Why is Directional Change suitable for Handling Tick-to-tick Data?

Edward Tsang

Working Paper WP088-20,

Centre for Computational Finance and Economic Agents (CCFEA), University of Essex

12th October 2020

Abstract

Time Series (TS) records transactions in a market at fixed intervals. In contrast, Directional Change (DC) records transactions that represent significant price changes in the opposite direction in a trend, where "significance" is observer-dependent. In this paper, we argue that DC is particularly suitable for recording and analysing tick-to-tick data. Firstly, significant data points and high activities between sampling points will not be recorded in a TS. Secondly, as transactions take place at irregular times, but TS records transactions at fixed intervals, adjustments are required in the recording process, which may distort the records. Thirdly, DC is data-driven: every new transaction could potentially provide us with valuable information on the pulse of the market; for that reason, DC is more suitable for tracking tick-to-tick data for signals.

I. Introduction

Time and prices are recorded as transactions are made in a market. The record of all transactions is referred to as tick-to-tick data (TD). It is important to note that transactions take place irregularly: while many transactions may take place in one second, no transaction may take place in the next second. To facilitate analysis, these transactions are often summarized as time series (TS), where transactions are sampled at fixed intervals. For example, one could take the final transaction of every day to form the daily closing time series.

Olsen et al [7][9] introduced the concept of Directional Change (DC) as an alternative way to sample transactions: the idea is to let data determines when a transaction should be sampled. A transaction is only recorded when a significant price change in the opposite direction of the current trend has taken place, where "significance" is observer-dependent. This will be explained with an example in Section IV. DC has been applied to forecasting [2], market analysis [13][10], monitoring [6] and trading [8][3][4][15][1][6].

This paper aims to argue that DC is more suitable for handling TD than TS.

II. An example set of TD

An example set of TD is shown in Columns 2 and 3 of Table 1. These transactions are shown in Figure 1. It is important to remember that transactions take place irregularly in the market. In this example, nine transactions took place between 01:00 and 02:00. These transactions were (01:08, 98), (01:13, 105), …, (01:53, 100). There were no transactions between 03:00 and 05:00.

Table 1. Data used for explaining the relationship between TD, TS and DC

Tick-to-tick Data (TD)			Time Series (TS)		Directional Change (DC) with a 5% threshold				
Data Point	Time (mm:ss)	Price	Time (mm:ss)	Price	Time (mm:ss)	Extreme Point	Last Hi/Low	Change from last hi/low	DC Confirmation
1	00:00	100	00:00	100	00:00	100	100		
2	00:10	110			00:10	110	110		Up
3	00:40	106					110	-3.64%	
4	00:50	107					110	-2.73%	
5	01:08	98	01:00	107	01:08	98	98	-10.91%	Down
6	01:13	105			01:13	105	105	7.14%	Up
7	01:23	90					105	-14.29%	Down
8	01:28	92					105	-12.38%	
9	01:33	83			01:33	83	83	-20.95%	
10	01:38	98					83	18.07%	Up
11	01:43	95					83	14.46%	
12	01:48	104			01:48	104	104	25.30%	
13	01:53	100					104	-3.85%	
14	02:08	103	02:00	100			104	-0.96%	
15	02:38	101	03:00	101			104	-2.88%	
16	05:08	98	04:00	101	05:08	98	98	-5.77%	Down
17	05:38	100	05:00	101			98	2.04%	
18	06:08	104	06:00	100			98	6.12%	Up
19	06:28	106			06:28	106	106	8.16%	
20	06:43	102					106	-3.77%	
21	07:03	100	07:00	102			106	-5.66%	Down
22	07:33	95					106	-10.38%	
23	07:58	98					106	-7.55%	
24	08:28	90	08:00	98	08:28	90	90	-15.09%	
25	08:43	92					90	2.22%	
26	09:08	97	09:00	92			90	7.78%	Up
27	09:40	99					90	10.00%	
28	10:00	100	10:00	100	10:00	100	90	11.11%	

Table 2. Time Series generated from the TD in Table 1

	Time	Price	Returns
1	00:00	100	
2	01:00	107	7.00%
3	02:00	100	-6.54%
4	03:00	101	1.00%
5	04:00	101	0.00%
6	05:00	101	0.00%
7	06:00	100	-0.99%
8	07:00	102	2.00%
9	08:00	98	-3.92%
10	09:00	92	-6.12%
11	10:00	100	8.70%

