STATISTICS IN TRANSITION new series, March 2025
Vol. 26, No. 1, pp. 155–181, https://doi.org/10.59139/stattrans-2025-009
Received – 13.02.2024; accepted – 04.12.2024

Improving detectability of the indicator saturation approach through winsorization: an empirical study in the cryptocurrency market

Suleiman Dahir Mohamed[1], Mohd Tahir Ismail[2], Majid Khan Bin Majahar Ali[3]

Abstract

Despite the introduction of several adjustments, mitigating data anomalies in financial datasets has proven challenging, particularly in the context of cryptocurrencies with extreme values and increased volatility. The progress in properly addressing these anomalies prior to testing remains restricted, highlighting the unique and complex nature of financial data in this domain. Thus, in this paper we propose a hybrid approach called the Win-IS strategy. It is meant to address the influence of extreme outliers in the tail and subsequently identify breaks, trend breaks and outliers in cryptocurrencies. This methodology uses the winsorization (Win) process to enhance the effectiveness of the indicator saturation (IS) approach. The study uses cryptocurrencies like Bitcoin (BTC), Ethereum (ETH), Litecoin (LTC), Tether (USDT), and Ripple (XRP). The results of the research indicate that the winsorization strategy improved the detectability of the IS approach, with Win-IS outperforming the IS method in terms of the Bayesian Information Criterion. Furthermore, the Win-IS technique uncovered additional breaks, trend breaks and outliers that were previously unknown and repeated in some cases as detected by the IS strategy. The effect of winsorization is dependent on the chosen percentile and dataset attributes. Through detailed examination and comparison, the findings of this research contribute to the improvement of other detection approaches, providing a valuable perspective for researchers and practitioners in the field. Additionally, this hybrid approach can improve decision-making, risk management and model creation, benefiting investors, legislators and scholars.

Key words: breaks, outliers, winsorization, indicator saturation, cryptocurrency.
JEL Classification: C22, C58, C61, G23, G32.

[1] School of Mathematical Sciences, Universiti Sains Malaysia, Minden, Pulau Pinang, 11800, Malaysia. E-mail: suleimandm2017@gmail.com. ORCID: http://orcid.org/0000-0002-3368-204X.

[2] Corresponding Author. School of Mathematical Sciences, Universiti Sains Malaysia, Minden, Pulau Pinang, 11800, Malaysia. E-mail: m.tahir@usm.my. ORCID: http://orcid.org/0000-0003-2747-054X.

[3] School of Mathematical Sciences, Universiti Sains Malaysia, Minden, Pulau Pinang, 11800, Malaysia. E-mail: majidkhanmajaharali@usm.my. ORCID: http://orcid.org/0000-0002-5558-5929.

1. Introduction

A structural break is an abrupt change in a time series of data. The structural break and outliers are an important aspect to consider in economics and statistics since they are unexpected. A structural break denotes a change in the behavior of a variable over time, such as a rise in the money stock, or a shift in a previously observed link between observable variables, such as inflation and unemployment, or the balance of trade and the exchange rate (Brooks, 2019). Outliers are data points that deviate from the norm (Hawkins, 1980). Extreme values can have a significant impact on the performance of statistical tests. As a result, correctly identifying changepoints in time series data becomes challenging when working with big samples including a large number of extreme values (outliers), which can either coincide with or disguise major shifts (breaks). Aside from the masking effect, recognizing and correctly identifying breaks and outliers concurrently is another significant challenge. According to Mulry et al. (2014), correctly detecting breaks and outliers is crucial for making educated investment decisions, managing risk, and maintaining the accuracy of financial analysis. Particularly when dealing with financial data, such as cryptocurrency, which is known to undergo major changes as a result of external factors such as wars, natural disasters, etc. (Chatzikonstanti, 2017). Satoshi Nakamoto, the alias for an anonymous computer programmer or group of programmers, invented the first cryptocurrency on January 3, 2009, when the Bitcoin software was made public. Later, numerous additional coins appeared. Cryptocurrencies have emerged as a disruptive digital asset class that has the potential to disrupt traditional financial systems. As these digital assets gain popularity, the need for reliable and effective solutions to monitor and manage data fluctuations grows.

So far, a variety of statistical methodologies have been used to identify breaks and anomalies within cryptocurrency datasets, including single change detection approaches such as Chow (1960) and Quandt (1960), two change detection approaches such as Papell and Prodan (2003), and multiple change detection approaches such as BP of Bai and Perron (1998, 2003), as well as the Iterative Cumulative Sum of Squares (ICSS) approach developed by Iclan and Tiao (1994). In addition, researchers used these strategies separately for break or outlier detection. Chatzikonstanti (2017) used the wavelet approach to handle outliers and the CUSUM approach to find breaks. Mandaci and Cagli (2022) used Bai and Perron to determine the frequency of cryptocurrency breaks. Yen et al. (2022) used the BP approach to analyze ten cryptocurrencies and found structural breaks in return, price, and squared return. Sahoo (2021) employed the Narayan and Popp (2010) endogenous two structural breakdowns unit root test to identify breaks in bitcoin returns. Canh et al. (2019) applied the Wald test and discovered structural fractures in all well-known cryptocurrencies. Thies and Molnár (2018) used Bayesian change point (BCP) analysis to examine the presence of many segments in the Bitcoin return distribution and demonstrated evidence of structural breaks in the first and second moments of the return distribution. Dutta and Bouri (2022) used Ane et al. (2008) technique and found no indication that any of the top cryptocurrencies have outliers except the Bitcoin return series. Kaseke et al. (2022) used the Pruned Exact Linear Time (PELT) method to identify breakpoints in the cryptocurrency market, which tests for changepoints in the mean, variance, and both mean/variance of the series. Abakah et al. (2020) used Bai and Perron approach and its extension to the fractional case and discovered the existence of breaks in the cryptocurrency market. Aharon (2023) found breaks in cryptocurrency by employing modified ICSS technique. Jiang and Yoon (2023) used BP and ICSS to detect cryptocurrency breaks.

However, many of these traditional techniques may struggle in situations when outliers exist (Fearnhead and Rigaill, 2019). Rodrigues and Rubia (2011) demonstrate that outliers can conceal the presence of structural breaks. Thus, the challenge is to determine which magnitude can be classified as a break or an outlier. The topic of distinguishing between changepoints and outliers has gotten very little consideration. As a result, when evaluating data for structural changes, outliers can frequently obscure major trends, yielding incomplete or misleading conclusions. Traditional approaches may not properly discriminate between genuine changes and those disguised by extreme values.

To solve this problem, Hendry (1999) suggested the indicator saturation strategy known as the IS approach. So far, this technology has been able to detect various data patterns such as breaks and outliers in the data simultaneously. However, because the IS technique can detect many data patterns at the same time, a masking effect may occur, as previously discussed. Specifically, when the IS technique is used in very high frequency datasets such as rapidly fluctuating markets like cryptocurrencies. Therefore, this study uses the IS technique to first identify and record the dates of the breaks, trend breaks, and outliers simultaneously in five distinct cryptocurrency log returns. Second, by ignoring the outcome of the first objective, lessens the impact of extreme observations in each data set using Winsorization technique. Thirdly, employs the IS approach to simultaneously re-identify and record the dates of outliers, trend breaks, and breaks in the Winsorized log returns of each cryptocurrency. Fourth, detects presence of masking effect by comparing and categorizing results into repeating and emerging changes.

The study is driven by the fact that outliers can influence data analysis, masking actual structural changes that are crucial for accurate interpretation. By tackling outliers first, we can uncover hidden breaks, resulting in a better comprehension of the data and more informed decision-making. When applying the IS technique to high-frequency data, a caution is to be exercised since outliers and rapid changes in the data may generate a masking effect, and high-frequency data with rapid ups and downs may confound the IS approach's automatic detection system. This paper presents a methodological framework for integrating the winsorization strategy into the indicator saturation approach. The hybrid technique might be referred to as the Win-IS approach. The hybrid technique identifies Hidden Patterns that are disguised by outliers, improving the accuracy of break identification. This method may improve risk management and assist investors, traders, and financial analysts in making sound judgments in the ever-changing world of digital assets and other financial markets. The paper is organized as follows. Section 2 contains the body of the current literature; Section 3 describes the approach used; Section 4 shows the results and the discussion; and Section 5 offers the conclusion.

2. Literature review

Literature documented various approaches for detecting data breaks and outliers. Most known break detection methods are based on regression. Chow (1960) pioneered structural break testing for regression models, developing the F-test for a single break, assuming that the break date is previously known under the null of no break. Quandt (1960) altered the Chow framework to consider the F-statistic with the highest value among all potential break dates to loosen the requirement that the candidate break date be known. Later research revealed the assumption of prior knowledge of the break dates, which expanded on previous experiments to allow for multiple breaks, particularly the Bai and Perron tests (Bai and Perron, 1998, 2003). The Bai and Perron approach is limited to trimming 15% of the data and a maximum of 5 breaks. Ohara (1999) also used a method based on Zivot and Andrews (2002) sequential t-tests to analyze the case with m breaks with ambiguous break dates. Papell and Prodan (2003) developed a test based on restricted structural change that explicitly enables two offsetting structural modifications. Detection of breaks in the case of variance has also been investigated using Iclan and Tiao (1994) iterative cumulative sum of squares (ICSS) approach.

