SILESIAN UNIVERSITY OF TECHNOLOGY PUBLISHING HOUSE

SCIENTIFIC PAPERS OF SILESIAN UNIVERSITY OF TECHNOLOGY 2024
ORGANIZATION AND MANAGEMENT SERIES NO. 213

ANOMALY DETECTION IN UNIVARIATE TIME SERIES USING A MULTI-CRITERIA APPROACH

Maciej WOLNY

Silesian University of Technology, Faculty of Organization and Management, Department of Economics and Computer Science; maciej.wolny@polsl.pl, ORCID: 0000-0002-8872-7794

Purpose: The purpose of this study is to propose and evaluate a multi-criteria framework for anomaly detection in univariate time series. By integrating various statistical and machine learning techniques, the study aims to enhance the accuracy and robustness of anomaly identification. This approach seeks to address the challenges posed by complex and dynamic datasets, providing a flexible methodology suitable for diverse applications.
Design/methodology/approach: The study employs a multi-criteria approach to anomaly detection in univariate time series. It integrates statistical methods, such as boxplots and deviation-based rules, with machine learning techniques like clustering (hierarchical and k-means). The framework includes three aggregation strategies: restrictive, liberal, and scoring-based, to evaluate anomalies based on different criteria. The methodology is demonstrated using synthetic time series data that incorporates trends, seasonality, noise, and controlled anomalies.
Findings: The study reveals significant differences in the performance of various anomaly detection methods applied to univariate time series. Restrictive approaches provide high specificity, minimizing false positives, while liberal methods are more inclusive but prone to false alarms. The scoring-based approach offers a balanced evaluation, enabling quantification of anomaly significance across multiple criteria. The results demonstrate that combining statistical and machine learning methods enhances detection precision. The proposed multi-criteria framework is adaptable to diverse applications, though further validation on real-world datasets is required to confirm its effectiveness and scalability.
Originality/value: This study introduces a novel multi-criteria framework for anomaly detection in univariate time series, combining statistical and machine learning techniques with aggregation strategies to enhance detection accuracy. Unlike existing approaches, it systematically integrates multiple criteria and evaluates their collective impact on anomaly identification. The framework provides flexibility through restrictive, liberal, and scoring-based methods. Its originality lies in the methodological synthesis and the potential to address complex challenges in anomaly detection across various domains, offering valuable insights for both researchers and practitioners.
Keywords: anomaly detection, outlier detection, univariate time series, multi-criteria analysis.
Category of the paper: research paper, technical paper.

http://dx.doi.org/10.29119/1641-3466.2024.213.46 http://managementpapers.polsl.pl/

1. Introduction

Anomaly detection is critically important across various domains (Mehrotra et al., 2017), as it enables the identification of atypical, potentially critical events or patterns in data. Such events can signal problems, risks, or even new, unexpected opportunities. The importance of anomaly detection can be considered in several key aspects (Box et al., 2015; Chandol et al., 2009; Mills, 2019; Nielsen, 2019), including:

- Preventing and detecting problems: Identifying equipment failures, irregularities in device performance, network disruptions, early warnings in critical systems, deviations in sensor readings (e.g., in industry, energy, aviation), data protection, and cybersecurity.
- Supporting decision-making processes: Detecting operational inefficiencies, analyzing production or logistics processes to identify anomalies that lead to time or resource losses.
- Managing finance and risk: Detecting financial fraud, unusual transactions, deviations from the planned budget, or anomalies in stock market operations.
- Enhancing safety and quality: Identifying deviations in patients' health parameters (medical field), anomalies in navigation systems, autonomous vehicles, or railways to prevent accidents (transport), as well as improving data quality by detecting errors, missing data, and irregular values.

Anomaly detection plays a pivotal role in improving efficiency, safety, and quality across nearly every aspect of life and business. Rapid identification of deviations enables better decision-making, risk minimization, and the enhancement of business data value. In the context of growing data volume and complexity, automation and fast anomaly detection algorithms have become indispensable tools for modern organizations (Kao, Jiang, 2019).

From a technical perspective, anomaly detection in time series involves identifying data points that significantly deviate from expected patterns, trends, or seasonality observed in historical data. This may include extreme values (outliers), sudden trend changes, unexpected seasonal patterns, or deviations in value distributions (Braei Wagner, 2020; Chandola et al., 2009). The general principle is to identify patterns or a theoretical model of the series and then examine the residual distribution (deviations from the pattern or errors). The core idea is illustrated in Figure 1, which demonstrates time series decomposition using the STL method: Seasonal-Trend Decomposition using Loess (Cleveland et al., 1990).