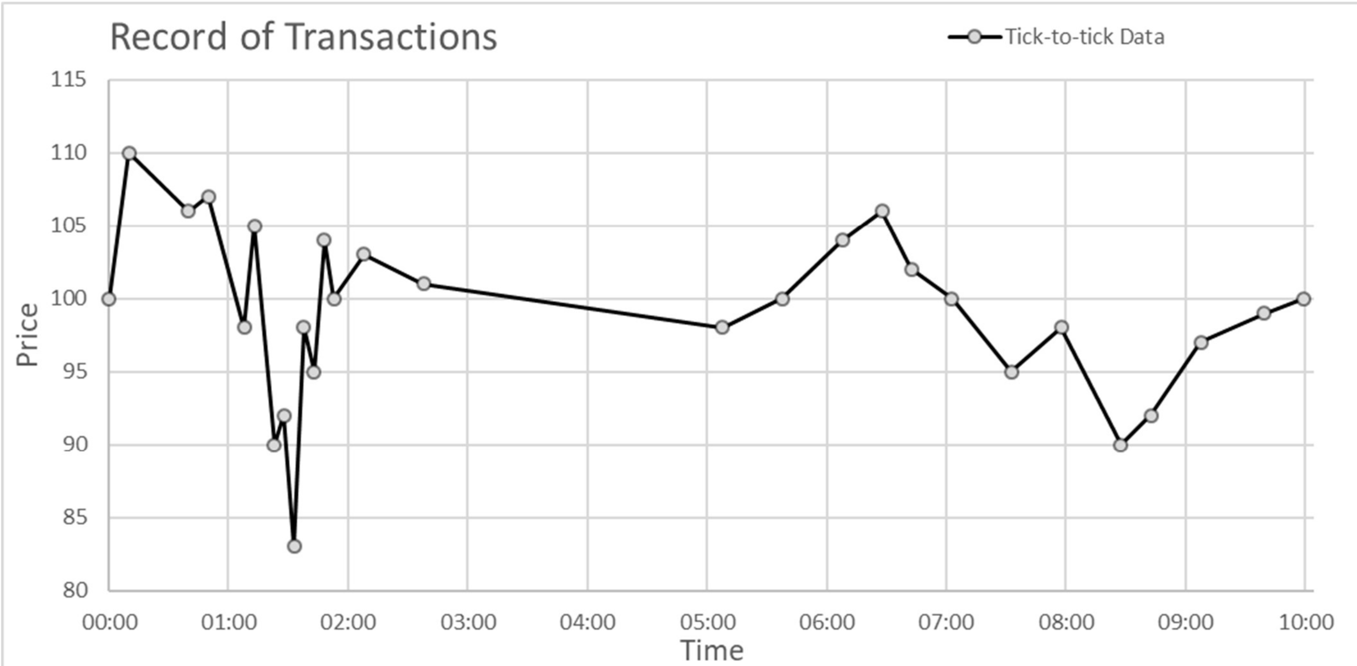

Figure 1. TD shown in Table 1

III. Recording transactions with TS

Given the TD shown in Columns 2 and 3 in Table 1, a time series is generated in Columns 4 and 5. This time series is summarised in Table 2 and plotted in Figure 2.

Observation 1. In TS, a data point is just an approximation of the nearest transaction times.

For example, TS records a price of 107 at time 01:00. This is a record of the transaction (00:50, 107). The price recorded for 03:00 is 101, which is the latest price recorded at 02:38.

Time approximation is not a problem if there are many transactions within each sampling interval. This will be the case in low-frequency data.

Observation 2. In TS, a data point must be created in the absence of transactions.

For example, as no transactions took place between 03:00 and 05:00, missing data points must be created in TS. One way to create a data point is to record the previous price. As the latest recorded price before 03:00 was 101, so this may be used as the price for 04:00 and 05:00.

Above is just one way to reconstruct missing data points. If we allow hindsight in creating the missing data points, we may reconstruct the missing data by assuming a linear decrease between (02:38, 101) and (05:08, 98).