Two popular ways to deal with outliers are trimming and winsorization (Moir, 1988). Winsorizing, also known as Winsorization, is a statistical transformation technique that limits extreme values in statistical data to lessen the impact of potentially inaccurate outliers. It is named after the engineer-turned-biostatistician Charles P. Winsor (1895–1951). According to Tukey (1962) when Winsor discovered an outlier in a sample, he did not just discard it; instead, he altered its value. Winsor proposed utilizing the size of the next greatest (or smallest) observation to estimate the magnitude of an extreme, poorly known, or unknown observation. According to Xao et al. (2014)

Winsorization is another reliable method for handling non-normal distributions to prevent information loss and preserve the original sample size. In another classification, the most common outlier treatments in finance are winsorizing, trimming, and dropping (Adams et al., 2019). To lessen the influence of the outlying points, robust solutions based on Winsorization are frequently used (Cheng & Young, 2023). Winsorization method provides adjustments for the observed influential value and winsorization processes can be one-sided or two-sided (Mulry et al., 2014). However, determining the cutoff sites is a vital part of these approaches (Cheng and Young 2023). The more the data is winsorized, the bigger the bias in the coefficient estimates for variables (Lien et al., 2005). Adams et al. (2019) extended the winsorization approach into multivariate case. Hamadani and Ganai (2023) cleaned and processed their data using the winsorization approach. Li et al. (2021) combined the change detection method and the Winsorization method into the prediction model based on the autoregressive moving average model.

On the other hand, Hendry (1999) suggested a strategy called indicator saturation (IS). The indicator saturation methodology employs an automatic multi-path search strategy that can handle more candidate variables N than observations T, separates variables into blocks, and records significant ones using impulse-indicator saturation (Castle et al., 2011). According to Pretis et al. (2017), indicator saturation offers an alternate technique based on an expanded general-to-specific methodology based on model selection. Starting with a full set of indicators and discarding all except the most significant ones, structural breaks can be identified without specifying a minimum break length, maximum break number, or imposed co-breaking. Hendry et al. (2008) demonstrated that different numbers of splits and uneven splits have no effect on the retention rate. According to Castle (2022), numerous indicator saturation estimators (ISEs) are available to model a wide range of non-stationarity phenomena. However, each ISE is created to address a particular issue. For example, the step indicator saturation (SIS) of Castle et al. (2015) for location shifts, the trend indicator saturation (TIS) of Pretis et al. (2015) for trend breaks, and the impulse indicator saturation (IIS) of Hendry et al. (2008) and Johansen and Nielsen (2016) for outliers. Pretis et al. (2018) developed an algorithm for the indicator saturation approach in general to specific modelling (Gets), which provides a straightforward computation of this approach. This method has been applied in a variety of fields in the literature. Marczak and Proietti (2016) used IIS and SIS in the framework of structural time series models. Ghouse (2021) employed IIS to discover structural breaks in Pakistan Islamic banks data. Pretis et al. (2015) used TIS and SIS to assess climate models. Castle et al. (2021) utilized TIS and SIS in identifying shifts in trends within a long-term UK production function. Ghouse et al. (2022) utilized the IIS technique to identify the structural breaks in the returns and volatility of commercial banks in Pakistan. Muhammadullah (2022) applied IIS to detect outliers in the cross-sectional analysis estimated through the application of regularization techniques with COVID-19 data. Panday (2015) utilized IIS to investigate the influence of monetary policy on the exchange rate of Nepal. Castle et al. (2012) employed U.S. real interest rates to assess the effectiveness of IIS in comparison to the BP approach and found that IIS successfully reproduces the results of the BP. Ismail and Nasir (2018) conducted a comparison between IIS and BP in ASEAN sharia-compliant indices and found almost identical results. However, IIS can identify significant breaks and outliers that occur at the start and end of a sample, while BP necessitates a designated percentage of the sample for analysis. Mohamed et al. (2023) conducted an empirical study to compare BP and IS in detecting cryptocurrency breaks and outliers and found that the IS approach produce more.

Structural break tests allow us to assess when and whether there is a major change in data. The indicator saturation approach uses regression model and identifies non stationarity shifts at any point and location. According to Choi (2009), empirical results can be misleading when researchers disregard abnormal observations, particularly when it comes to dependent variables. The study assumes that the winsorization approach helps the IS approach in lessening the extreme values. Brownen (2019) and Afanasyev et al. (2019) investigated how the winsorization technique influences the performance of regression models and discovered three factors: the level of data inaccuracies in the tails, the characteristics of enterprises affected by the process, and the usage of scaling. Moir (1988) found that when the distribution is non-normal, winsorization is suggested as an alternative to trimming. Rivest (1994) also recommended the use of winsorization for skewed distributions. Dixon (1960) claims that maintaining symmetric winsorizaiton and making suitable changes will result in improved estimators. Sharma and Chatterjee (2021) discovered that Winsorization is a versatile strategy for compensating for data outliers. Yuliyani and Indahwati (2017) used winsorization in linear mixed models to detect violations of the normalcy assumption in national exam data. In this study, just 1% of the observations are winsorized in order to prevent winsorization from affecting all observations. Adams et al. (2019) found that the estimation of trimmed or winsorized least squares can still be influenced (possibly dramatically) by a single lingering outlier

Finally, the study combines two existing methodologies. The two approaches are Hendry's indicator saturation technique (1999) and Charles P. Winsor's winsorization strategy (1895-1951). There is a huge knowledge gap regarding the combination of these two approaches to increase the detectability of the indicator saturation strategy. This study addresses this gap by offering a hybrid strategy that combines the winsorization and indicator saturation approaches. The IS strategy is responsible for recognizing outliers, breaks, and trend breaks all at once, whereas winsroization aids the IS technique in reducing and mitigating extreme outliers.

3. Methodology

3.1. Datasets

The data used in this study consists of five different cryptocurrencies: Tether (USDT), Litecoin (LTC), Ripple (XRP), Ethereum (ETH), and Bitcoin (BTC). The data was obtained from https://finance.yahoo.com/. All prices are through June 30, 2023. The total number of observations for BTC, which began on November 22, 2014, LTC, which began on September 22, 2014, and XRP, ETH, and USDT, which began on November 13, 2017, are 3143, 3204, and 2056, respectively. The study employed price log-returns calculated using the following formula:

$$\mathrm{r_t} = \ln\left(\frac{\mathrm{P_t}}{\mathrm{P_{t-1}}}\right) = \ln(\mathrm{P_t}) - \ln(\mathrm{P_{t-1}}) \tag{1}$$

In this case, r_t represents returns, P_t is the lag price at time t, and P_{t-1} is the lag price from time $t-1$.

3.2. Theoretical framework

3.2.1. Winsorization Approach

The Winsorization approach is a robust test that targets reducing the impact of outliers using certain percentile values and dealing with non-normality. This method preserves the original sample size by replacing the tail of the data rather than removing it. The default percentiles of the winsorization approach are the 5th and 95th percentiles. Our strategy assigns the extreme values to a very tight percentile of returns. So, the study uses a 99% winsorization that would assign all returns below the 1st percentile to the 1st percentile value and data above the 99th percentile to the 99th percentile value. Mathematically speaking, given a sample of T observations, Winsorizing entails replacing the k highest values with the $k - 1$ value and the l lowest values with the $l - 1$ value, then calculating the desired statistic on the T values. For example, consider the situation of BTC with $T = 3142$ observed ordered values and $k = l = 32$, where the winsorized values are the $k + l = 32 + 32 = 64$ substituted values.

The winsorized vector of returns is achieved by

$$w(r_t) = \begin{cases} -v & \text{For } r_t \leq -v \\ r_t & \text{For } |r_t| < v \\ v & \text{For } r_t \geq v \end{cases}$$

where $w(r_t)$ represents the winsorized vector, r_t represents the original returns of each cryptocurrency, v is the $k-1$ observation (nearest 99th percentile) and $-v$ is the $l-1$ observation (nearest 1st percentile).

To winsorize the returns of every cryptocurrency, the following steps have been taken. The process is based on Bitcoin, however, Table 1 shows the outcomes of applying the same process to other currencies:

1. Order each cryptocurrency data in ascending order.
2. Decide length of k and l.
3. For the case of BTC, the length of l can be found by the 1st percentile × sample size $(0.01 \times 3142 \approx 32)$, while the length of k is found by 99th percentile × sample size $(0.99 \times 3142 - 3142 \approx 3110)$.
4. In this case we have symmetric values. So, $k = l = 32$ matching 32 observations less than 33rd observation and another 32 observations greater than 3111th observation.
5. Pick the $-v$ value which is 33rd observation and v value which is 3111th observation.
6. Replace Values Greater than the 99th percentile $(k\ values)$. So, all values greater than the 99th percentile are replaced with the v.
7. Replace Values Less than the 1st percentile. So, all values less than the 1st percentile are replaced with the $-v$.

Table 1: The framework of winsorized observations

Series	Sample size (T)	Winsorized observations	minVal($-v$)	maxVal($-v$)
BTCW	3142	$k = 32, l = 32, k + l = 64$	-0.11092	0.104009
LCTW	3203	$k = 32, l = 32, k + l = 64$	-0.14447	0.169612
ETHW	2055	$k = 20, l = 20, k + l = 40$	-0.14658	0.127828
USDTW	2055	$k = 20, l = 20, k + l = 40$	-0.01277	0.013738
XRPW	2055	$k = 20, l = 20, k + l = 40$	-0.15583	0.200854

Note: symbol BTCW stands for winsorized Bitcoin return.

3.2.2. Indicator Saturation Approach

Hendry (1999) introduced the IS approach. Pretis et al. (2018) state that the IS technique was created in order to identify and model outliers as well as structural breaks in the mean. Several IS estimators (ISEs) can be utilized to model distinct aspects of wide-sense non-stationarity (Castle & Hendry, 2022). Step-indicator saturation (SIS) for location shifts, impulse indicator saturation (IIS) for outliers, and trend indicator saturation (TIS) for trend breaks are a few examples. IIS is a reliable and effective statistical method, according to Hendry (1999) and Johansen et al. (2009). Both Castle et al. (2015) and Pretis et al. (2015) broadened the definition of IIS to encompass SIS and TIS.

