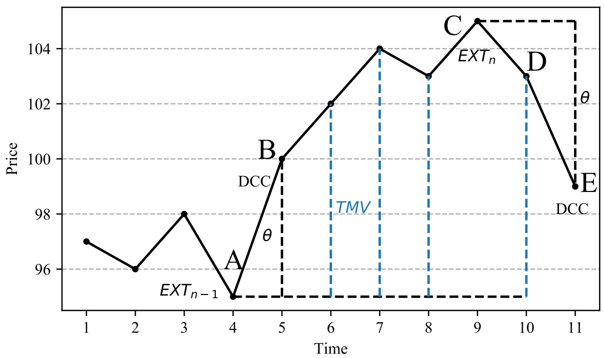

Fig. 1. A hypothetical example of tracking the unfinished DC trends.

With DC, many techniques for analysing time series do not apply. For example, daily returns are used in time series to measure volatility, but DC does not record daily returns – it only records a data point as required. Before analysis is possible, a new vocabulary is needed to describe the DC series. Tsang et al have introduced a number of DC indicators [6]. Here we introduce two indicators, which are relevant to the rest of this paper.

$$TMV(i) = \frac{P(i) - P_{EXT_i}}{P_{EXT_i} \times \theta}, \tag{1}$$

where $P(i)$ is the price at time i, P_{EXT_i} is the price of the last extreme point, θ is the threshold. $TMV(i)$ is the Total Movement from the last extreme point known to time i. The change is normalised with θ so that TMV values computed from different thresholds can be compared.

$$T(i) = t(i) - T_{EXT_i}, \tag{2}$$

where $t(i)$ is the time of point i and $T_{EXT_i}$is the time of the last extreme point. $T(i)$ is the time elapsed since the last extreme point known to time i.

In [6] we explained that TMV and T are measures of volatility under DC. TMV and T combine to define Return (R), which was used in [1] and [2] to detect regimes. In the next section, we explain how we could track the market by monitoring the values of TMV and T continuously.

III. Methodology

In this section, we propose a method for tracking price movements dynamically with the aim to detect what are the likely regime changes in the market. The idea is to observe the TMV and T in the current trend, and compare these values to those found in the normal regimes indicated in the past [2]. A naive Bayes classifier is applied to compute two probabilities independently: (i) the probability of the market being in the normal regime; and (ii) the probability of the market being in the abnormal regime. These two probabilities are combined to conclude what regime the market is currently in – two decision methods will be proposed below.

A. Tracking DC trends

Summarising the financial data into DC trends using the DC approach enabled us to focus on what are the significant price changes. However, according to the definition of DC [5], one DC trend will not be confirmed until the next DC event is triggered. This may cause a delay when tracking regime changes based on DC trends. Therefore, a dynamic way is required to track the DC trends to take the research further.

We shall use Figure 1 to explain what tracking means in this paper. At every time point i, the values of $TMV(i)$ and $T(i)$ are calculated. With the help of past data, one attempted to infer from these values whether the market is in or out of the normal regime at time i. It is important to note that tracking uses data up to now. For example, the last known extreme point at time 10 is Point A at time 4, because C is not confirmed as an extreme point until time 11 (when the price drops from C by θ). Therefore, TMV(10) should be calculated with the price at A as P_{EXT_i}; T(10) should be calculated with the time at A as and T_{EXT_i}.

B. Naive Bayes model

The price movements tracked are compared with what has been observed in the past. We have shown in our previous work how to characterise normal and abnormal regimes [2]. Given the tracked TMV and T in the on-going trend, we can compare their values with those observed in normal and abnormal regimes in the past (which we refer to as training data) to calculate the probability of the current market being in either regime. For that, we employ the use of a naive Bayes model. A naive Bayes model calculates the probability of the data belonging to a particular class, given what has been observed in the past. It is called "naive" because it makes a strong assumption that the existence of a particular feature of a class is independent or unrelated to the existence of all other features. This means that the existence of one particular feature does not affect the other.

The naive Bayes model is established based on applying the Bayes theorem. In classification, Bayes' theorem is used to calculate the probabilities of the classes [7].

Using the Bayes theorem, we can calculate the conditional probability $p(C_k|x)$, after seeing the observation, x. The formula is given by:

$$p(C_k|x) = \frac{p(C_k)p(x|C_k)}{p(x)}, \tag{3}$$

where C is the class variable of k possible outcome. In our case, two different market regimes are considered. So there are two possible outcomes of the class variable, where Regime 1 is denoted by C_1 and Regime 2 is denoted by C_2. And $p(C_k)$ is the probability that the market being in regime k, regardless of the input features x.