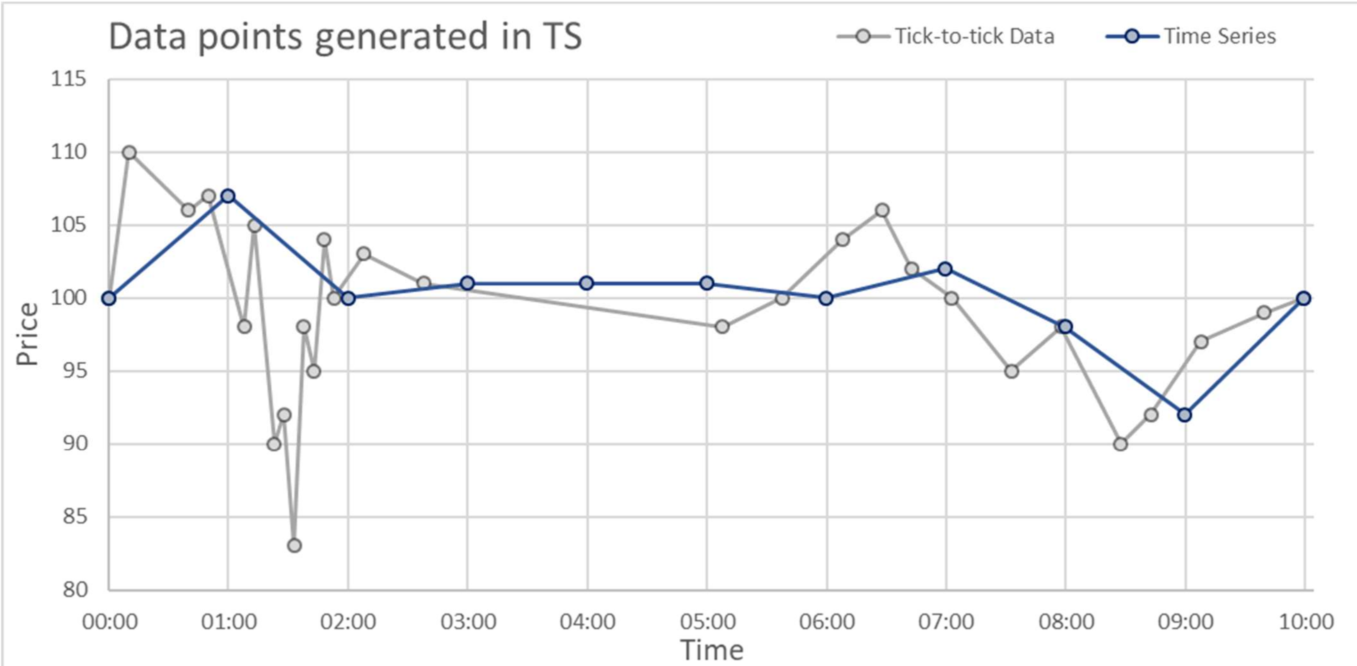

Figure 2. TD and TS shown in Table 1

In general, if we sample between large intervals, then the chance of missing data is low. The error of approximation is also proportionally reduced.[1] However, if we want to support high-frequency trading, we must sample frequently. The more frequently we sample, the more chance that we need to reconstruct missing data points.

IV. Recording transactions with DC

First, we shall explain how DC records transactions in a market. Columns 6 to 10 of Table 1 show DC summarises of the TD shown in Columns 2 and 3. DC records an extreme point (Column 7) when price reverses by a significant amount from the current trend. In Table 1, 5% is taken to be the threshold of a significant change. DC events are recorded in hindsight. At (01:08, 98), the first DC event is confirmed. This event started at (00:10, 110), which is called an extreme point. This DC event is confirmed because 98 is 10.91% (which is above the 5% threshold) below 110. Here (01:08, 98) is called a DC Confirmation (DCC) point. The next transaction (01:13, 105) confirms that the DCC point (01:08, 98) is itself an extreme point at which ends the downtrend and start the next uptrend. This is because 105 is 7.14% above 98.

A DC summary records the extreme points (Columns 6 and 7 in Table 1) and DCC points (Columns 8 and 9 in Table 1). The key points recorded by DC are summarised in Table 3. The extreme points are plotted in Figure 3. It is worth clarifying that the first (00:00, 100) and final points (10:00, 100) in Table 3 and Figure 3 are not extreme points in DC. They are plotted to show the first (00:10, 110) and the final (08:28, 90) extreme points in this summary.

[1] For example, for a time series with daily closing prices, even if the data point is one minute after the actual transaction, the error is only 0.07% of the 24 hours period. However, in a minutely time series, an approximation by 1 second incurs an error of 1.7%.