The IS method is based on a regression model that uses a general-to-specific modeling strategy to produce indicator variables for every observation. A different approach that uses a general-to-specific procedure based on model selection is provided by IS, claim Pretis et al. (2018). In other words, a regression model is overloaded with indicators, which are then chosen using the general-to-specific at a predefined level of significance. All indicators that are not significant are then removed without imposing co-breaking or defining a minimum or maximum break segment. By starting with a general model (the GUM) and lowering variables along search paths while assessing the diagnostics at each stage, the general-to-specific technique (GETS) offers an organized search. Marczak and Proietti (2016) state that IS has been shown to be effective and useful when used with a dynamic regression model. This IS method was developed by Pretis et al. (2018) using the GETS package in the R programming language. Three IS estimators—SIS, TIS, and IIS—are used in this study to identify breaks, trend breaks, and outliers. The following is the formulation of these estimators:

$$\text{IIS } y_t = \mu + \sum_{j=1}^{n} \delta_j 1_{\{t=j\}} + \varepsilon_t \tag{2}$$

$$\text{SIS } y_t = \mu + \sum_{j=2}^{n} \delta_j 1_{\{t\geq j\}} + \varepsilon_t \tag{3}$$

$$\text{TIS } y_t = \mu + \sum_{j=1}^{n} \delta_j 1_{\{t>j\}}(t-j) + \varepsilon_t \tag{4}$$

A break, trend break, or outlier's size is denoted by δ, errors are represented by ε, the BTC return over time is represented by y_t, and the constant term is denoted by μ. We regressed with the constant and used y_t as a dependent variable in order to apply the IS technique. According to suggestion by Ismail and Nasir (2018) we set alpha value that is based on the sample size, $\alpha = 1/T$. With an alpha value determined by sample size, the three equations were executed concurrently. The tight alpha value allows us to limit the number of significant indicators (dates); however, larger alpha values can be allowed if the interest is to increase number of changepoints.

During the application, the IS approach creates dummy variables automatically when the algorithm is executed. The IS approach will first create dummy variables representing each estimator that is equal to the number of observations in the returns. The BTC dataset generates 9423 indicators when three IS estimators (IIS, SIS, and TIS) are performed concurrently, divided into 105 blocks of 30 indicators each. Table 2 shows details of the dummy variables and their blocks.

Table 2: Dummy variables

Returns	Sample size	Alpha	Dummy Variables	
			Indicators	Blocks
BTC	3142	0.0003	9423	105
LCT	3203	0.0003	9606	107
ETH	2055	0.0005	6162	69
USDT	2055	0.0005	6162	69
XRP	2055	0.0005	6162	69

3.2.3. Incorporating Winsorization into IS approach

The incorporation begins by winsorizing the returns of each digital coin as stated. Then, each winsorized return is considered as the dependent variable. Since the IS approach is based on regression model, we regress a constant to the winsorized returns (dependent variable). Then, the IS approach automatically allows the incorporation of dummy variables into the regression equation to detect either breaks, trend breaks or outliers. Different alpha values can be considered under the null of no breaks or outlier or trend breaks. The algorithm of the IS approach estimators can be executed either separately or jointly. We allow simultaneously detection of breaks by SIS, trend breaks by TIS and outliers by IIS. The general formula of the Win-IS approach is as follows:

$$W(\mathrm{r_t}) = \mu + IS + \varepsilon_t \tag{5}$$

This is a quite general formula but since the IS approach has estimators including IIS, SIS and TIS, this general formula can be reformatted as:

$$W(\mathrm{r_t}) = \mu + \mathrm{IIS} + \mathrm{SIS} + \mathrm{TIS} + \varepsilon_t \tag{6}$$

Pretis et al. (2018) formulated the three types of IS approaches. Equation 6 can be rewritten as:

$$W(\mathrm{r_t}) = \mu + \textstyle\sum_{\mathrm{j=1}}^{\mathrm{n}} \delta_{\mathrm{j}} 1_{\{\mathrm{t=j}\}} + \sum_{\mathrm{j=2}}^{\mathrm{n}} \delta_{\mathrm{j}} 1_{\{\mathrm{t \geq j}\}} + \sum_{\mathrm{j=1}}^{\mathrm{n}} \delta_{\mathrm{j}} 1_{\{\mathrm{t>j}\}}(\mathrm{t-j}) + \varepsilon_t \tag{7}$$

In this case, IS denotes the indicator saturation approach, r_t represents original returns of each coin, $W(\mathrm{r_t})$ represents winsorized vector of returns of each coin, μ represents constant term, δ represents the size of break, trend break, or outlier, and ε represents errors. Equation 7 represents the developed hybrid approach. The strategy is simply winsorizing the dependent variable and then applying IS approach. However, the Win-IS has sub formulas that can be derived from equation 4 including Win-IIS for outlier identification, Win-SIS for break identification and Win-TIS for trend break identification.

3.3. Empirical Application of Win-IS Approach

The process of empirical application of the Win-IS approach begins by:

1. Applying IS technique to identify and record the location and dates of breaks, trend breaks, and outliers in the original returns of each cryptocurrency.
2. Then winsorize each returns following steps given above.
3. Again, apply the IS approach to identify and record the location and dates of breaks, trend breaks, and outliers in the winsorized returns of each cryptocurrency.
4. Compare the performance of the IS approach in the two returns.
5. Report improvements achieved.

4. Results and Discussions

4.1. Descriptive Statistics

Table 3 displays descriptive statistics for each cryptocurrency's winsorized and original returns. The table is split into two panels. Panel A for the original returns, whereas Panel B for the winsorized returns. The results in Table 3 can be classified as central tendency measures, variability measures, and distribution tests. Except for USDT, the original and winsorized returns for each coin have positive mean as expected. This suggests that holders of these coins profited during the examined period, whilst USDT holders lost, signifying its tendency to underperform and generate losses on average. The standard deviation for both the original and winsorized results of each coin is quite high. These high standard deviations indicate significant risk and that returns are very varied or spread around the mean. However, winsorization lowered the risk marginally. According to the data range for each coin, the winsorized returns had a smaller data range than the original returns, indicating that extreme values were pushed closer to the mean. For example, in BTC, the original series ranged from -0.465 to 0.2251, but when winsorized, the minimum value increased to -0.11 and the maximum value reduced to 0.104. This broader adjustment implies a compression of the data range, bringing both extremes closer to zero. The original and winsorized returns for BTC and ETH have negative skewness, indicating a left-skewed distribution. However, LTC, USDT, and XRP have a positively skewed distribution. In contrast, all coins have positive kurtosis greater than +3 in both returns, indicating that the distributions are heavy-tailed and non-normal. Furthermore, the heavier tail of original returns suggests that they are riskier. Following winsorization, each kurtosis decreased, indicating that the tails are becoming less heavy. The Jarque-Bera (JB) test statistic confirms the kurtosis values. Original returns have significant and exceedingly high JB, indicating a large departure from the normally distribution. In comparison, the JB test for winsorized returns is substantially lower, but still significant. This means that, while both the original and winsorized returns are non-normal, the adjustment technique has reduced the divergence from normality by some amount. However, the JB test does not take into consideration variables other than skewness and kurtosis. The JB test still yields a low p-value since the non-normality may be caused by other non-normal features, outliers, or structural breaks in the data.

Table 3: Descriptive statistics

Panel A: Original Returns							
Series	**Min**	**Max**	**Mean**	**Std. Dev.**	**Skewness**	**Kurtosis**	**JB**
BTC	-0.465	0.2251	0.001419	0.038	-0.7895	14.2458	16883.03
LCT	-0.515	0.512	0.001011	0.055	0.103561	15.851	22044.84
ETH	-0.551	0.235	0.000880	0.0497	-0.923868	13.145	9104.12
USDT	-0.053	0.0567	-0.000005	0.004283	0.745575	53.31	216890.1
XRP	-0.551	0.6068	0.000412	0.0615	0.850	20.35	26032.13
Panel B: Winsorised Returns							
Series	**Min**	**Max**	**Mean**	**Std. Dev.**	**Skewness**	**Kurtosis**	**JB**
BTC	-0.1109	0.104009	0.001569	0.0347	-0.174	4.9258	501.45
LCT	-0.14447	0.169612	0.001014	0.048	0.222771	5.239	695.69
ETH	-0.146580	0.127828	0.001096	0.046	-0.243741	4.428	195.03
USDT	-0.012770	0.013738	-0.000016	0.003285	0.10443	8.9266	3011.28
XRP	-0.155830	0.200854	0.000173	0.0514	0.496705	6.2204	972.52