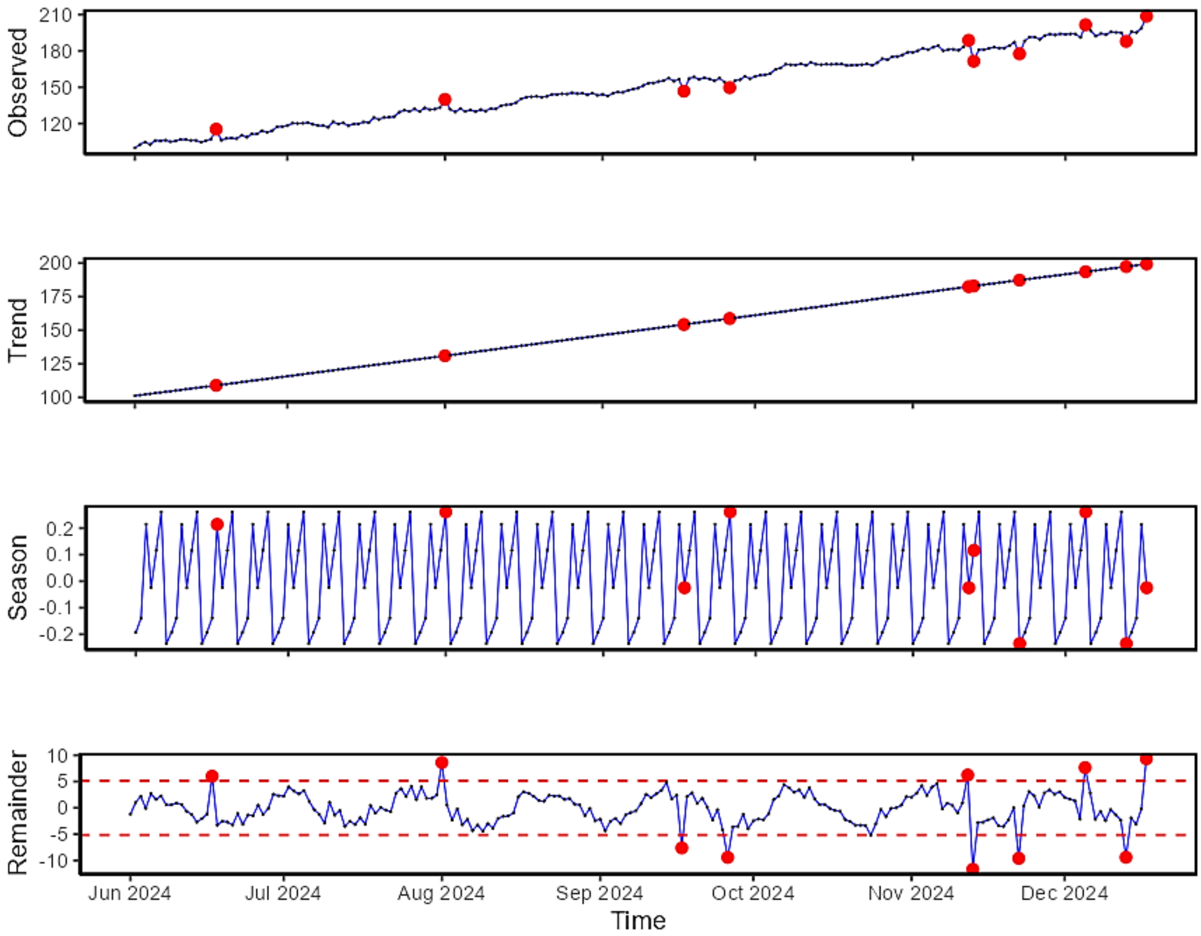

Figure 1. Presentation of identified anomalies (red points) on the decomposition plot of the time series. Source: own elaboration.

Figure 1 illustrates the additive decomposition of a time series into four main components: Observed, Trend, Season, and Remainder (Observed = Trend + Season + Remainder). Each plot highlights the analyzed series with a blue line, with anomalies marked as red points. The Remainder component represents the residual values, which are not explained by the systematic patterns of the series: trend and seasonality. The fundamental idea behind anomaly detection is to analyze the Remainder component based on statistical criteria. The basis lies in the variability of the series and deviations from the average error level. Outlier detection methods are most commonly used, with the anomaly region located outside the range defined by the dashed lines on the Remainder plot.

Outliers are identified based on various distance-based criteria, with the most commonly used methods including:

- basic statistical methods, such as boxplot-based approaches (IQR test), standard deviation thresholds, and statistical tests (e.g., Grubbs' test, Dixon's test),
- quantitative methods: a fixed number (or fraction) of the most deviating values,

- clustering-based criteria: deviations are identified as clusters (groups) with the smallest sizes, utilizing unsupervised learning mechanisms,
- machine learning-based criteria: approaches leveraging various supervised learning algorithms.

Although various metrics are used for anomaly detection, relatively few studies focus directly on a multi-criteria approach. Two studies indexed in the Scopus database explicitly refer to anomaly detection and a multi-criteria approach in their titles (Hsiao et al., 2015; Zafari et al., 2022). In Zafari et al. (2022), different anomaly metrics are evaluated to establish a final threshold defining anomalies. This study focuses on detecting irregularities in the circulation of prescriptions in the pharmaceutical market and identifying potential fraud. The final assessment is based on three measures related to the Gini index and drug (Opioid) scoring. The study by Hsiao et al. (2015) presents an anomaly detection method based on multi-criteria similarity measures and introduces a new metric – Pareto depth. This study proposes the Pareto Depth Analysis (PDA) method for anomaly identification.

In other bibliographic databases, relatively few studies reference a multi-criteria approach to anomaly detection. Ribeiro et al. (2020) proposed a multi-criteria approach to detect anomalies in drinking water quality, aiming to balance false alarms and missed detections during anomaly occurrences. Multi-criteria methods were also employed to evaluate machine learning algorithms for network access classification in the context of anomaly detection (Nascimento, Santos, 2022). Similarly, Wu et al. (2021) presented a multi-criteria approach to identifying anomalies in the energy systems of the iron and steel industry. Additionally, Dauwe et al. (2014) utilized a multi-criteria quality assessment model for noise monitoring networks to automatically detect anomalies in measurement data.

The purpose of this study is to present and propose a multi-criteria approach, along with the aggregation of these criteria, to provide a quantified assessment of observations that are potential anomalies.

2. Methods

2.1. Incorporating Multiple Criteria in Anomaly Detection

The proposed approach incorporates multiple criteria (methods, techniques) for anomaly detection. A method for aggregating anomaly detection results is introduced in three variations:

1. restrictive aggregation (restrictive aggregation) – an anomaly is defined as a value identified by all the considered criteria,
2. liberal aggregation (liberal detection) – an anomaly is defined as a value identified by any of the considered criteria,

3. aggregation by summing criteria indicating anomalies (scoring-based aggregation) – a potential anomaly is assigned a score based on the number of criteria under which the observation in the series is classified as an anomaly.

The framework of the proposed approach is illustrated in Figure 2.

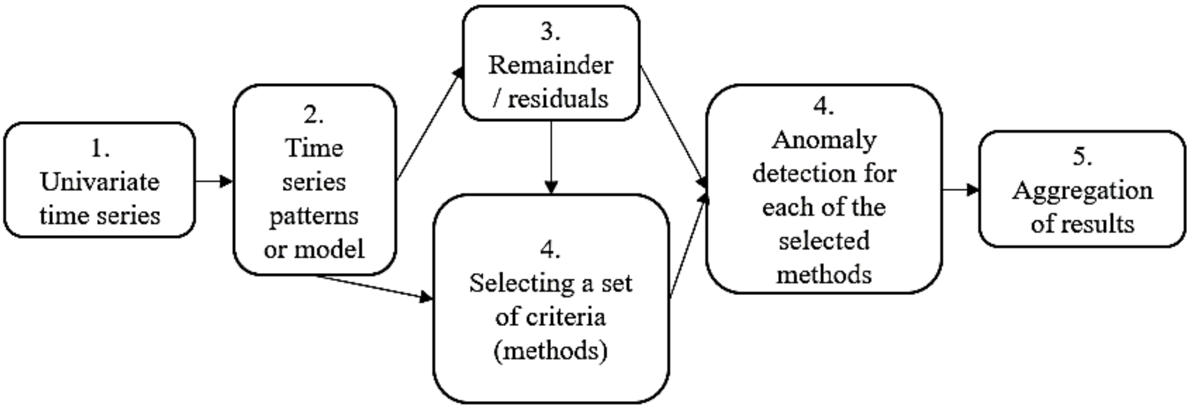
1. Univariate time series
2. Time series patterns or model
3. Remainder / residuals
4. Selecting a set of criteria (methods)
4. Anomaly detection for each of the selected methods
5. Aggregation of results

Figure 2. Process diagram of anomaly detection in univariate time series using multiple criteria.

Source: own elaboration.

Figure 2 illustrates the process diagram for anomaly detection in univariate time series. The process begins with the analysis of the time series (1) and the identification of patterns or the development of a time series model (2). Subsequently, after removing the identified patterns, the residual component (3) is obtained and analyzed for anomalies. For this purpose, a set of criteria (4) is selected, enabling anomaly detection for each chosen approach. Finally, the analysis results are aggregated (5) to provide a final assessment of the anomalies.