The input features is represented by a vector $x = (x_1, ..., x_n)$, which represented n features. As discussed in Section III-A, two DC indicators are considered as input features: TMV and T. The variable x can be written as $x_i = (TMV_i, T_i)$.

$p(x|C_k)$ is the probability of seeing the input x when it is known to belong to regime C_k. In our case, we apply the naive Bayes model to calculate two conditional probabilities:

1) $p(C1)$: The probability of the market being in the normal regime.
2) $p(C2)$: The probability of the market being in the abnormal regime.

We assume, for simplicity, that each feature is distributed according to a Gaussian distribution:

$$P(x|C) = \frac{1}{\sqrt{2\pi\sigma_C^2}} exp\left(-\frac{(x-\mu_C)^2}{2\sigma_C^2}\right), \tag{4}$$

where μ_C and σ_C represent the mean and the variance of each input feature x in each class C. We first separated the observed instances in the training data by the class, then the mean and variance of each input feature can be computed according to each class.

$p(x)$ is the marginal probability that a input feature x is seen, regardless of whether it is in Regime 1 or Regime 2. By substituting the variable x with our input features, the Bayes model can be described as:

$$p(C_k|TMV_i, T_i) = \frac{p(C_k)p(TMV_i, T_i|C_k)}{p(TMV_i)p(T_i)}, \tag{5}$$

Now given the observation of the input features and the past class values, the Bayes model can be established. With it, the probability of the market belonging to a particular regime can be calculated. We shall use such probabilities to estimate the class in unseen data.

C. Data

The empirical study of this paper focused focuses on three stock indices: the Dow Jones Industrial Average (DJIA) index, the FTSE 100 index, and the S&P 500 index. The three stock indices were chosen because they were linked to the 2007-2008 global financial crisis, where regime changes were considered to have taken place in financial markets. We used daily closing prices sampled from January 2007 to December 2012 to cover periods of the financial crisis.

To examine the proposed method, the dataset was separated into two datasets: training dataset and test dataset (see Table I). What were the parameters of the Bayes model was estimated from the training datasets. Then the model was used to detect regime changes on the test dataset.

The training datasets were comprised of the data of the completed DC trends. This is because regime changes were detected by analysing the value of DC trends, on the basis of the regime change detection approach proposed in [1].

The raw financial data in the training datasets is summarised into DC trends, under a threshold of 0.3%. The DC trends were then measured by two DC indicators: TMV_{EXT} and T_{EXT}.

In the test datasets, the raw financial data was summarised into the on-going DC trends, which were then measured by two DC indicators TMV and T. As discussed in section III-A, their values were used to track the market. In [2], we show that the two regimes were clearly separable on the TMV-T space when both TMV and T are normalised. Therefore, the training and test data were both normalised before modelling by the Bayes classifier. The data was normalised using the min-max normalisation approach.

In the empirical study, the parameter of the Bayes model was learnt from the training datasets. And the model was used to recognise market regimes for each pair of input features from the test dataset.

TABLE I. TIME PERIODS OF TRAINING AND TEST DATASET.

Data	Time periods	
	Training	Test
Dow Jones	03/01/2007 - 05/04/2010	
FTSE 100	03/01/2007 - 06/04/2010	
S&P 500	03/01/2007 - 06/04/2010	07/04/2010 - 28/12/2012

IV. EMPIRICAL RESULTS

In this section, we will analyse the effectiveness of the Bayes model and the two decision rules. The purpose of this analysis is to investigate whether the proposed method is able to track regime changes on the test dataset.

The Bayes model is established with observations in the training datasets. The model is then used to monitor the market, calculating for each day in the test dataset the probabilities of the market being in the two regimes, $p(C1)$ and $p(C2)$. With the probabilities calculated, we attempt to combine them and determine which regime the current market belongs to. Two decision rules are designed and compared. The regimes classified by these rules are compared with the regimes computed by the method presented in [1]. This comparison allows us to assess the performance of our classification approach.