These extreme points shown in Figure 3 correspond to the "zig-zag indicator" in technical analysis [12]. The formal definition of DC[2] is presented in [6] (Appendix A).

Observation 3. DC records transactions as they are. Unlike TS, no adjustment is needed for data.

By definition, every extreme point and DC confirmation point recorded in a DC summary is a transaction in the TD.

Table 3. A 5% DC Summary generated from the TD in Table 1

Extreme Points		DC Confirmation Points		Trends started
Time	Price	Time	Price	
00:00	100	← For reference; this is not really an extreme point		
00:10	110	01:08	98	Down
01:08	98	01:13	105	Up
01:13	105	01:23	90	Down
01:33	83	01:38	98	Up
01:48	104	05:08	98	Down
05:08	98	06:08	104	Up
06:28	106	07:03	100	Down
08:28	90	09:08	97	Up
10:00	100	← For reference; this is not really an extreme point		

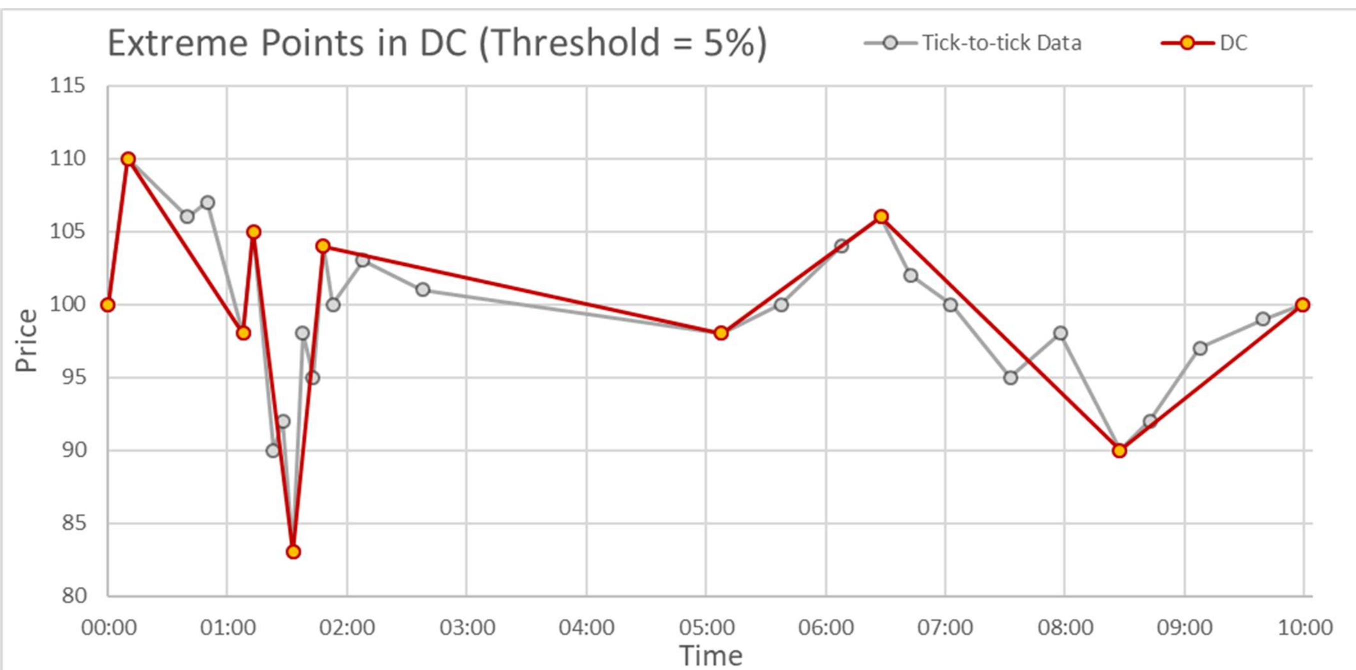

Figure 3. TD and DC extreme points in Table 1

[2] The formal definition of DC is mutually recursive. For simplicity, this is not elaborated in this paper.

V. Contrasting TS and DC

Figure 4 highlights the difference between the TS and DC summaries. As we shall explain, although there are 11 data points in the TS and only 8 extreme points in DC, the latter does not capture less information than the former.