The fundamental premise of the proposed approach is the identification of anomalies within a fixed time window, represented by the entire analyzed time series. Consequently, this process can be implemented in more complex anomaly detection systems.

A more detailed description of the proposed analysis is presented using a numerical example.

2.2. Detailed description of the proposed approach using a numerical example

2.2.1. Generating Test Data for Anomaly Detection in Time Series

To demonstrate the proposed approach for anomaly detection, a time series was generated as an example for testing anomaly detection algorithms. This series combines characteristics of real-world data (trend, seasonality, noise) with controlled deviations, enabling the evaluation of method effectiveness in a controlled environment.

First, a linear trend was defined, introducing a systematic increase in values over time. White noise and random changes from a uniform distribution were added to simulate natural variability. Next, a periodic component was introduced to incorporate cyclical patterns. Subsequently, ten time points from the series were randomly selected and modified by adding or subtracting anomaly values calculated as a multiple of the original series values.

This modification aimed to simulate deviations from the pattern, representing anomalies in the data. The simulated time series, along with the generated deviations from the original series, is shown in Figure 3.

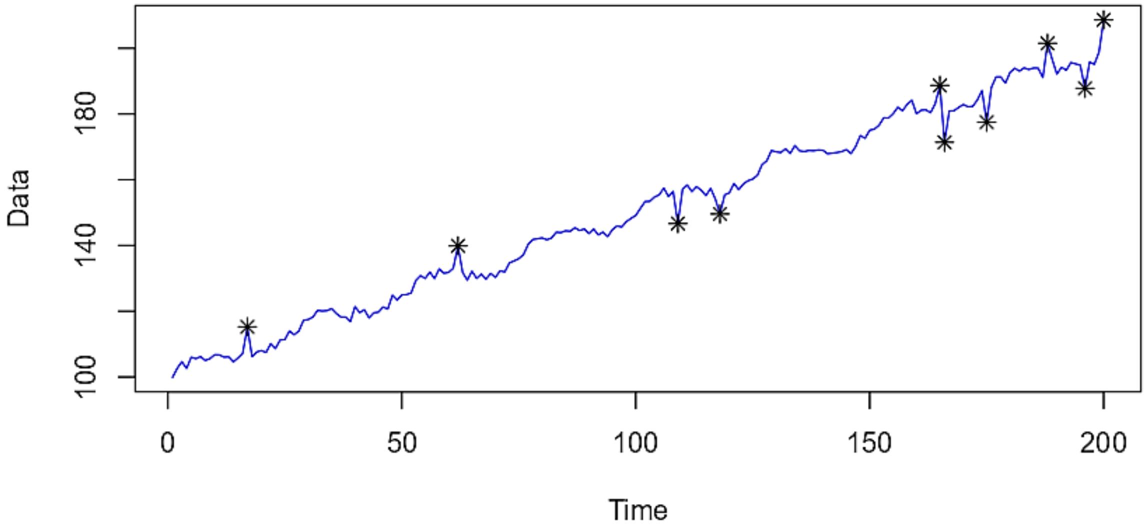

Figure 3. Visualization of a time series with marked deviations from the pattern.

Source: own elaboration.

2.2.2. Selected Methods for Time Series Modeling and Anomaly Detection

Eight methods were selected for anomaly detection, combining time series modeling with anomaly identification techniques. Time series representation utilized STL decomposition and the ARIMA model with automatic parameter selection (Hyndman, Khandakar, 2008; Wang et al., 2006). Anomaly detection on the residual component employed the following approaches:

- three-sigma rule (anomaly defined as deviations beyond 3 standard deviations): an anomaly is defined as a residual value deviating by more than three standard deviations,
- boxplot rule: an anomaly is a value that lies outside the whiskers of a classic boxplot, where whiskers extend up to 1.5 times the interquartile range,
- hierarchical clustering into four clusters: anomalies are values belonging to the two clusters with the smallest proportion of total observations (it is assumed that one cluster represents extremely negative values, the other represents extremely positive values, while the remaining two clusters contain typical values).
- K-means clustering into four clusters: similar to hierarchical clustering, anomalies are values belonging to the two clusters with the smallest proportion of total observations.

Based on this, the selected methods can be specifically defined as criteria for anomaly identification:

- Method 1: time series modeled with ARIMA, anomalies identified as observations deviating from the mean residuals by more than three standard deviations.
- Method 2: time series modeled with ARIMA, anomalies identified as outliers based on a classic boxplot.
- Method 3: time series modeled with ARIMA, anomalies identified as observations extracted using hierarchical clustering.
- Method 4: time series modeled with ARIMA, anomalies identified as observations extracted using k-means clustering.
- Method 5: time series decomposed using STL, anomalies identified as observations deviating from the mean residuals by more than three standard deviations.
- Method 6: time series decomposed using STL, anomalies identified as outliers based on a classic boxplot.
- Method 7: time series decomposed using STL, anomalies identified as observations extracted using hierarchical clustering.
- Method 8: time series decomposed using STL, anomalies identified as observations extracted using k-means clustering.

A common feature of the methods is that they are independent of the assumed significance level and do not account for the maximum or minimum number of identified anomalies.

The relationships between time series models and anomaly identification methods applied to the residual component are presented in Figure 4.

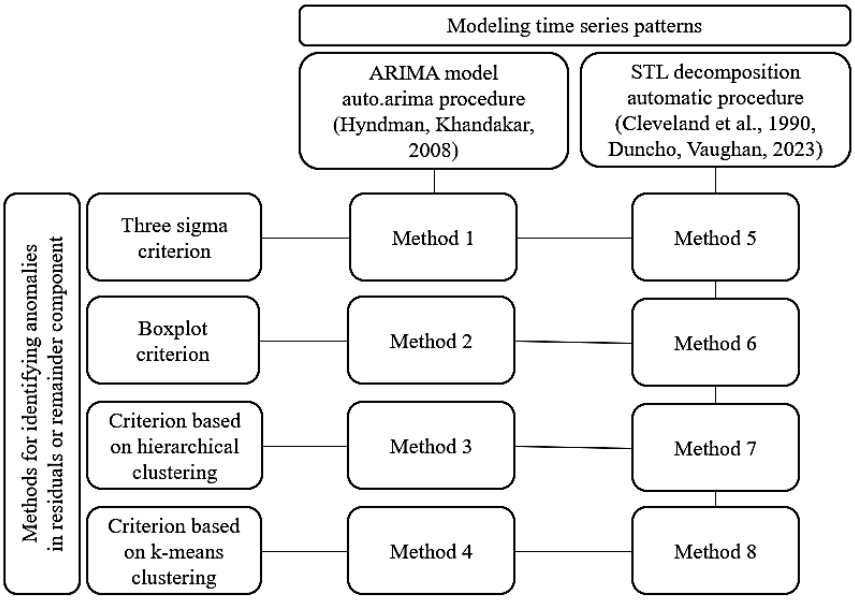
Modeling time series patterns
ARIMA model auto.arima procedure (Hyndman, Khandakar, 2008)
STL decomposition automatic procedure (Cleveland et al., 1990, Duncho, Vaughan, 2023)
Methods for identifying anomalies in residuals or remainder component
Three sigma criterion
Boxplot criterion
Criterion based on hierarchical clustering
Criterion based on k-means clustering
Method 1
Method 2
Method 3
Method 4
Method 5
Method 6
Method 7
Method 8

Figure 4. Approach for anomaly detection in time series – combining modeling and identification methods.