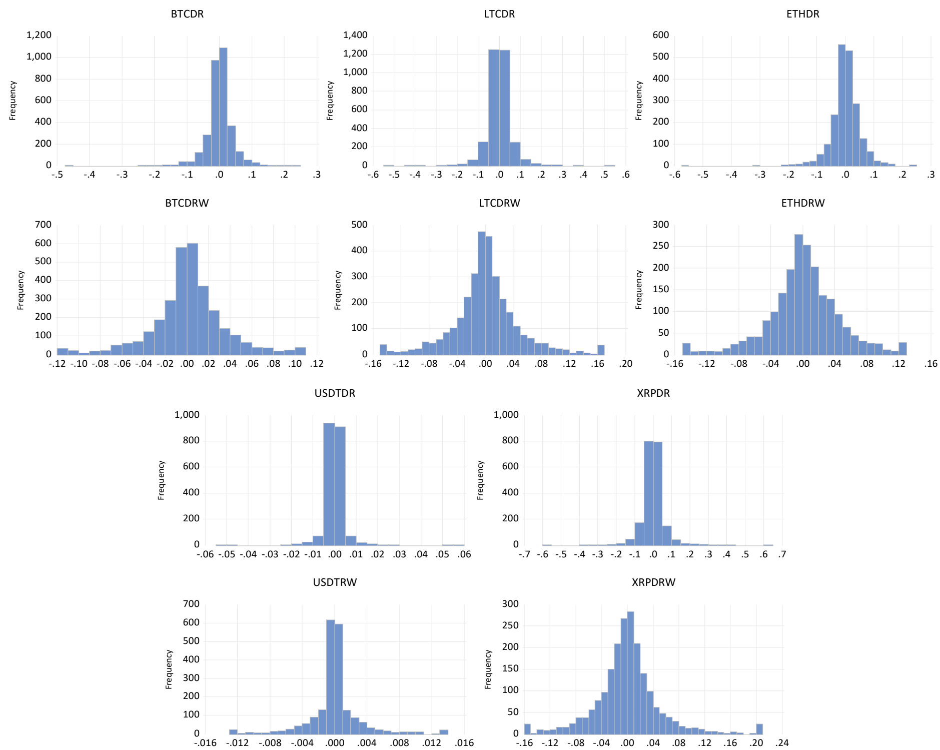

Figure 1: Histograms of Winsorized and unwinsorized returns

Figure 1 shows how the tails of each two histograms presenting winsorized and original returns have changed. This suggests that the distribution of the returns data for each cryptocurrency has been impacted by winsorizing. The winsorizing strategy effectively capped or shaped the extreme results to a less extreme value, reducing the impact of outliers. The original data histograms show a wider distribution with more noticeable tails. The inherent uncertainty and possible non-normality in financial log returns are reflected in this broader spread. These features can make it more difficult to identify structural breaks by adding noise, even if they are crucial for capturing real-world financial dynamics. Detecting structural breaks in the dataset entails locating the points at which the data's statistical characteristics alter. Maintaining a certain amount of non-normality with returns is crucial for capturing dynamics and changes over time in the real world. We reduced the impact of extreme outliers, which otherwise have the potential to distort the results and make it difficult to identify structural fractures, by winsorizing just 1% of the data. Without completely enforcing normalcy, this adjustment results in a distribution that is more symmetrical and closer to normal. The integrity of the data's dynamic features is preserved by such a slight modification, which is essential when looking for structural breaks in time series. Thus, the winsorization acts as a tactical middle ground, improving the data's regression analysis applicability while maintaining important non-stationary components.

4.2. Comparison of IS and Win-IS Performance

4.2.1. Breaks, Trend breaks, and Outliers Detected by IS and Win-IS

Descriptive statistics in Table 3 reveal that even when severe outlier are mitigated by the winsorization strategy, their influence persists since winsorized returns exhibit non-normality. Outliers have a significant effect on summary statistics such as mean and standard deviation. With such high frequency and non-normal data returns, we investigated existing breaks, trend breaks, and outliers that continue to cause this non-normality. This section discusses the performance of IS and Win-IS approaches in detecting breaks, trend breaks, and outliers. Table 4 provides an overview of the performance of the two approaches, IS and Win-IS. Tables 5, 6, and 7 report the significant indicators recognized and retained as outliers, breaks, or/and trend breaks, respectively. Significant indicators are those retained by the method after rejecting the null hypothesis of no break, outlier, or trend break. Each table is divided into two panels, panel A gives the IS approach results and panel B the Win-IS approach results to compare the dates and total of the identified outliers, breaks, and trend breaks in each cryptocurrency. The brackets indicated by positive (+) or negative (-) sign, represent whether the shock on the identified date is up or down. Tables 5, 6, and 7 also provide additional information, such as the dates of the breaks, trend breaks, and outliers found with the two techniques. The findings of Mohamed et al. (2024) are in line with the some of the IS method results shown in these tables.

Table 4: Overall Performance of IS and Win-IS approaches

Returns	IS approach Original Returns				Win-IS approach Winsorized Returns			
	IIS (Outliers)	SIS (Breaks)	TIS (Trend Breaks)	BIC	Win-IIS (Outliers)	Win-SIS (Breaks)	Win-TIS (Trend Breaks)	BIC
BTC	25	12	5	-3.88	0	16	3	-3.89
LCT	28	11	2	-3.24	0	29	3	-3.26
ETH	13	10	8	-3.33	0	12	0	-3.34
USDT	28	8	5	-8.79	32	17	23	-8.91
XRP	26	11	9	-3.19	17	20	15	-3.24

Table 4 compares the overall performance of both IS and Win-IS approaches with 3 estimators for each. On all coins except USDT, the Win-IIS reduced number of outliers to zero in the cases of BTC, LTC, ETH and XRP while increased slightly from 28 to 32 in USDT. Win-SIS revealed more breaks in each coin than SIS. The Win-TIS technique identified fewer trend breaks than TIS tests for BTC, LTC, and ETH, but more trend breaks in USDT and XRP. Overall, the developed Win-IS reduced the number of outliers and trend breaks while revealed hidden breaks. Hence, the Win-IS approach enabled detecting hidden important breaks, trend breaks and outliers (see Table 8). Digital markets perform differently, and BTC, LTC, and ETH appear to behave similarly, as do USDT and XRP.

The comparison of the two approaches IS vs Win-IS is based on Bayesian Information Criteria (BIC). The selection is based on the lowest BIC criterion values. Table 4 shows that the Win-IS technique had the lowest BIC values in all comparisons, indicating that it outperformed the IS approach. As a result, including the winsorization strategy into the IS approach improved its effectiveness. This also emphasizes how significant the newly detected breaks and trend breaks were and how the existence of outliers could obscure them if not handled carefully (masking effect).

Table 5: IIS and Win-IIS Results (Outliers)

Series	Alpha	Panel A: IIS results (Outlier)	Total
BTC	0.0003	1/13/2015(-),1/14/2015(-), 1/15/2015(+), 8/18/2015(-), 1/15/2016(-), 1/11/2017(-), 7/17/2017(+), 7/20/2017(+), 9/14/2017(-), 9/15/2017(+), 12/06/2017(+), 12/07/2017(+), 1/16/2018(-), 2/05/2018(-), 4/02/2019(+), 6/27/2019(-), 7/16/2019(-), 10/25/2019(+), 3/12/2020(-), 3/19/2020(+), 1/21/2021(-), 2/08/2021(+), 5/12/2021(-),5/19/2021(-), 6/13/2022(-)	25
ETH	0.0005	12/22/2017(-), 9/05/2018(-), 10/11/2018(-), 9/24/2019(-), 3/08/2020(-), 3/12/2020(-), 3/13/2020(+), 3/19/2020(+), 1/03/2021(+), 1/21/2021(-), 5/19/2021(-), 5/24/2021(+), 6/21/2021(-)	13

Table 5: IIS and Win-IIS Results (Outliers) (cont.)

Series	Alpha	Panel A: IIS results (Outlier)	Total
LTC	0.0003	1/03/2015(-), 1/14/2015(-), 5/22/2015(+), 6/16/2015(+), 7/10/2015(-), 6/22/2016(-), 12/23/2016(+), 3/30/2017(+), 4/05/2017(+), 5/03/2017(+), 5/23/2017(+), 9/14/2017(-), 12/08/2017(+), 12/09/2017(+), 12/11/2017(+), 12/12/2017(+), 1/16/2018(-), 2/14/2018(+), 2/08/2019(+), 4/02/2019(+), 3/12/2020(-), 1/11/2021(-) , 5/12/2021(-), 5/19/2021(-), 5/24/2021(+), 6/21/2021(-), 9/07/2021(-), 6/30/2023(+)	28
USDT	0.0005	11/30/2017(+), 12/07/2017(+), 12/08/2017(-), 12/12/2017(+), 12/13/2017(-), 12/14/2017(-), 12/24/2017(-), 12/30/2017(+), 1/16/2018(+), 1/17/2018(-), 1/19/2018(-), 2/08/2018(+), 3/24/2018(+), 11/14/2018(-), 11/15/2018(+), 11/23/2018(-), 12/08/2018(+), 6/28/2019(+), 3/12/2020(+), 3/13/2020(-), 3/17/2020(-), 3/19/2020(+), 3/27/2020(+),3/28/2020(-), 5/06/2020(+), 5/07/2020(-), 7/03/2020(-), 8/14/2020(-)	28
XRP	0.0005	12/12/2017(+), 12/13/2017(+), 12/14/2017(+), 12/21/2017(+), 12/29/2017(+), 1/03/2018(+), 1/08/2018(-), 1/16/2018(-), 8/17/2018(+), 9/20/2018(+),9/21/2018(+), 5/14/2019(+),3/12/2020(-),11/21/2020(+), 11/23/2020(+), 12/23/2020(-), 12/24/2020(+), 1/07/2021(+), 1/30/2021(+), 2/01/2021(-), 4/10/2021(+), 4/26/2021(+), 5/19/2021(-), 5/24/2021(+), 5/11/2022(-), 3/21/2023(+)	26
Series	**Alpha**	**Panel B: WIN-IIS results (Outlier)**	**Total**
BTC	0.0003	No	0
ETH	0.0005	No	0
LTC	0.0003	No	0
USDT	0.0005	11/30/2017(+), 12/08/2017(-), 12/12/2017(+), 12/13/2017(-), 12/14/2017(-), 1/14/2018(+), 1/16/2018(+), 1/19/2018(-), 2/05/2018(-), 2/08/2018(+), 2/09/2018(-), 3/19/2018(-), 3/24/2018(+), 4/25/2018(-), 11/23/2018(-), 11/28/2018(+), 3/29/2019(-), 4/25/2019(-), 5/19/2019(+), 6/28/2019(+), 7/16/2019(-), 8/06/2019(-), 11/25/2019(-), 12/18/2019(+), 3/13/2020(-), 3/27/2020(+), 3/28/2020(-), 5/06/2020(+), 5/07/2020(-), 7/02/2020(+), 7/03/2020(-), 8/14/2020(-)	32
XRP	0.0005	12/12/2017(+), 12/13/2017(+), 12/14/2017(+), 12/21/2017(+), 12/29/2017(+), 1/03/2018(+), 8/17/2018(+), 11/21/2020(+), 11/23/2020(+), 12/24/2020(+), 4/10/2021(+),4/13/2021(+), 4/26/2021(+), 5/24/2021(+), 2/07/2022(+), 9/22/2022(+), 3/21/2023(+)	17

Table 6: SIS and Win-SIS Results (Breaks)

Series	Alpha	Panel A: SIS results (Breaks)	Total
BTC	0.0003	11/02/2015(+), 11/04/2015(-), 6/21/2016(-), 6/23/2016(+), 1/05/2017(-), 1/07/2017(+), 12/07/2017(-), 12/17/2017(-), 11/19/2018(-), 11/21/2018(+), 11/08/2022(-), 11/10/2022(+)	12
ETH	0.0005	12/11/2017(+), 12/13/2017(-), 2/06/2018(+), 11/19/2018(-), 11/21/2018(+), 5/21/2021(-), 6/10/2022(-) ,6/14/2022(+), 11/08/2022(-), 11/10/2022(+)	10
LTC	0.0003	1/24/2015(+), 1/26/2015(-), 7/05/2015(+), 5/03/2017(+), 5/08/2017(-), 5/25/2017(-), 5/27/2017(+), 6/16/2017(+), 6/18/2017(-), 5/21/2021(-), 5/25/2021(+)	11

Table 6: SIS and Win-SIS Results (Breaks) (cont.)