Observation 4. DC records all the extreme points which could be missed by TS sampling.

By nature, a DC summary records all the extreme points. In the TS, the lowest price (01:33, 83) is not recorded.

Observation 5. TS may miss high activities in the market, which will be captured by DC.

By recording at fixed intervals, TS may not be able to record high volatility in the market. For example, the market fluctuated between 01:00 and 02:00. As DC is data-driven, DC will record all significant changes that took place. Between 01:00 and 02:00, DC records four extreme points.

Remarks on sampling frequency under TS:

By sampling more frequently under TS, one could increase the chance of capturing extreme points. However, by increasing the sampling frequency under TS, the errors in the approximations could potentially increase. For example, if one samples minutely, the last data point that is five seconds before the sampling time would incur an error of (2 ÷ 60 =) 8.3%. Besides, potentially more missing data will need filling.

Does one have the same dilemma in DC? What threshold should one use? By using different thresholds, one sees the market in different levels of details. By and large, the market exhibits similar properties under different thresholds.

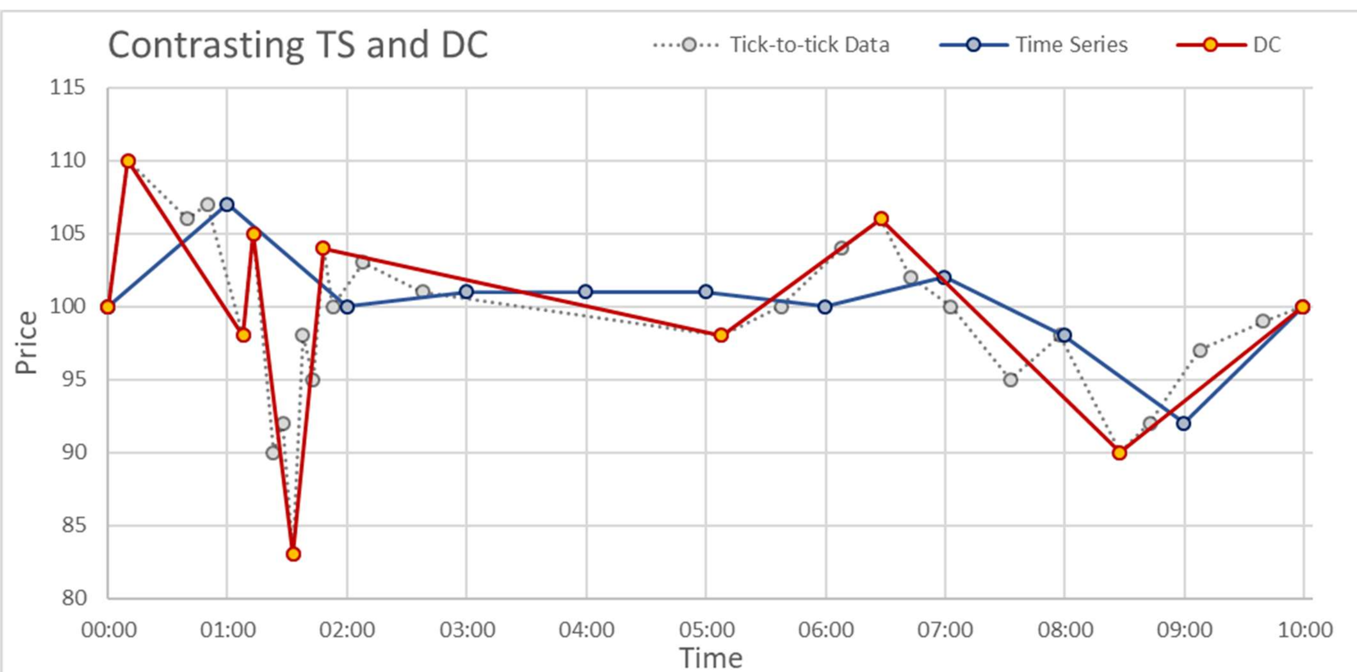

Figure 4. Contrasting TS and DC in recording the same TD

VI. DC can be used to track the market with every transaction

Volatility in TS is measured by the standard deviation of returns over a period. For example, in TS, the final three data points were (08:00, 98), (09:00, 92), (10:00, 100). While the two returns (-6.1% and +8.7%) are big, taking their standard deviation is not a very meaningful measure of the volatility of the market. Risk in TS is normally measured with a reasonably large number of returns.