Source: own elaboration.

2.2.3. Aggregation of results

The aggregation of results is based on the application of three approaches:

- Restrictive approach: a value in the series is considered an anomaly only if it is identified by all the considered methods (criteria).
- Liberal approach: a value in the series is considered an anomaly if it is identified by at least one method.
- Scoring approach: for each value in the series identified as an anomaly, the number of methods (criteria) recognizing it as an anomaly is summed.

The restrictive approach results in the smallest set of values classified as anomalies, while the liberal approach produces the largest set. The scoring approach provides a quantified evaluation of values deviating from the pattern, which can serve as a foundation for further studies and analyses.

Ultimately, the time series receives a scoring evaluation of potential anomalies alongside the extreme approaches to anomaly detection.

3. Results

The results of anomaly identification for each method are presented in Figures 5-12.

In each figure, the blue line represents the time series under analysis, while the red line shows the predicted (fitted) values. Artificially generated deviations are marked with asterisks, and anomalies identified by the methods are depicted as red filled circles.

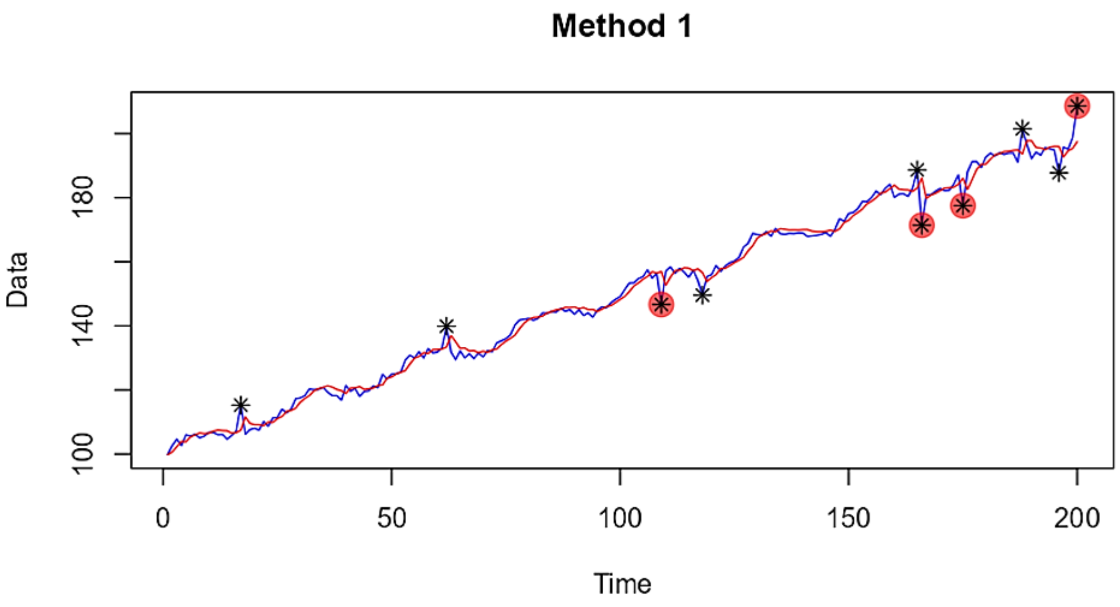

Figure 5. Presentation of identified anomalies using method 1 (ARIMA model with 3-sigma rule).
Source: own elaboration.

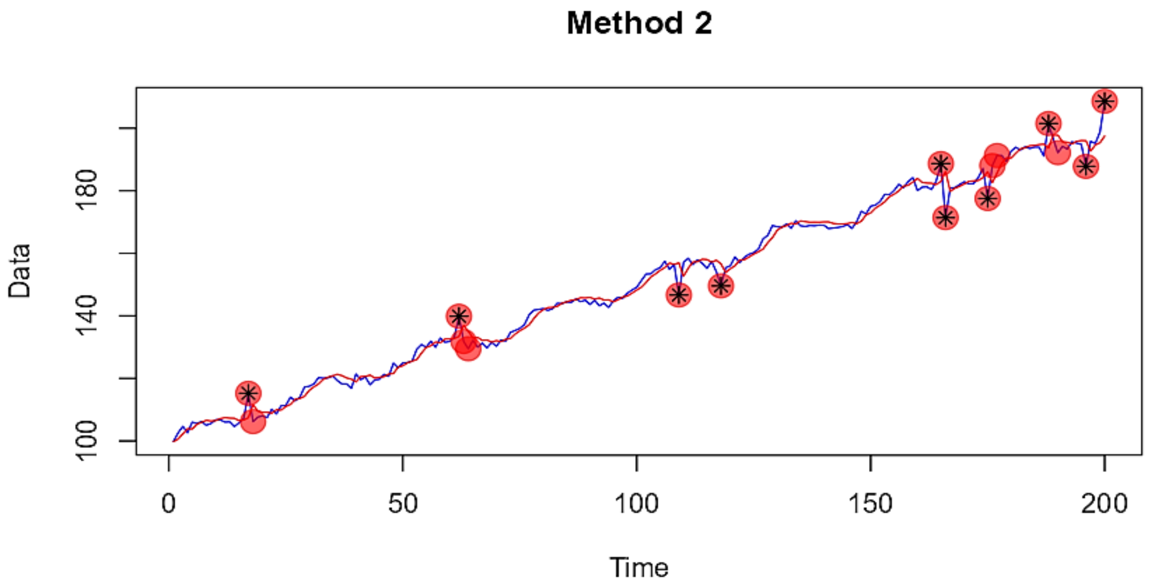

Figure 6. Presentation of identified anomalies using method 2 (ARIMA model with boxplot rule).

Source: own elaboration.