A. Calculating Probability

As discussed in Section III-B, the conditional probability of the occurrence of the current market, given the TMV and T values in the current trend, can be calculated by the Bayes model. Figure 2 shows the calculated probabilities of the market belonging to Regime 1 (the normal regime) and to Regime 2 (the abnormal regime), $p(C1)$ and $p(C2)$, over time. A higher $p(C1)$ value means the market is more likely to be in the normal regime; similarly, a higher $p(C2)$ means the higher likelihood that the current market is in the abnormal regime. For instance, as shown in Figure 2, from the middle of 2011 to the early of 2012, the probabilities of being in Regime 2 are much higher than that of being in Regime 1. This may imply that the market fell into Regime 2 in that period.

For each (TMV, T) pair, two probabilities are calculated independently. For $p(C1)$ and $p(C2)$ to be useful, we need to combine them in order to decide whether the market is in Regime 1 or Regime 2. Two decision rules are proposed in the next sections.

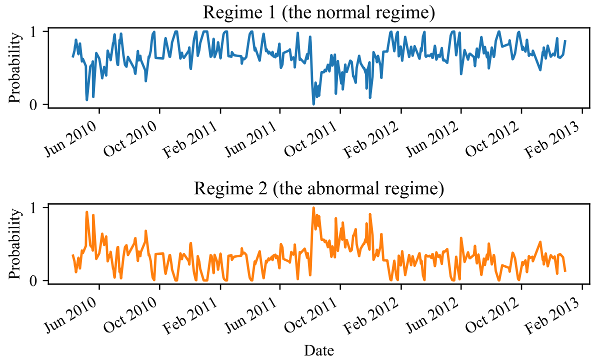

Fig. 2. Estimated probabilities for testing data. The blue line indicates the probability of the market being in Regime 1, and the orange line indicates the probability of the market being in Regime 2

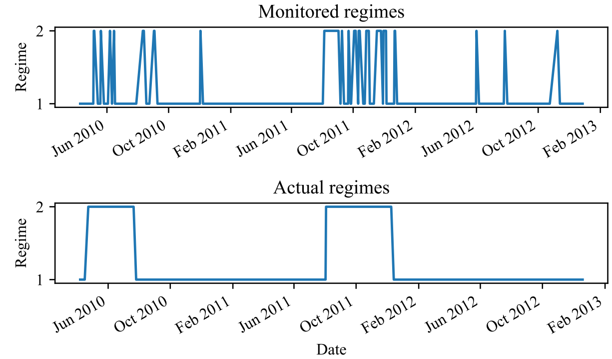

Fig. 3. Comparison of the tracked market regimes and the actual regimes. The tracked regimes are determined using Equation 6.

B. B-Simple for Regime Classification

The Bayes model compared each (TMV, T) pair with those found in the training data, and calculated the probabilities of the market belonging to Regime 1 and Regime 2 independently. A Simple Rule is where the hypothesis picked is most probable. In our case, it meant choosing the regime with the highest probability:

$$\begin{array}{ll} \text{choose } C_1 & \text{if } p(C_1) > p(C_2) \\ \text{choose } C_2 & \text{if } p(C_2) > p(C_1), \end{array} \tag{6}$$

where C_1 and C_2 represent Regime 1 and Regime 2, $p(C_1)$ and $p(C_2)$ denoted the probability of the market belonging to Regime 1 and Regime 2, respectively. We call this approach that combined the Bayes model, with the Simple Rule B-Simple.

The top part of Figure 3 shows the regime classification using B- Simple. The lower part of Figure 3 shows the market regimes computed by the method proposed in [1]; we call them the actual regimes as they were computed with the benefit of hindsight.

Figure 3 showed the performance of tracking using the B-Simple. The key issue to observe is: does the tracking mechanism has the ability to detect Regime 2 when it happened? If so, how long does it take the tracking mechanism to realise regime change has occurred after it has taken place?

Firstly, both spells of Regime 2 are detected. This means, by using data up to the time when regime classification is made, B-Simple detected Regime 2 when it took place. As tracking does not have the benefit of hindsight, it is reasonable to expect delay: in other words, it may take some time before the tracking mechanism realised that regime change has taken place. The first spell of Regime 2 in the monitored period (as computed by the method in [1]) took place from 27 April 2010 to 26 July 2010. Tracking B-Simple, regime change was detected on 6 May 2010, ten days in arrears. This is not a bad outcome, because the actual regime changes were computed with the benefit of hindsight. The tracking method proposed here lets the data tell us what happens in the market: no forecasting is attempted.