Series	Alpha	Panel A: SIS results (Breaks)	Total
USDT	0.0005	12/24/2017(-), 2/05/2018(-), 2/06/2018(+), 2/07/2018(-), 2/10/2018(+), 11/19/2018(-), 11/21/2018(+), 11/24/2018(-)	8
XRP	0.0005	1/08/2018(-), 1/17/2018(+), 1/19/2018(-), 2/11/2018(-), 11/24/2020(+), 4/07/2021(-), 4/08/2021(+), 4/14/2021(-), 5/25/2021(+), 6/21/2021(-), 6/23/2021(+)	11
Series	**Alpha**	**Panel B: WIN-SIS results (Breaks)**	**Total**
BTC	0.0003	1/13/2015(-), 1/15/2015(+), 11/02/2015(+), 11/04/2015(-), 6/21/2016(-), 6/23/2016(+), 1/05/2017(-), 1/07/2017(+), 3/16/2017(-), 3/19/2017(+), 11/19/2018(-), 11/21/2018(+), 12/24/2020(+), 1/09/2021(-), 11/08/2022(-), 11/10/2022(+)	16
ETH	0.0005	1/29/2018(-), 2/06/2018(+), 11/19/2018(-), 11/21/2018(+), 12/17/2018(+), 12/25/2018(-), 9/02/2020(-), 9/06/2020(+), 6/10/2022(-), 6/19/2022(+), 11/08/2022(-), 11/10/2022(+)	12
LTC	0.0003	1/24/2015(+), 1/26/2015(-), 7/10/2015(-), 7/13/2015(-), 7/17/2015(+), 3/30/2017(+), 4/06/2017(-), 4/20/2017(+), 5/10/2017(-), 5/25/2017(-), 5/27/2017(+), 6/16/2017(+), 6/18/2017(-),7/02/2017(+), 7/05/2017(-), 8/27/2017(+), 9/02/2017(-), 12/08/2017(+), 12/13/2017(-),4/02/2019(+), 4/04/2019(-), 12/16/2020(+), 12/20/2020(-), 5/21/2021(+), 5/24/2021(+), 11/08/2021(+), 11/10/2021(+), 11/08/2022(-), 11/10/2022(+)	29
USDT	0.0005	12/09/2017(-), 12/21/2017(+), 12/24/2017(-), 12/31/2017(-), 1/18/2018(+), 2/07/2018(-), 3/25/2018(+), 11/15/2018(+), 11/19/2018(-), 11/21/2018(+), 11/23/2018(-), 12/09/2018(-), 12/30/2018(-), 1/01/2019(+), 3/20/2020(-), 8/12/2020(+), 8/15/2020(-)	17
XRP	0.0005	1/08/2018(-), 1/17/2018(+), 1/19/2018(-), 2/11/2018(-), 4/18/2018(+), 4/21/2018(-), 9/18/2018(+), 9/22/2018(-), 5/16/2019(-), 11/25/2020(-), 12/25/2020(+), 1/06/2021(+), 1/08/2021(-), 2/01/2021(-), 4/07/2021(-), 5/21/2021(-), 5/25/2021(+), 8/15/2021(-), 11/08/2022(-), 11/10/2022(+)	20

Table 7: TIS and Win-TIS Results (Trend Breaks)

Series	Alpha	Panel A: TIS results (Trend Breaks)	Total
BTC	0.0003	7/13/2017(-), 7/15/2017(+), 7/17/2017(-), 12/06/2017(+), 12/23/2017(-)	5
ETH	0.0005	1/14/2018(-), 1/16/2018(+), 1/20/2018(-), 1/21/2018(+), 1/27/2018(-), 2/06/2018(+), 5/21/2021(+), 5/25/2021(-)	8
LTC	0.0003	7/12/2015(-), 7/13/2015 (+)	2
USDT	0.0005	12/20/2017(+), 12/24/2017(-), 1/18/2018(-), 1/30/2018(+), 2/03/2018(-)	5
XRP	0.0005	2/05/2018(+), 2/09/2018(-), 11/23/2020(-), 11/26/2020(+), 11/27/2020(-), 4/03/2021(+), 4/05/2021(-), 5/20/2021(-), 5/25/2021(+)	9
		Panel B: WIN-TIS results (Trend Breaks)	
BTC	0.0003	12/16/2017(-), 12/22/2017(+), 12/23/2017(-)	3
ETH	0.0005	No	0
LTC	0.0003	6/24/2015(+), 7/10/2015(+), 7/12/2015(-)	3

Table 7: TIS and Win-TIS Results (Trend Breaks) (cont.)

Series	Alpha	Panel B: WIN-TIS results (Trend Breaks)	
USDT	0.0005	12/02/2017(+), 12/14/2017(-), 12/20/2017(+), 1/03/2018(-), 1/04/2018(+), 1/06/2018(-), 1/30/2018(+), 2/10/2018(-), 3/18/2018(+), 3/19/2018(-), 3/25/2018(+), 11/12/2018(-), 11/15/2018(+), 12/05/2018(+), 12/09/2018(-), 11/21/2019(+), 11/23/2019(-), 11/26/2019(+), 3/08/2020(-), 3/09/2020(+) , 3/13/2020(-), 3/17/2020(+), 3/20/2020(-)	23
XRP	0.0005	2/05/2018(+), 2/09/2018(-), 5/12/2019(+), 5/14/2019(-), 11/26/2020(+), 11/27/2020(-), 12/15/2020(+), 12/16/2020(-), 12/25/2020(+), 1/28/2021(+), 1/30/2021(-), 4/03/2021(+), 4/08/2021(-), 8/08/2021(+), 8/15/2021(-)	15

4.2.2. Improvements of the Detectability IS Approach

In Tables 5,6, and 7 we presented the dates of the detected breaks, trend breaks, and outliers in the five cryptocurrencies. The results show that new breaks, trend breaks, and outliers were revealed after extreme observations are lessened by the winsorization approach. As discussed, this become evidence that some outliers mask some breaks or trend breaks. Table 8 show the overall number of new breaks, trend breaks, and outliers emerged together with those repeated.

Table 8: Win-IS performance and discovery

Series	Win-IIS		Win-SIS		Win-TIS		Total
	Repeated Outliers	New Outliers	Repeated Breaks	New Breaks	Repeated Trend Breaks	New Trend Breaks	
BTCDW	0	0	10	6	1	2	19
ETHDW	0	0	6	6	0	0	12
LTCDW	0	0	6	23	1	2	32
USDTDW	20	12	3	14	2	21	72
XRPDW	14	3	5	15	6	9	52
Total	**34**	**15**	**30**	**64**	**10**	**34**	**187**

Table 8 shows that, in contrast to those detected by IIS alone, Win-IIS did not find any new or recurring outliers in BTC, ETH, or LTC. Furthermore, Win-IIS identified three new outliers in XRP and 12 new outliers in USDT, whereas 14 and 20 outliers, respectively, repeated. Win-SIS discovered 94 breaks across the five markets, 30 of which were repeated as SIS detected, and revealed 64 new breaks. Win-TIS, on the other hand, revealed 44 trend breaks across five markets, 10 of which were previously spotted by TIS and 34 of which were new. Table 8 displays the distribution of 94 breaks and 44 trend breaks among the five markets. Additional details about the overall outcomes shown in Table 8 are provided in Tables 9–12. These details include the type of date (original value or winsored value), the type of estimator captured (Win-IIS, Win-SIS, and Win-TIS), and the status of each date (repeated, new, or changed to another).

By comparing and classifying the data into recurrent and emerging changes, these tables also demonstrate the presence of the masking effect.