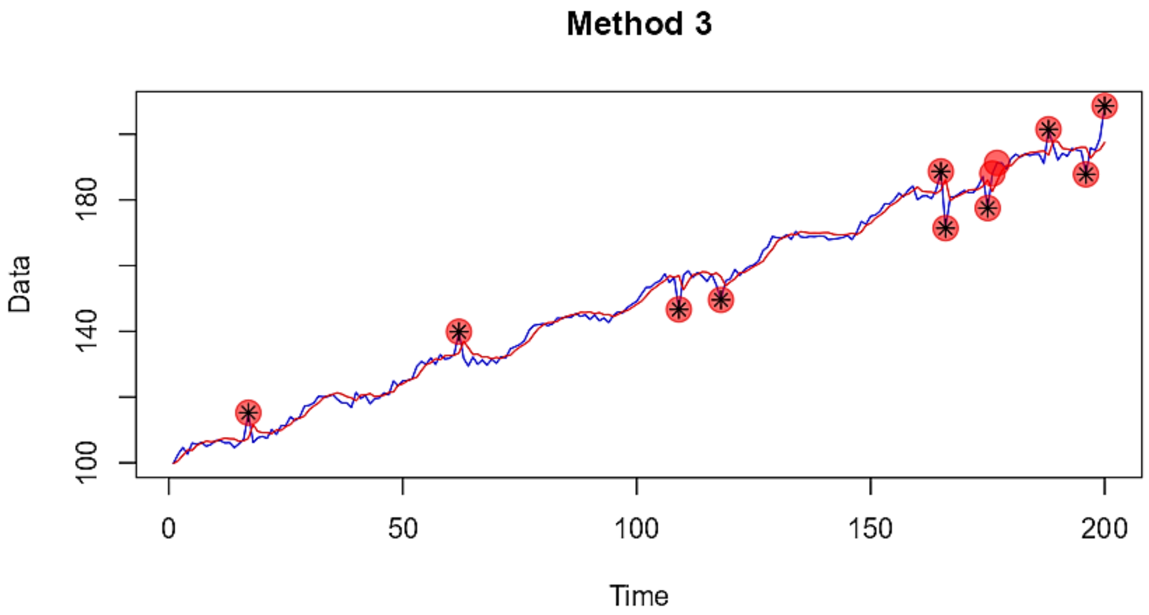

Figure 7. Presentation of identified anomalies using method 3 (ARIMA model with hierarchical clustering rule).

Source: own elaboration.

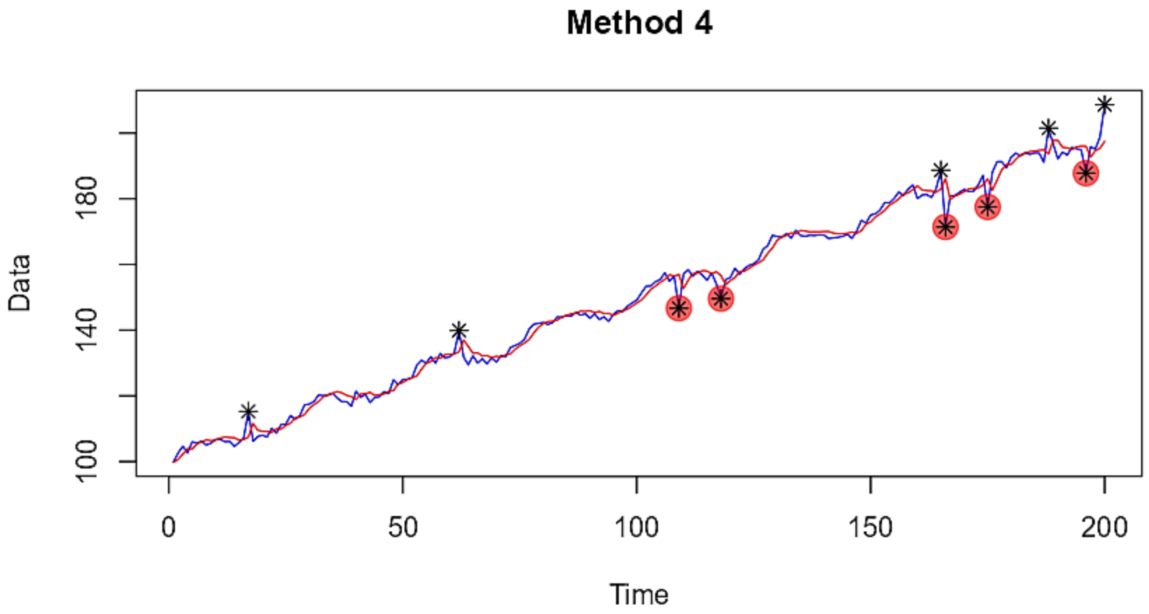

Figure 8. Presentation of identified anomalies using method 4 (ARIMA model with k-means clustering rule).

Source: own elaboration.

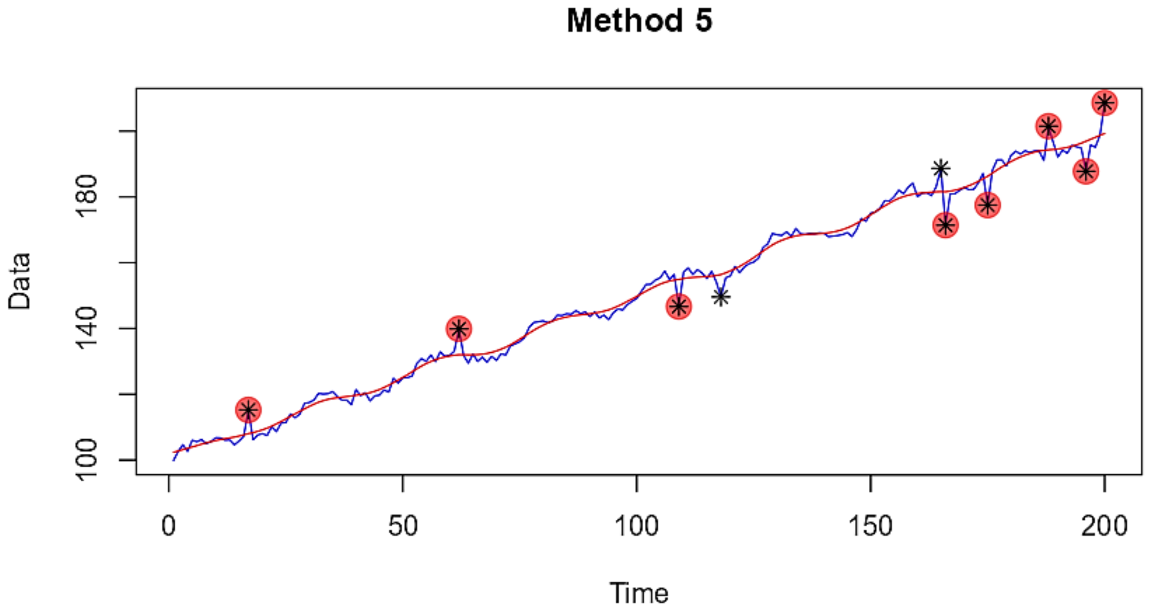

Figure 9. Presentation of identified anomalies using method 5 (STL model with 3-sigma rule).

Source: own elaboration.