However, in the second period of Regime 2, B-Simple suggested that regime change occurred ahead of the actual change. The actual regime changes took place from 8 August 2011 to 14 December 2011. B-Simple suggested regime 2 took place on 5 August 2011, three days ahead of the actual regime change. This is possibly because in [1], regime changes are computed based on completed trends. In tracking, we are dealing with on-going trends. Therefore, when the TMV value goes sufficiently high within a short time, regime 2 could have been concluded.

The second point to note is that B-Simple raised the alarm of regime change repeatedly, as opposed to raising persistent alarms throughout the Regime 2 spell. This is understandable because the method proposed in [1] attempted to model the hidden Markov state, which carried a momentum. In B-Simple, only the current (TMV, T) reading is used for decision making. Besides, in practice, traders could react when such an alarm is first raised. So the alarms raised by B-Simple do not have to be persistent to be useful to users. Repeated alarms would simply reinforce the message.

Thirdly, even though the tracking managed to detect both regime changes, twelve false alarms are also raised. This means where B-Simple suggested that the market was in Regime 2, the market was actually in Regime 1, according to the method proposed in [1]. In fact, the detection of the second spell of Regime 2 three days ahead may in fact be a lucky false alarm. In general, as long as false alarms do not happen too often, which they do not, the ability to track signals would be useful to market participants.

C. B-Strict for Regime Classification

As discussed in the previous section, some false alarms are reported in the mentioned regimes, with B-Simple (see the Simple Rules defined in Equation 6). If we want to reduce false alarms, we could combine the outcome probability of the Bayes model with a stricter classification rule. What makes up the designed stricter decision rule is made up as follows:

$$\begin{aligned}&\text{choose } C_1 \quad \text{if } p(C_1) > p(C_2)\\&\text{choose } C_2 \quad \text{if } p(C_2) > p(C_1) \text{ and } p(C_2) > threshold_2,\end{aligned} \tag{7}$$

where C_1 and C_2 represented Regime 1 and Regime 2, $p(C_1)$ and $p(C_2)$ denoted the probability of the market belonging to Regime 1 and Regime 2, respectively. Here $threshold_2$ is the lower bound value of $p(C_2)$ for the market to be concluded in Regime 2. The value of $threshold_2$ is a parameter defined by investors, reflecting its cautiousness of concluding Regime 2. The Stricter Rule is exactly the same as the Simple Rule, except that a minimal probability of $p(C_2)$ must be observed before concluding Regime 2.

We call the method of combining Bayes model with the stricter rule - B-Strict. Figure 4 compared the B-Strict-monitored regimes and the actual regimes. The threshold2 value was set to 0.8. B-Strict will only classify the current market to be in Regime 2 if $p(C2)$ is not only greater than $p(C1)$, but $p(C2)$ is also above 0.8. With B-Strict, fewer periods of Regime 2 are concluded.

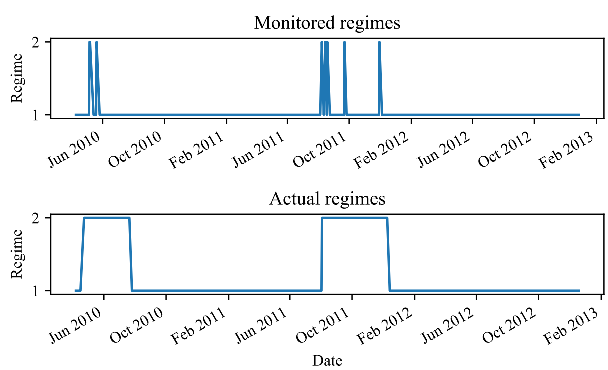

Fig. 4. Comparison of the monitored market regimes and the actual regimes. The monitored regimes are determined using the Equation 7.

The performance of B-Strict can be observed in Figure 4. Firstly, B-Strict was also able to detect both spells of Regime 2. In the first spell of Regime 2, the monitored regimes reported the first regime change on 6 May 2010, while the actual regime happened on 27 April 2010. Like B-Simple, there is a ten day delay in detecting regime change in B-Strict. As explained above, this is acceptable as the actual regime change was computed with the benefit of hindsight.

In the second spell of Regime 2, which appeared from 8 August 2011 to 14 December 2011, B-Strict reported the first regime change on 8 August 2011, which was spot on. This is a positive result.

Secondly, B-Strict also raised alarms repeatedly (as opposed to continuously) during the Regime 2 spells. B-Strict raised fewer alarms than B-Simple. As explained above, this does not prevent B-Strict from being useful to market participants.

Finally, with B-Strict, no false alarms were reported in the monitored period. The cautious approach by B-Strict may be preferred by some traders and those with oversight responsibilities in financial markets who need to react to the earliest alarms.