Table 9: BTCDW and ETHDW: Emerging and Repeated Patterns Detected by Win-IS

No.	Win-IS Performance in BTCDW				Win-IS Performance in ETHDW			
	Date	Type	Estimator	Status	Date	Type	Estimator	Status
1.	1/13/2015	Winsored	Win-SIS	IIS→Win-SIS	1/29/2018	Origin	Win-SIS	New
2.	1/15/2015	Winsored	Win-SIS	IIS→Win-SIS	2/6/2018	Origin	Win-SIS	Repeated
3.	11/2/2015	Winsored	Win-SIS	Repeated	11/19/2018	Winsored	Win-SIS	Repeated
4.	11/4/2015	Origin	Win-SIS	Repeated	11/21/2018	Origin	Win-SIS	Repeated
5.	6/21/2016	Origin	Win-SIS	Repeated	12/17/2018	Origin	Win-SIS	New
6.	6/23/2016	Origin	Win-SIS	Repeated	12/25/2018	Origin	Win-SIS	New
7.	1/5/2017	Origin	Win-SIS	Repeated	9/2/2020	Origin	Win-SIS	New
8.	1/7/2017	Origin	Win-SIS	Repeated	9/6/2020	Origin	Win-SIS	New
9.	3/16/2017	Origin	Win-SIS	New	6/10/2022	Origin	Win-SIS	Repeated
10.	3/19/2017	Origin	Win-SIS	New	6/19/2022	Origin	Win-SIS	New
11.	11/19/2018	Winsored	Win-SIS	Repeated	11/8/2022	Winsored	Win-SIS	Repeated
12.	11/21/2018	Origin	Win-SIS	Repeated	11/10/2022	Winsored	Win-SIS	Repeated
13.	12/24/2020	Origin	Win-SIS	New				
14.	1/9/2021	Origin	Win-SIS	New				
15.	11/8/2022	Origin	Win-SIS	Repeated				
16.	11/10/2022	Origin	Win-SIS	Repeated				
17.	12/16/2017	Origin	Win-TIS	New				
18.	12/22/2017	Winsored	Win-TIS	New				
19.	12/23/2017	Origin	Win-TIS	Repeated				

Table 10: LTCDW: Emerging and Repeated Patterns Detected by Win-IS

Win-IS Performance in LTCDW									
No.	Date	Type	Estimator	Status	No.	Date	Type	Estimator	Status
1.	1/24/2015	Winsor	Win-SIS	Repeated	21.	4/4/2019	Origin	Win-SIS	New
2.	1/26/2015	Origin	Win-SIS	Repeated	22.	12/16/2020	Origin	Win-SIS	New
3.	7/10/2015	Winsor	Win-SIS	IIS→Win-SIS	23.	12/20/2020	Origin	Win-SIS	New
4.	7/13/2015	Origin	Win-SIS	TIS→Win-SIS	24.	5/21/2021	Winsor	Win-SIS	New
5.	7/17/2015	Origin	Win-SIS	New	25.	5/24/2021	Winsor	Win-SIS	IIS→Win-SIS
6.	3/30/2017	Winsor	Win-SIS	IIS→Win-SIS	26.	11/8/2021	Origin	Win-SIS	New
7.	4/6/2017	Origin	Win-SIS	New	27.	11/10/2021	Origin	Win-SIS	New
8.	4/20/2017	Origin	Win-SIS	New	28.	11/8/2022	Winsor	Win-SIS	New
9.	5/10/2017	Origin	Win-SIS	New	29.	11/10/2022	Winsor	Win-SIS	New
10.	5/25/2017	Origin	Win-SIS	Repeated	30.	6/24/2015	Origin	Win-TIS	New
11.	5/27/2017	Origin	Win-SIS	Repeated	31.	7/10/2015	Winsor	Win-TIS	IIS→Win-TIS
12.	6/16/2017	Origin	Win-SIS	Repeated	32.	7/12/2015	Winsor	Win-TIS	Repeated
13.	6/18/2017	Origin	Win-SIS	Repeated					
14.	7/2/2017	Origin	Win-SIS	New					
15.	7/5/2017	Origin	Win-SIS	New					
16.	8/27/2017	Origin	Win-SIS	New					
17.	9/2/2017	Origin	Win-SIS	New					
18.	12/8/2017	Winsor	Win-SIS	IIS→Win-SIS					
19.	12/13/2017	Origin	Win-SIS	New					
20.	4/2/2019	Winsor	Win-SIS	IIS→Win-SIS					

Table 11: USDTDW: Emerging and Repeated Patterns Detected by Win-IS

Win-IS Performance in USDTW									
No.	**Date**	**Type**	**Estimator**	**Status**	**No.**	**Date**	**Type**	**Estimator**	**Status**
1.	11/30/2017	Winsor	Win-IIS	Repeated	37.	1/18/2018	Origin	Win-SIS	New
2.	12/8/2017	Origin	Win-IIS	Repeated	38.	2/7/2018	Origin	Win-SIS	New
3.	12/12/2017	Winsor	Win-IIS	Repeated	39.	3/25/2018	Origin	Win-SIS	New
4.	12/13/2017	Winsor	Win-IIS	Repeated	40.	11/15/2018	Winsor	Win-SIS	IIS→Win-SIS
5.	12/14/2017	Winsor	Win-IIS	Repeated	41.	11/19/2018	Origin	Win-SIS	Repeated
6.	1/14/2018	Winsor	Win-IIS	New	42.	11/21/2018	Winsor	Win-SIS	Repeated
7.	1/16/2018	Winsor	Win-IIS	Repeated	43.	11/23/2018	Winsor	Win-SIS	IIS→Win-SIS
8.	1/19/2018	Winsor	Win-IIS	Repeated	44.	12/9/2018	Origin	Win-SIS	New
9.	2/5/2018	Winsor	Win-IIS	SIS→Win-IIS	45.	12/30/2018	Origin	Win-SIS	New
10.	2/8/2018	Winsor	Win-IIS	Repeated	46.	1/1/2019	Origin	Win-SIS	New
11.	2/9/2018	Winsor	Win-IIS	Repeated	47.	3/20/2020	Origin	Win-SIS	New
12.	3/19/2018	Winsor	Win-IIS	Repeated	48.	8/12/2020	Origin	Win-SIS	New
13.	3/24/2018	Winsor	Win-IIS	Repeated	49.	8/15/2020	Origin	Win-SIS	New
14.	4/25/2018	Winsor	Win-IIS	New	50.	12/2/2017	Origin	Win-TIS	New
15.	11/23/2018	Winsor	Win-IIS	Repeated	51.	12/14/2017	Winsor	Win-TIS	IIS→Win-TIS
16.	11/28/2018	Winsor	Win-IIS	New	52.	12/20/2017	Origin	Win-TIS	Repeated
17.	3/29/2019	Origin	Win-IIS	New	53.	1/3/2018	Origin	Win-TIS	New
18.	4/25/2019	Origin	Win-IIS	New	54.	1/4/2018	Origin	Win-TIS	New
19.	5/19/2019	Winsor	Win-IIS	New	55.	1/6/2018	Origin	Win-TIS	New
20.	6/28/2019	Winsor	Win-IIS	Repeated	56.	1/30/2018	Origin	Win-TIS	Repeated
21.	7/16/2019	Origin	Win-IIS	New	57.	2/10/2018	Origin	Win-TIS	SIS→Win-TIS
22.	8/6/2019	Origin	Win-IIS	New	58.	3/18/2018	Origin	Win-TIS	New
23.	11/25/2019	Winsor	Win-IIS	New	59.	3/19/2018	Winsor	Win-TIS	New
24.	12/18/2019	Origin	Win-IIS	New	60.	3/25/2018	Origin	Win-TIS	New
25.	3/13/2020	Winsor	Win-IIS	Repeated	61.	11/12/2018	Origin	Win-TIS	New
26.	3/27/2020	Winsor	Win-IIS	Repeated	62.	11/15/2018	Winsor	Win-TIS	IIS→Win-TIS
27.	3/28/2020	Winsor	Win-IIS	Repeated	63.	12/5/2018	Origin	Win-TIS	New
28.	5/6/2020	Origin	Win-IIS	Repeated	64.	12/9/2018	Origin	Win-TIS	New
29.	5/7/2020	Winsor	Win-IIS	Repeated	65.	11/21/2019	Origin	Win-TIS	New
30.	7/2/2020	Winsor	Win-IIS	New	66.	11/23/2019	Origin	Win-TIS	New
31.	7/3/2020	Winsor	Win-IIS	Repeated	67.	11/26/2019	Origin	Win-TIS	New
32.	8/14/2020	Winsor	Win-IIS	Repeated	68.	3/8/2020	Origin	Win-TIS	New
33.	12/9/2017	Origin	Win-SIS	New	69.	3/9/2020	Origin	Win-TIS	New
34.	12/21/2017	Origin	Win-SIS	New	70.	3/13/2020	Winsor	Win-TIS	IIS→Win-TIS
35.	12/24/2017	Winsor	Win-SIS	Repeated	71.	3/17/2020	Winsor	Win-TIS	IIS→Win-TIS
36.	12/31/2017	Origin	Win-SIS	New	72.	3/20/2020	Origin	Win-TIS	New

Repeated breaks, trend breaks, and outliers as shown in Tables 9-12 signify their importance and persistence despite certain observations being weighted down. The appearance of new outliers, breaks, and trend breaks suggests that they were important and if extreme values are not winsored, they will be buried. The Win-IS technique also enabled to redetect some of the treated data, as shown in Tables 9–12. Table 9 demonstrates, for instance, that four of the winsored extreme values in BTC appear as breaks and one as a trend break, whereas three winsored observations were classified as breaks in ETH. Tables 10–12 demonstrate that for the remaining coins, a portion of the winsorized observations is identified as either outliers, trend breaks, or breaks. Some of the winsored observations in these tables shift to either break or trend break, indicating that they remain noteworthy. It also shows that Win-IS can reveal previously unseen data points with possibly unique characteristics or behaviors. This emphasizes the need of using sophisticated techniques, such as Win-IS, to improve the detection sensitivity of IS approaches and acquire a more thorough understanding of the changing dynamics inside market data.