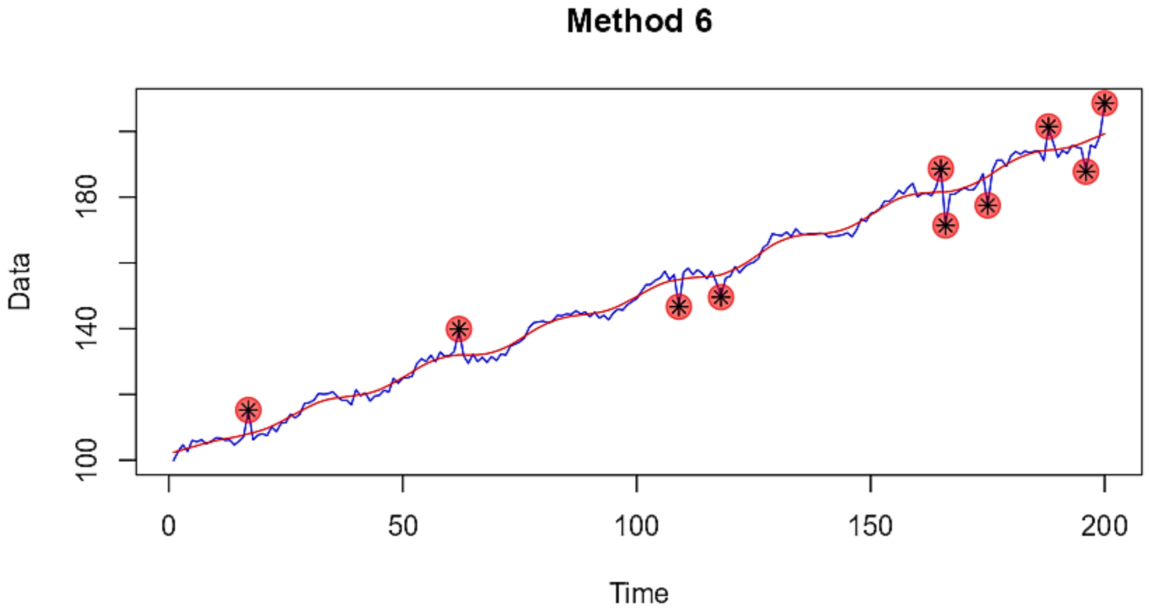

Figure 10. Presentation of identified anomalies using method 6 (STL model with boxplot rule).

Source: own elaboration.

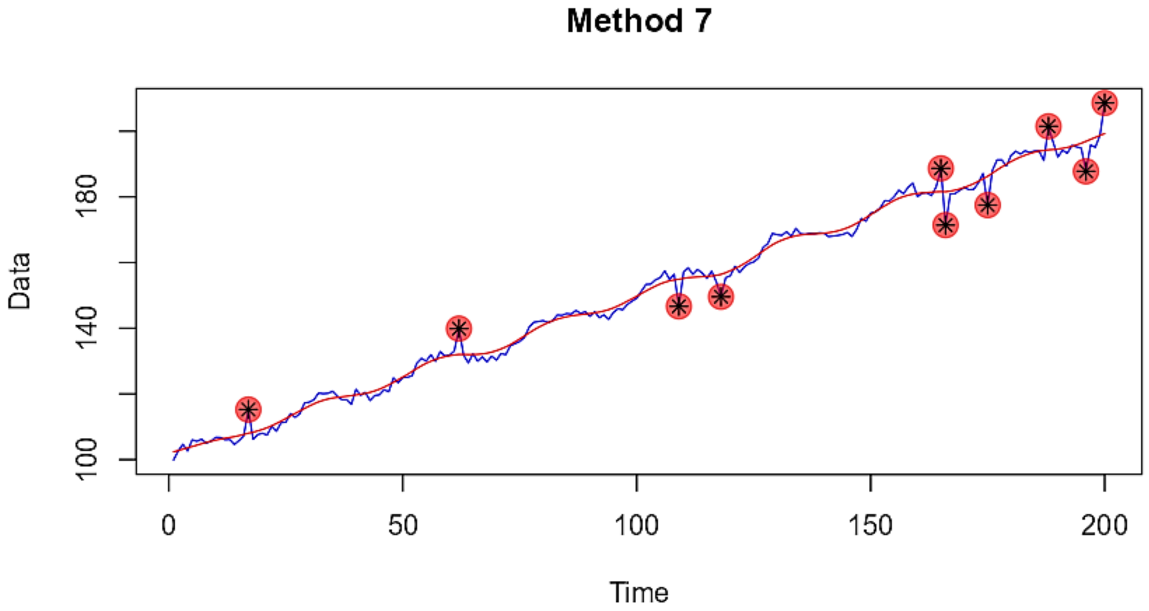

Figure 11. Presentation of identified anomalies using method 7 (STL model with hierarchical clustering rule).

Source: own elaboration.

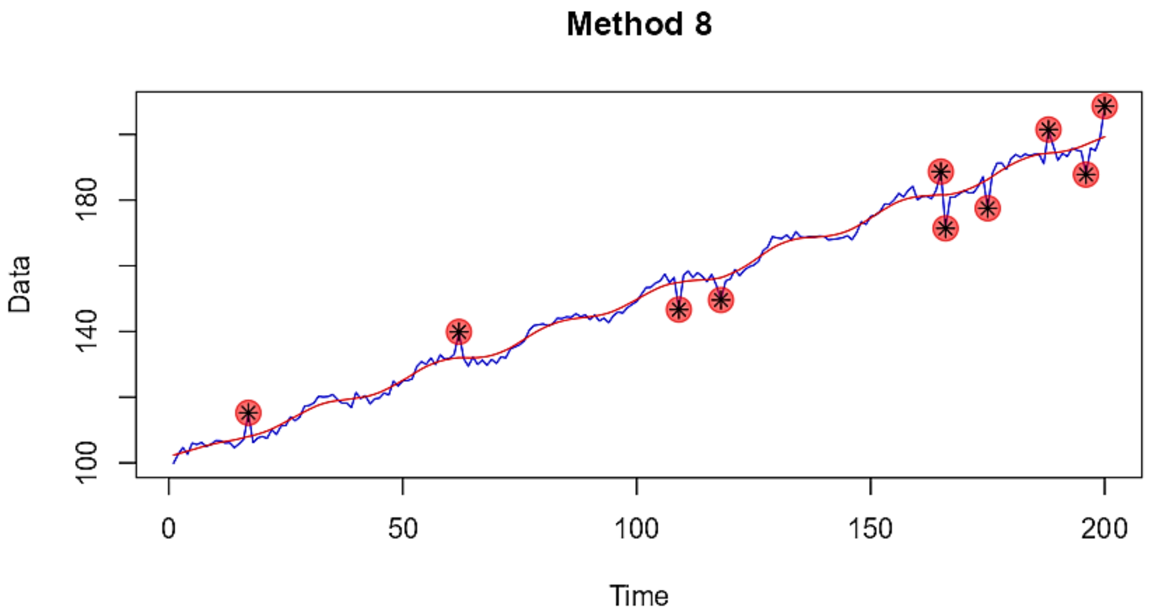

Figure 12. Presentation of identified anomalies using method 8 (STL model with k-means clustering rule).

Source: own elaboration.

The results of aggregated anomaly detection using the restrictive approach are presented in Figure 13. Figure 14 shows the results for the liberal approach, while Figure 15 illustrates the scoring approach.

Additionally, in Figure 15, each identified point is annotated with the number of methods by which that point was classified as an anomaly. The size of the circles reflects this number, meaning the larger the circle, the stronger the justification for classifying the observation as an anomaly.