D. Discussion

Table II summarises and compares the performance of tracking under the two classification rules presented above. We focus on the rules' ability to pick up Regime 2 because market participants would benefit from alarms being registered when the market moved into a volatile regime, which is what Regime 2 represents.

Firstly, both B-Simple and B-Strict pick out regime changes as they happen. So they are useful to market participants. We compared the length of delay in raising an alarm of regime change under the two classification rules. In the first spell of Regime 2, both rules reported regime change after a ten days delay. However, in the second spell of Regime 2, B-Simple reported an alarm of regime change three days ahead of the actual regime change, while B-Strict reported a regime change on the spot. Both rules are doing their jobs properly, which is to alert traders of regime changes.

Secondly, the number of true alarms is compared between the two classification rules. A true alarm is an alarm raised when the market is actually in Regime 2. As shown in Table II, in both spells of Regime 2, more true alarms can be seen by B-Simple than B-Strict. In total, 45 alarms were raised by B-Simple while 12 of them were considered as false alarms. On the other hand, only eight alarms were raised by B-Strict, all of which were true alarms.

The number of false alarms generated by B-Simple is not excessive. So both the B-Simple and B-Strict are perfectly usable. Which rule a market participant might prefer depends on the market participant's attitude towards false alarms and how the signals are used.

TABLE II. COMPARISON OF B-SIMPLE AND B-STRICT IN THEIR DETECTION OF REGIME 2.

	B-Simple		B-Strict	
	1st spell	2nd spell	1st spell	2nd spell
Length of delay (days)	10	3 days ahead	10	0
Number of true alarms	5	28	3	5
Total alarms	45		8	
Total false alarms	12		0	

V. CONCLUSION

In [1] we presented a way to detect regime changes in hindsight. In [2] we showed that normal regimes share similar characteristics – in their normalised TMV and T values. In this paper, we have shown that results in the above two papers support a tracking mechanism. We have provided such a data-led mechanism to track regime changes dynamically. This is a practical method, as it uses data up to the present to monitor the likelihood of the market entering a volatile regime.

The proposed approach used TMV and T values observed in the two regimes in the past to establish a Bayes model. For each pair of (TMV, T) values observed in the current market, the Bayes model calculated two probabilities: one for the market being in Regime 1 (the "normal regime" in [2]) and the one for the market being in Regime 2, (the "abnormal regime" in terms of volatility). These two probabilities are used to decide which regime the market is in. Two classification rules were examined: a Simple Rule and a Stricter Rule. Combined with the Bayes model, the tracking systems are called B-Simple and B-Strict, respectively.

Dow Jones, FTSE 100 and S&P 500 data from 2007 to 2010 were used to build the Bayes model. This was used to track S&P 500 prices from 2010 to 2012. By using the method presented in [1], we concluded, with the benefit of hindsight, two spells of Regime 2 in the test period. Both B-Simple and B-Strict managed to pick up both spells, with ten days delay in the first spell and zero delay in the second. In our view, these results are very positive. The tracking signals could be useful to market participants. This work potentially lays the foundation for a financial early warning system, warning market participants of market instability, which influence the outcome of local, national and international financial markets.

However, this paper is a proof of concept, and thus part of a beginning of the research on this topic. It uses a naive Bayes model and two very simple classification rules for its proof. No doubt more experiments will need to be done and more advanced methods could be developed in the future to improve the reliability and usability of the tracking in future research.

REFERENCES

[1] E. P. Tsang and J. Chen, “Regime change detection using directional change indicators in the foreign exchange market to chart Brexit,” *IEEE Transactions in Emerging Technology in Computational Intelligence (TETCI)*, vol. 2, no. 3, pp. 185–193, 2018.

[2] ——, “Classification of normal and abnormal regimes in financial markets,” *submitted to Algorithms, Special Issue on “Algorithms in Computational Finance”*, 2018. [Online]. Available: available in http://www.bracil.net/download/Chen-NormalRegimes-Manuscript-2018.pdf

[3] 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.

[4] A. Sklarew, *Techniques of a professional commodity chart analyst*. Commodity Research Bureau, 1980.

[5] E. Tsang, “Directional changes, definitions,” *Working Paper WP050-10, Centre for Computational Finance and Economic Agents (CCFEA), University of Essex*, 2010.

[6] 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.

[7] E. Alpaydin, *Introduction to machine learning*. MIT press, 2014.