Table 12: XRPDW: Emerging and Repeated Patterns Detected by Win-IS approach

Win-IS Performance in XRPTW

No.	Date	Type	Estimator	Status	No.	Date	Type	Estimator	Status
1.	12/12/2017	Winsor	Win-IIS	Repeated	27.	11/25/2020	Origin	Win-SIS	New
2.	12/13/2017	Winsor	Win-IIS	Repeated	28.	12/25/2020	Origin	Win-SIS	New
3.	12/14/2017	Winsor	Win-IIS	Repeated	29.	1/6/2021	Origin	Win-SIS	New
4.	12/21/2017	Winsor	Win-IIS	Repeated	30.	1/8/2021	Origin	Win-SIS	New
5.	12/29/2017	Winsor	Win-IIS	Repeated	31.	2/1/2021	Winsor	Win-SIS	IIS→Win-SIS
6.	1/3/2018	Winsor	Win-IIS	Repeated	32.	4/7/2021	Winsor	Win-SIS	Repeated
7.	8/17/2018	Winsor	Win-IIS	Repeated	33.	5/21/2021	Winsor	Win-SIS	New
8.	11/21/2020	Winsor	Win-IIS	Repeated	34.	5/25/2021	Origin	Win-SIS	TIS→Win-SIS
9.	11/23/2020	Winsor	Win-IIS	Repeated	35.	8/15/2021	Origin	Win-SIS	New
10.	12/24/2020	Winsor	Win-IIS	Repeated	36.	11/8/2022	Origin	Win-SIS	New
11.	4/10/2021	Winsor	Win-IIS	Repeated	37.	11/10/2022	Origin	Win-SIS	New
12.	4/13/2021	Origin	Win-IIS	New	38.	2/5/2018	Winsor	Win-TIS	Repeated
13.	4/26/2021	Winsor	Win-IIS	Repeated	39.	2/9/2018	Origin	Win-TIS	Repeated
14.	5/24/2021	Winsor	Win-IIS	Repeated	40.	5/12/2019	Origin	Win-TIS	New
15.	2/7/2022	Origin	Win-IIS	New	41.	5/14/2019	Winsor	Win-TIS	Repeated
16.	9/22/2022	Origin	Win-IIS	New	42.	11/26/2020	Winsor	Win-TIS	Repeated
17.	3/21/2023	Winsor	Win-IIS	Repeated	43.	11/27/2020	Origin	Win-TIS	Repeated
18.	1/8/2018	Winsor	Win-SIS	Repeated	44.	12/15/2020	Origin	Win-TIS	New
19.	1/17/2018	Origin	Win-SIS	Repeated	45.	12/16/2020	Origin	Win-TIS	New
20.	1/19/2018	Origin	Win-SIS	Repeated	46.	12/25/2020	Origin	Win-TIS	New
21.	2/11/2018	Origin	Win-SIS	Repeated	47.	1/28/2021	Origin	Win-TIS	New
22.	4/18/2018	Origin	Win-SIS	New	48.	1/30/2021	Winsor	Win-TIS	IIS→Win-TIS
23.	4/21/2018	Origin	Win-SIS	New	49.	4/3/2021	Origin	Win-TIS	Repeated
24.	9/18/2018	Origin	Win-SIS	New	50.	4/8/2021	Origin	Win-TIS	New
25.	9/22/2018	Origin	Win-SIS	New	51.	8/8/2021	Origin	Win-TIS	New
26.	5/16/2019	Origin	Win-SIS	New	52.	8/15/2021	Origin	Win-TIS	New

Overall, the paper firstly detected breaks, trend breaks, and outliers using the IS approach in each coin concurrently. Secondly, we show that extreme values in the data may hide some significant changes due to the simultaneous execution of the three estimators of IS and due to the highly fluctuated data, by undertaking a strategy to lessen only 1% of the extreme observations in each using the winsorization approach. Thirdly, as discussed in the results, the Win-IS approach enabled to reveal new breaks, trend breaks, and outliers after extreme value were treated. The BIC value led to the decision that Win-IS results outperform the IS results. The p-values of the hypothesis also reveal significance of the new results revealed. Win-IS estimators also repeated some of the observations as identified by IS estimators. Therefore, we emphasize that excessive values can make it difficult for statistical change tests to accurately detect where the breaks and trend breaks occur. This highlights the necessity of developing precise hybrid methodologies that might assist the existing tests in obtaining reliable results.

Finally, results from Tables 5-7 implies that the market returns encountered both upward and downward movements on different occasions which suggests that the coin market encountered phases of instability and unpredictability. Most outliers and breaks detected fall in the years 2017, 2018, 2020, and 2021. 2018 witnessed a total of 100 breaks, trend breaks, and outliers. This was followed by 88 in 2017, 59 in 2021, and 57 in 2020. These shifts can be caused by a variety of factors, including economic events, political developments, and investor sentiment. Specifically, in 2017, BTC achieved an all-time high of $20,000 and saw a rise in interest; in 2018, it marked crypto winter; and in 2020, it saw the Covid-19 epidemic.

5. Conclusions

This article improves the detectability of the IS approach by combining it with the winsorization strategy and hence proposes a technique known as Win-IS. The performance of Win-IS is then empirically compared to IS in five cryptocurrency markets. The study identified and dated outliers, breaks, and trend breaks in each market using both IS and Win-IS estimators. The Win-IS strategy outperformed the IS technique, as demonstrated by BIC scores. Furthermore, the Win-IS technique reduced severe outliers in four coins while revealing new outliers, breaks, and trend breaks, some of which were duplicated from the IS results. The repeated outliers, breaks, and trend breaks show their importance in this market because they remained constant in both winsored and original returns. The new findings demonstrate that if extreme values are not addressed, they will not be discovered. This highlights the importance of thoroughly evaluating the data before using any detection strategy, as some outliers disguise potential breaks. Subsequent research efforts may focus on adapting and expanding this hybrid methodology, as well as its relevance to other financial markets. Other methods can be compared to ours as well. The study concentrated on five digital currencies and only winsorized their first and 99th percentiles.

Although the suggested technique was first used for cryptocurrency datasets, it could have broad relevance in fields including technology research, financial markets, and economic forecasts. Additionally, the work tackles the problem of severe outliers by enhancing detectability of IS technique using Winsorization, which successfully handled tail outliers. Without making any assumptions beforehand, the enhanced IS technique is reliable in identifying outliers, trend breaks, and structural breaks, guaranteeing thorough analysis across datasets. The study also emphasizes how the Win-IS and IS technique may concurrently capture outliers, trend breaks, and breaks. Lastly, the technique is resilient under fat-tailed distributions, even if the underlying data-generating mechanism assumes near-normal behavior. Future research might go deeper into these areas to improve the technique's resilience and usefulness.

References

Abakah, E. J. A., Gil-Alana, L. A., Madigu, G. and Romero-Rojo, F., (2020). Volatility persistence in cryptocurrency markets under structural breaks. *International Review of Economics & Finance*, 69, pp. 680–691. https://doi.org/10.1016/j.iref.2020.06.035.

Adams, J., Hayunga, D., Mansi, S., Reeb, D. and Verardi, V., (2019). Identifying and treating outliers in finance. *Financial Management*, 48(2), pp. 345–384. https://doi.org/ 10.1111/fima.12269.

Afanasyev, D. O., Fedorova, E. A., (2019). On the impact of outlier filtering on the electricity price forecasting accuracy. *Applied Energy*, 236, pp. 96–210. https://doi.org/10.1016/j.apenergy.2018.11.076.

Aharon, D. Y., Butt, H. A., Jaffri, A. and Nichols, B., (2023). Asymmetric volatility in the cryptocurrency market: New evidence from models with structural breaks. *International Review of Financial Analysis*, 87, pp. 102651. https://doi.org/10.1016/j.irfa.2023.102651.

Ané, T., Ureche-Rangau, L., Gambet, J. B. and Bouverot, J., (2008). Robust outlier detection for Asia–Pacific stock index returns. *Journal of International Financial Markets, Institutions and Money*, 18(4), pp. 326–343. https://doi.org/10.1016/j.intfin.2007. 03.001.

Bai, J., Perron, P., (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 66(1), pp. 47–78. http://doi/10.2307/2998540.

Bai, J., Perron, P., (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1), pp. 1–22. https://doi.org/10.1002/jae.659.

Brooks, C., (2019). Introductory Econometrics for Finance, Fourth Edition. *Cambridge University Press*. https://doi.org/10.1017/9781108524872.

Brownen-Trinh, R., (2019). Effects of winsorization: The cases of forecasting non-GAAP and GAAP earnings. *Journal of Business Finance & Accounting*, 46(1–2), pp. 105–135. https://doi.org/10.1111/jbfa.12365.

Canh, N. P., Wongchoti, U., Thanh, S. D. and Thong, N. T., (2019). Systematic risk in cryptocurrency market: Evidence from DCC-MGARCH model. *Finance Research Letters*, 29, pp. 90–100. https://doi.org/10.1016/j.frl.2019.03.011.

Castle, J. L., Fawcett, N. W. and Hendry, D. F., (2011). Forecasting breaks and forecasting during breaks. *The Oxford Handbook of Economic Forecasting.* https://doi.org/ 10.1093/oxfordhb/9780195398649.013.0012.

Castle, J. L., Doornik, J. A. and Hendry, D. F., (2012). Model selection when there are multiple breaks. *Journal of Econometrics,* 169(2), pp. 239–246. https://doi.org/ 10.1016/j.jeconom.2012.01.026.

Castle, J. L., Doornik, J. A., Hendry, D. F. and Pretis, F., (2015). Detecting Location Shifts During Model Selection by Step-Indicator Saturation. *Econometrics*, 3(2), pp. 240–264. DOI: 10. 3390/econometrics3020240.

Castle, J. L., Doornik, J. A. and Hendry, D. F., (2021). Forecasting facing economic shifts, climate change and evolving pandemics. *Econometrics*, 10(1), 2. https://doi.org/ 10.3390/econometrics10010002.

Castle, J.L., Hendry, D.F., (2022). Econometrics for modelling climate change. *Oxford Research Encyclopedia of Economics and Finance*. https://doi.org/10.1093/acrefore/ 9780190625979.013.675.

Chatzikonstanti, V., (2017). Breaks and outliers when modelling the volatility of the US stock market. *Applied Economics*, *49*(46), pp. 4704-4717.

Cheng, K., Young, D. S., (2023). An Approach for Specifying Trimming and Winsorization Cutoffs. *Journal of Agricultural, Biological and Environmental Statistics*, 28, pp. 299–323. https://doi.org/10.1007/s13253-023-00527-4.

Choi, S. W., (2009). The effect of outliers on regression analysis: regime type and foreign direct investment. *Quarterly Journal of Political Science*, *4*(2), pp. 153–165.

Chow, G.C., (1960). Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica*, 28(3), pp. 591–605. https://doi.org/10.2307/1910133.