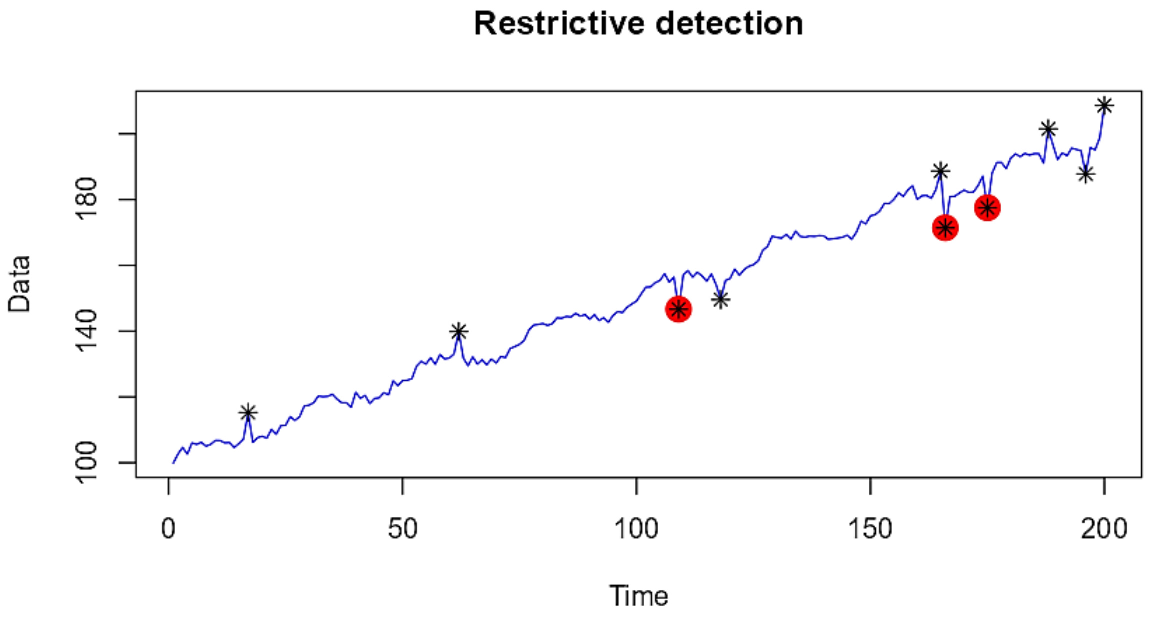

Figure 13. Identification of anomalies using the restrictive approach.

Source: own elaboration.

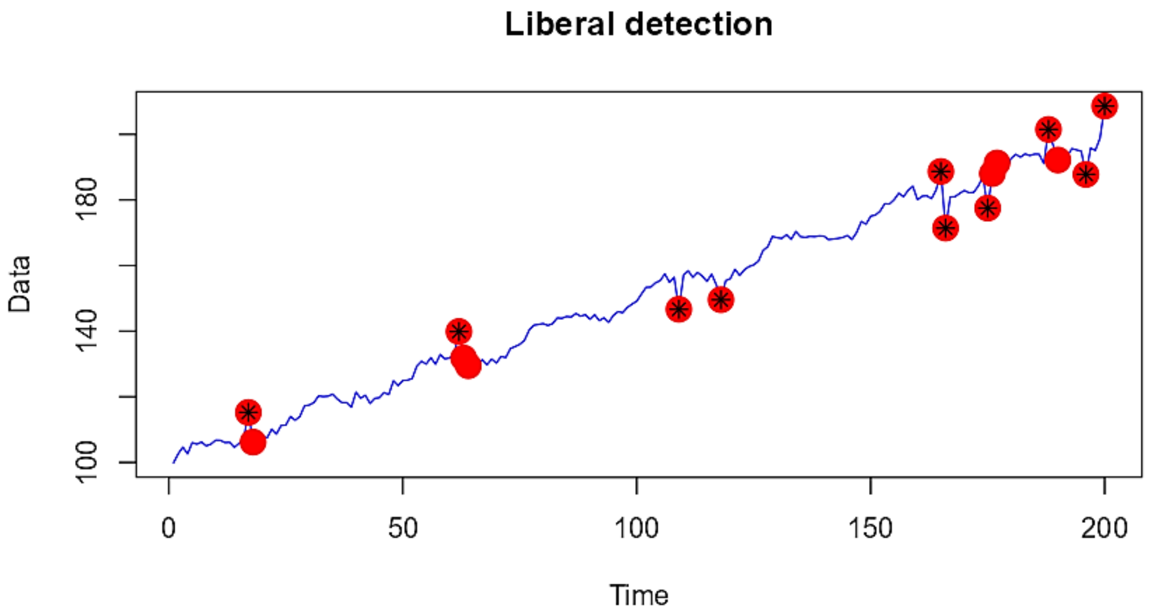

Figure 14. Identification of anomalies using the liberal approach.

Source: own elaboration.

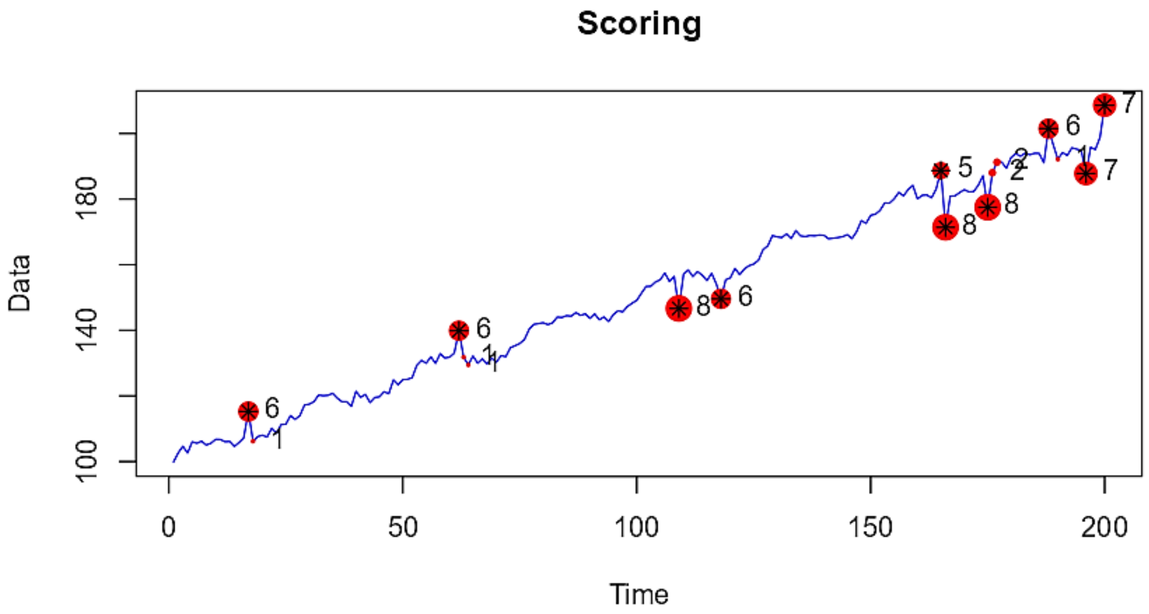

Figure 15. Identification of anomalies using the scoring approach.