Tsang et al [13] introduced several measures of risk in DC. Two of them are adapted here for our discussion:

Definition 1. Absolute Total Movement (**aTMV**) in DC:

$$aTMV = (|P_c - P_{EP}| \div P_{EP}) \div Threshold$$

Where P_c is the current price, P_{EP} is the preceding extreme price, Threshold is the threshold used to determine significance in the DC summary. By normalisation with the Threshold, the aTMV values obtained in different DC summaries can be compared, even if they were derived from different thresholds.

In the above example, the final extreme point in the DC summary in Table 1 is (08:28, 90). At (09:08, 97) (data point 26 in Table 1), the aTMV is ((|97 – 90| ÷ 90) ÷ 5% =) 1.556. At the next transaction point (09:40, 99), the aTMV is ((|99 – 90| ÷ 90) ÷ 5% =) 2. At (10:00, 100), the aTMV is 2.222.

Observation 6. The aTMV can be calculated after each transaction. This gives us a transaction-by-transaction measure of the volatility in the market in DC, which is not available under TS.

For example, the extreme point (01:13, 105) is immediately followed by the transaction (01:23, 90) (TD items 6 and 7 in Table 1). The aTMV value at (01:23, 90) is (|90 – 105| ÷ 105) ÷ 5% =) 2.8571. Olsen et al [5][9] discovered that, on average, price reverse when the price reaches twice the threshold; this means, on average, DC events take place at aTMV equals to 2. The aTMV at (01:33, 83) is 4.1905, which indicates an increased chance of price reversion.

We also know from historical data that aTMV values roughly follow a power-law distribution [14]: the chance of a DC event happening increases exponentially as aTMV increases. Being able to monitor every transaction allows one to compare the current aTMV with the historical aTMV distribution. A transaction-by-transaction measure of aTMV allows us to estimate the chance of a directional change after the current transaction, according to historical data.

On the other hand, the high activities between 01:00 and 02:00 are not recorded by TS (Observation 5). TS does not identify extreme points. Therefore, the data points that TS chooses to record does not allow it to monitor the market transaction-by-transaction that follow the sampled data points.

Definition 2. Absolute Return **aR** in DC:

$$aR = (|P_c - P_{EP}| \div P_{EP}) \div (T_c - T_{EP})$$

Where T_c and P_c are the time and price of the current transaction, T_{EP} and P_{EP} are the time and price of the preceding extreme point.

For example, the final extreme point was (08:28, 90) (TD item 24 in Table 1). At the DCC point (09:08, 97) (TD item 26), aR is (|97 – 90| ÷ (09:08 – 08:28) =) 11.67% per minute[3]. At the next transaction point (09:40, 99), aR is (|99 – 90| ÷ (09:40 – 08:28) =) 8.33%.

Observation 7. The absolute return aR can be calculated after each transaction. This gives us a transaction-by-transaction measure of the return since the previous extreme point in DC. This measure is not available under TS.

If we look at the transaction (01:23, 90) (TD item 7 in Table 1), we can see that it is a sharp drop from the preceding extreme point (01:13, 105). The aR value is (|90 – 105| ÷ (01:23 – 01:13) =) 85.71%. Unlike aTMV, aR takes time into consideration. While the aTMV (2.8571) mentioned above may not be big enough to cause any concern[4], this aR value is alarmingly large. When such a big drop is found in the market, traders and regulators should be alerted.

Can we monitor the price change and return at each transaction-by-transaction in TS? Yes, we can monitor the price change and return from the last sampled point. However, the last sampled point may not (and most likely will not) be the lowest or highest point in the current trend. So these measures under TS do not bear the same significance as TMV and R in DC. Chen & Tsang [6] showed how regime changes in the market can be tracked under DC.

VII. Concluding Summary

We have used an example to compare and contrast how we summarise price changes under TS and DC. We summarise our observations below:

1. Data points in a TS approximate transaction times (Observation 1), whereas DC records transactions as they are (Observation 3).
2. As transactions take place at irregular times in TD, reconstruction of missing data may be needed in TS (Observation 2).
3. All extreme points are recorded in DC summaries; they may be missed by TS sampling (Observation 4).
4. High activities in the market captured by DC may be missed in TS (Observation 5).
5. To increase the chance of capturing extreme points and high activities in TS, one could choose to increase the frequency of sampling. But doing so increases the error in approximations (Observation 1) and the need to fill in missing data (Observation 2).
6. DC is more sensitive to price movements, hence more suitable for market monitoring: Under DC, a single transaction may provide us with valuable information about the market. With aTMV (Definition 1) and aR (Definition 2), one could monitor the price movement transaction-by-transaction. The same cannot be done in TS, as it does not record the extreme points (Observation 6 and Observation 7).