Dixon, W. J., (1960). Simplified estimation from censored normal samples. *The Annals of Mathematical Statistics*, pp. 385–391.

Dutta, A., Bouri, E., (2022). Outliers and time-varying jumps in the cryptocurrency markets. *Journal of Risk and Financial Management*, 15(3), pp.128. https://doi.org/ 10.3390/jrfm15030128.

Fearnhead, P., Rigaill, G., (2019). Changepoint detection in the presence of outliers. *Journal of the American Statistical Association,* 114(525), pp. 169-183. https://doi.org/10.1080/01621459.2017.1385466.

Ghouse, G., Aslam, A. and Bhatti, M. I., (2021). Role of Islamic banking during COVID-19 on political and financial events: Application of impulse indicator saturation. *Sustainability*, 13(21), pp. 11619. https://doi.org/10.3390/su132111619.

Ghouse, G., Bhatti, M. I. and Shahid, M. H., (2022). Impact of COVID-19, political, and financial events on the performance of commercial banking sector. *Journal of Risk and Financial Management,* 15(4), pp. 186. https://doi.org/10.3390/jrfm15040186.

Hamadani, A., Ganai, N. A., (2023). Artificial intelligence algorithm comparison and ranking for weight prediction in sheep. *Scientific Reports*, 13(1), pp. 13242. https://doi.org/10.1038/s41598-023-40528-4.

Hawkins, D. M., (1980). Identification of outliers. London: Chapman and Hall. https://doi.org/10.1007/978-94-015-3994-4.

Hendry, D. F., (1999). An econometric analysis of US food expenditure, 1931–1989. In J. R. Magnus, & M. S. Morgan (Eds.). *Methodology and tacit knowledge: two experiments in econometrics,* pp. 341–361. Chichester: John Wiley and Sons.

Hendry, D. F., Johansen, S., Santos, C., (2008). Automatic Selection of Indicators in a Fully Saturated Regression. *Computational Statistics*, 23(2), pp. 317–335. doi:10.1007/ s00180-007-0054-z.

Inclán, C., Tiao, G. C., (1994). Use of cumulative sums of squares for retrospective detection of changes of variance. *Journal of the American Statistical Association*, 89(427), pp. 913–923. https://doi.org/10.1080/01621459.1994.10476824.

Ismail, M. T., Nasir, I. N. M., (2018). *Structural breaks and outliers in ASEAN Shariah compliant indices: The impulse indicator saturation approach*. In AIP Conference Proceedings 1 2013.

Jiang, Z., Mensi, W. and Yoon, S. M., (2023). Risks in Major Cryptocurrency Markets: Modeling the Dual Long Memory Property and Structural Breaks. *Sustainability*, 15(3), pp. 2193. https://doi.org/10.3390/su15032193.

Johansen, S., Nielsen, B., (2016). Asymptotic Theory of Outlier Detection Algorithms for Linear Time Series Regression Models. *Scandinavian Journal of Statistics*, 43(2), pp. 321–348. doi: 10.1111/sjos.12174.

Kaseke, F., Ramroop, S. and Mwambi, H., (2022). A comparative analysis of the volatility nature of cryptocurrency and JSE market. *Investig. Manag. Financ. Innov*, 19, pp. 23–39. http://dx.doi.org/10.21511/imfi.19(4).2022.03.

Li, L., Xie, X. and Yang, J., (2021). A predictive model incorporating the change detection and Winsorization methods for alerting hypoglycemia and hyperglycemia. *Medical & Biological Engineering & Computing*, 59, pp. 2311–2324. https://doi.org/ 10.1007/s11517-021-02433-8.

Lien, D., Balakrishnan, N., (2005). On regression analysis with data cleaning via trimming, winsorization, and dichotomization. *Communications in Statistics–Simulation and Computation*, 34(4), pp. 839–849. https://doi.org/10.1080/03610910500307695.

Mandaci, P. E., Cagli, E. C., (2022). Herding intensity and volatility in cryptocurrency markets during the COVID-19, *Finance Research Letters*, 46, pp. 102382. https://doi.org/10.1016/j.frl.2021.102382.

Marczak, M., Proietti, T., (2016). Outlier detection in structural time series models: The indicator saturation approach. *International Journal of Forecasting*, 32(1), pp. 180–202. https://doi.org/10.1016/j.ijforecast.2015.04.005.

Mohamed, S. D., Ismail, M. T. and Ali, M. K. M, (2023). Detecting Structural Breaks in Cryptocurrency Market, *Jurnal Ekonomi Malaysia*, 57(2), pp. 107–22. *Jurnal Ekonomi Malaysia.* http://dx.doi.org/10.17576/JEM-2023-5702-08.

Mohamed, S. D., Ismail, M. T. and Ali, M. K. B. M. (2024). Cryptocurrency Returns Over a Decade: Breaks, Trend Breaks and Outliers. *Scientific Annals of Economics and Business*, *71*(1), pp. 1–20.

Moir, R., (1998). A Monte Carlo analysis of the Fisher randomization technique: reviving randomization for experimental economists. *Experimental Economics*, 1, pp. 87–100. https://doi.org/10.1023/A:1009961917752.

Muhammadullah, S., Urooj, A., Mengal, M. H., Khan, S. A. and Khalaj, F., (2022). Cross-Sectional Analysis of Impulse Indicator Saturation Method for Outlier Detection Estimated via Regularization Techniques with Application of COVID-19 Data. *Computational and Mathematical Methods in Medicine*. https://doi.org/10.1155/ 2022/2588534.

Mulry, M. H., Oliver, B. E. and Kaputa, S. J., (2014). Detecting and treating verified influential values in a monthly retail trade survey. *Journal of Official Statistics*, 30(4), pp. 721–747. https://doi.org/10.2478/jos-2014-0045.

Narayan, P. K., Popp, S., (2010). A new unit root test with two structural breaks in level and slope at unknown time. *Journal of Applied Statistics*, 37(9), pp. 1425–1438. https://doi.org/10.1080/02664760903039883.

Ohara, H.I.,(1999). A unit root test with multiple trend breaks: A theory and an application to US and Japanese macroeconomic time-series. *Japanese Economic Review*, 50(3), pp. 266–290. https://doi.org/10.1111/1468-5876.00119.

Panday, A., (2015). Impact of monetary policy on exchange market pressure: The case of Nepal. *Journal of Asian Economics*, 37, pp. 59–71. https://doi.org/10.1016/j.asieco.2015.02.001.

Papell, D. H., Prodan, R., (2003). The Uncertain Unit Root in U.S. Real GDP: Evidence with Restricted and Unrestricted Structural Change. *Journal of Money, Credit and Banking*, 36(3). 10.1353/mcb.2004.0059.

Pretis, F., Mann, M. L. and Kaufmann, R. K., (2015). Testing Competing Models of the Temperature Hiatus: Assessing the Effects of Conditioning Variables and Temporal Uncertainties through Sample-Wide Break Detection. *Climatic Change*, 131(4), pp. 705–718. doi: 10.1007/s10584-015-1391-5.

Pretis, F., Schneider, L., Smerdon, J. E. and Hendry, D. F., (2017). Detecting volcanic eruptions in temperature reconstructions by designed break-indicator saturation. *Environmental Economics and Sustainability*, pp. 7–37. https://doi.org/10.1002/9781119328223.ch2.

Pretis, F., Reade, J. J., Sucarrat, G., (2018). Automated general-to-specific (GETS) regression modeling and indicator saturation for outliers and structural breaks. *Journal of Statistical Software,* 86 (3), pp. 1–44. 10.18637/jss.v086.i03.

Quandt, R. E., (1960). Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes Tests of the hypothesis that a linear regression system obeys two separate regimes. *Journal of the American Statistical Association*, 55(290), pp. 324–330, DOI: 10.1080/01621459.1960.10482067.

Rivest, L. P., (1994). Statistical properties of Winsorized means for skewed distributions. *Biometrika*, 81(2), pp. 373–383. https://doi.org/10.1093/biomet/81.2.373.

Rodrigues, P. M. M., Rubia, A., (2011). The effects of additive outliers and measurement errors when testing for structural breaks in variance. *Oxford Bulletin of Economics and Statistics*, 73, pp. 449–68.

Sahoo, P. K., (2021). COVID-19 pandemic and cryptocurrency markets: an empirical analysis from a linear and nonlinear causal relationship. *Studies in Economics and Finance*, 38(2), pp. 454–468. https://doi.org/10.1108/SEF-09-2020-0385.

Sharma, S., Chatterjee, S., (2021). Winsorization for robust Bayesian neural networks. *Entropy,* 23(11), pp. 1546. https://doi.org/10.3390/e23111546.

Thies, S., Molnár, P., (2018). Bayesian change point analysis of Bitcoin returns. *Finance Research Letters*, 27, pp. 223–227. https://doi.org/10.1016/j.frl.2018.03.018.

Tukey, J. W., (1962). The future of data analysis. *The annals of mathematical statistics*, *33*(1), pp. 1–67.

Xao, O. G., Yahaya, S. S. S., Abdullah, S., Yusof, Z. M., (2014, July). H-statistic with winsorized modified one-step M-estimator for two independent groups design. *AIP Conference Proceedings,* Vol. 1605, No. 1, pp. 928–931. American Institute of Physics.

Yen, T. C., Beng, K. Y., Haur, N. K., Huat, N. K., (2022). Structural change analysis of active cryptocurrency market. *Asian Academy of Management Journal of Accounting and Finance*, 18(2), pp. 63–85. https://doi.org/10.21315/aamjaf2022.18.2.4.

Yuliyani, L., Kurnia, A., Indahwati, I., (2017, March). Winsorization on linear mixed model (Case study: National exam of senior high school in West Java). In *AIP Conference Proceedings*, Vol. 1827, No. 1. AIP Publishing.

Zivot, E., Andrews, D. W. K., (2002). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business and Economic Statistics* 20(1), pp. 25–44. https://doi.org/10.1198/073500102753410372.