Source: own elaboration.

Using the restrictive approach, only three observation was identified as an anomaly. The liberal approach resulted in the largest number of observations classified as anomalies. The scoring approach provided a quantified assessment of the identified anomalies.

4. Discussion

The presented results highlight key differences among various anomaly detection methods. The applied approaches (restrictive, liberal, and scoring-based) offer diverse strategies for detecting deviations, each suitable for specific problem contexts.

The restrictive approach ensures high specificity, reducing false alarms, but at the potential cost of missing significant anomalies. This makes it advantageous in scenarios where minimizing false positives is critical, such as in safety-critical systems. In contrast, the liberal approach, while more inclusive, may generate an excess of false positives, limiting its applicability for analyzing large datasets where precision is paramount. The scoring approach provides a balanced compromise, quantifying the significance of observed deviations, which can be valuable in decision-making processes.

It is worth noting that the results vary depending on the modeling method used (ARIMA, STL decomposition) and the anomaly detection criteria. The agreement observed in the outcomes of Methods 7 and 8 suggests they may be more universally applicable in the studied context, though further research is necessary.

One limitation of this study is the lack of evaluation of the methods' effectiveness in real-world applications, representing an important area for future research. A next step could involve a comparative analysis of the proposed approach’s efficiency across various domains, such as industry or finance.

Another crucial aspect is that time series anomaly detection frequently involves sliding time windows and artificial intelligence methods (Kao, Jiang, 2019; Lu et al., 2023). Future research could incorporate the proposed methodology into analyses based on sliding time windows. Expanding the set of methods may also enhance the precision of quantifying anomalies in time series.

The proposed approach may also prove useful for examining forecasting error series, which can exhibit certain patterns, where patterns and significant deviations from them might be classified as anomalies (Wolny, 2023).

5. Summary

This paper explores anomaly detection in time series through a multi-criteria approach, addressing the increasing complexity and volume of data in modern applications. Anomaly detection plays a vital role across various domains, such as industry, finance, healthcare, and cybersecurity, by identifying critical deviations that may indicate potential risks or opportunities. The study focuses on integrating multiple criteria to enhance anomaly detection precision, utilizing both statistical and machine learning techniques.

The proposed methodology includes three aggregation strategies for anomaly identification: restrictive, liberal, and scoring-based approaches. These strategies balance specificity and sensitivity, catering to diverse application requirements. The restrictive approach emphasizes minimizing false positives, while the liberal approach prioritizes inclusivity. The scoring method provides a quantifiable assessment of potential anomalies, combining insights from multiple detection criteria.

The results demonstrate significant differences between detection methods, with models based on ARIMA and STL decomposition exhibiting varied performance. A comparative analysis of methods highlights the potential universality of some approaches, such as those involving clustering algorithms. However, further research is required to validate these findings across real-world datasets and dynamic applications, including sliding time windows and advanced artificial intelligence techniques.

In conclusion, the study underscores the importance of adopting a multi-criteria perspective in anomaly detection to achieve more robust, adaptable, and precise solutions. Future work should extend the proposed framework to include additional criteria and dynamic contexts, enhancing its applicability and reliability in practical scenarios.

References

1. Box, G.E., Jenkins, G.M., Reinsel, G.C., Ljung, G.M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
2. Braei, M., Wagner, S. (2020). Anomaly detection in univariate time-series: A survey on the state-of-the-art. *arXiv preprint arXiv:2004.00433*.
3. Chandola, V., Banerjee, A., Kumar, V. (2009). Anomaly detection: A survey. *ACM computing surveys (CSUR)*, *41(3)*, pp. 1-58.
4. Cleveland, R.B., Cleveland, W.S., McRae, J.E., Terpenning, I. (1990). STL: A seasonal-trend decomposition. *J. off. Stat*, *6(1)*, pp. 3-73.
5. Dancho, M., Vaughan, D. (2023). Anomalize: Tidy anomaly detection. *R package version 0.3.0*. Retrieved from: https://CRAN.R-project.org/package=anomalize.
6. Dauwe, S., Oldoni, D., De Baets, B., Van Renterghem, T., Botteldooren, D., Dhoedt, B. (2014). Multi-criteria anomaly detection in urban noise sensor networks. *Environmental Science: Processes & Impacts*, *16(10)*, pp. 2249-2258.
7. Hsiao, K.J., Xu, K.S., Calder, J., Hero, A.O. (2015). Multicriteria similarity-based anomaly detection using Pareto depth analysis. *IEEE transactions on neural networks and learning systems*, *27(6)*, pp. 1307-1321.
8. Hyndman, R.J., Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. *Journal of statistical software*, *27*, pp. 1-22.

9. Kao, J.B., and Jiang, J.R. (2019). Anomaly detection for univariate time series with statistics and deep learning. *IEEE Eurasia conference on IOT, communication and engineering (ECICE, October 2019),* pp. 404-407.
10. Lu, T., Wang, L., Zhao, X. (2023). Review of anomaly detection algorithms for data streams. *Applied Sciences*, *13(10), 6353.*
11. Mehrotra, K.G., Mohan, C.K., Huang, H. (2017). *Anomaly Detection Principles and Algorithms. Terrorism, Security, and Computation.* Springer International Publishing, doi:10.1007/978-3-319-67526-8
12. Mills, T.C. (2019). *Applied time series analysis: A practical guide to modeling and forecasting.* Academic press. Berlin: Springer.
13. Nascimento, G.B. Santos, M. (2022). Performance evaluation of machine learning algorithms for network anomaly detection: an approach through the AHP-TOPSIS-2N method. *Procedia Computer Science, 214*, pp. 164-171.
14. Nielsen, A. (2019). *Practical time series analysis: Prediction with statistics and machine learning*. O'Reilly Media.
15. Ribeiro, V.H.A., Moritz, S., Rehbach, F., Reynoso-Meza, G. (2020). A novel dynamic multi-criteria ensemble selection mechanism applied to drinking water quality anomaly detection. *Science of The Total Environment, 749, 142368.*
16. Wang, X., Smith, K., Hyndman, R. (2006). Characteristic-based clustering for time series data. *Data mining and knowledge Discovery*, *13, 335-364.*
17. Wolny, M. (2023). A Decomposition Study of the Time Series of Electricity Consumption Forecasting Errors. *Zeszyty Naukowe Organizacja i Zarządzanie, 171.* Politechnika Śląska, pp. 173-182.
18. Wu, H., Jin, F., Zhao, J., Wang, W. (2021). Anomaly detection method based on multi-criteria evaluation for energy data of steel industry. *IEEE 10th Data Driven Control and Learning Systems Conference (DDCLS)*, pp. 741-747.
19. Zafari, B., Ekin, T., Ruggeri, F. (2022). Multicriteria decision frontiers for prescription anomaly detection over time. *Journal of Applied Statistics, 49(14)*, pp. 3638-3658.