For these reasons, we argue that DC is more suitable for recording and analysing TD.

Acknowledgements

Thanks to Richard Olsen for introducing the concept of Directional Change to me in mid-2000. My research took a directional change since. This paper started in 2013 through discussions with Ao Han.

---

[3] The returns in this artificial data set are large; this is designed so for illustration purpose.

[4] This value (2.8571) is not much bigger than the average aTMV at extreme points (which is 2, according to [5][9]).

References

[1] Ao, H. & Tsang, E.P.K., Trading algorithms built with Direction Changes, IEEE Conference on Computational Intelligence for Financial Engineering and Economics (CIFEr) 2019 Conference, Shenzhen, China, 4th-5th May 2019

[2] Bakhach, A., Tsang, E.P.K. & Jalalian, H., Forecasting Directional Changes in FX Markets, IEEE Symposium on Computational Intelligence for Financial Engineering & Economics (IEEE CIFEr'16), Athens, Greece, 6-9 December 2016

[3] Bakhach, A., Raju Chinthalapati, V.L., Tsang, E.P.K. & El Sayed, A.R., Intelligent Dynamic Backlash Agent: a trading strategy based on the directional change framework, Algorithms, Special Issue on Algorithms in Computational Finance, MDPI Open Access Publishing (ISSN 19994893), 11(11), 2018

[4] Bakhach, A., Tsang, E.P.K. & Raju Chinthalapati, V.L., TSFDC: A trading strategy based on forecasting directional change, Intelligent Systems in Accounting, Finance and Management, Vol.25, Issue 3, May 2018

[5] Bisig, T., Dupuis, A., Impagliazzo, V & Olsen, R.B., The scale of market quakes, Quantitative Finance Vol.12, No.4, 2012, 501-508

[6] Chen, J. & Tsang, E.P.K., Detecting Regime Change in Computational Finance, Data Science, Machine Learning and Algorithmic Trading, CRC Press, September 2020

[7] Dacorogna, M.M., Gencay, R., Muller, U., Olsen, R.B. & Picktet, O.V., An introduction to high-frequency finance, Academic Press 2001

[8] Golub, A., Glattfelder, J. & Olsen, R.B., The Alpha Engine: Designing an Automated Trading Algorithm, High-Performance Computing in Finance, Chapman & Hall/CRC Series in Mathematical Finance, April 2017

[9] Glattfelder, J.B., Dupuis, A. & Olsen, R.B., Patterns in high-frequency FX data: discovery of 12 empirical scaling laws, Quantitative Finance, Volume 11 (4), 2011, 599-614

[10] Ma, J., Xiong, X., He, F. & Zhang, W., Volatility measurement with directional change in Chinese stock market: Statistical property and investment strategy, Physica A: Statistical Mechanics and its Applications, 2017, vol. 471, issue C, 169-180

[11] Ma, S. & Tsang, E.P.K., Guidelines for determining the useable range of thresholds for Directional Change profiling, Working paper WP081-19, Centre for Computational Finance and Economic Agents (CCFEA), University of Essex, 11 August 2019

[12] Sklarew, A., Techniques of a Professional Commodity Chart Analyst. New York: Commodity Research Bureau, 1980

[13] Tsang, E.P.K., Tao, R., Serguieva, A. & Ma, S., Profiling High-Frequency Equity Price Movements in Directional Changes, Quantitative Finance, Vol.17, Issue 2, 2017, 217-225

[14] Tsang, E.P.K. & Ma, M., Distribution of aTMV, an empirical study, Working Paper WP089-20, Centre for Computational Finance and Economic Agents (CCFEA), University of Essex, September - October 2020 (to appear)

[15] Ye, A., Chinthalapati, V.L.R., Serguieva, A. & Tsang, E., Developing sustainable trading strategies using directional changes with high-frequency data, IEEE International Conference on Big Data, Boston, 11-14 December 2017