PRODUCTION ENGINEERING ARCHIVES 2026, 32(3), 473-485

PRODUCTION ENGINEERING ARCHIVES

ISSN 2353-5156 (print)
ISSN 2353-7779 (online)

Exist since 4th quarter 2013
Available online at https://pea-journal.eu

Exploratory analysis of variability in municipal waste collection using RFID-based operational data: A case study from the Slovak Republic

Roman Lacko[1*], Radúz Dula[1], Martin Kuchta[2], Zuzana Hajduová[3], Marcin Zawada[4], Patrik Břečka[5]

[1] Bratislava University of Economics and Business, Faculty of Commerce, Department of Tourism, Dolnozemská cesta 1, 852 35, Bratislava, Slovak Republic; roman.lacko@euba.sk (RL); raduz.dula@euba.sk (RD)
[2] Bratislava University of Economics and Business, Faculty of Commerce, Department of Marketing, Dolnozemská cesta 1, 852 35, Bratislava, Slovak Republic; martin.kuchta@euba.sk
[3] Bratislava University of Economics and Business, Faculty of Business Management, Department of Business Finance, Dolnozemská cesta 1, 852 35, Bratislava, Slovak Republic; zuzana.hajduova@euba.sk
[4] Częstochowa University of Technology, Faculty of Management, Dąbrowskiego 69, 42-201 Częstochowa; marcin.zawada@pcz.pl
[5] Asseco Central Europe, Galvaniho 19, 821 04 Bratislava, Slovakia; patrik.brecka@asseco-ce.com
*Correspondence: roman.lacko@euba.sk

Article history
Received 20.03.2026
Accepted 01.06.2026
Available online 30.06.2026

Keywords
waste management,
variability,
optimisation,
Slovak Republic,
waste predictability.

Abstract
The objective of this case study is to identify and assess the potential for reducing variability in municipal waste collection in a selected town. The research was conducted in Michalovce, a town located in eastern Slovakia. In 2024, the municipality introduced weighing of individual municipal waste collections using modern collection vehicles and RFID-based identification, creating a detailed operational database of collection events at individual sites. Based on this dataset, we examined the variability in waste amounts at both the collection-event and collection-site levels. The analysis confirmed significant relationships between the number of containers and the amount of waste collected, and between collection frequency and total waste volume. At the same time, the results revealed operational inefficiencies, suggesting that some collection schedules could be better aligned with actual waste generation patterns. The methodology and findings can be applied directly in the town under study and may also serve as a useful reference for other municipalities in Slovakia and neighbouring countries with similar waste collection systems.

DOI: 10.30657/pea.2026.32.39

1. Introduction

The collection of municipal waste is a key aspect of human existence, as it contributes not only to improving people's quality of life but also to enhancing environmental quality. Insufficient management of municipal waste can create unhygienic conditions and, ultimately, lead to increased healthcare costs (Erçin et al., 2021). These effects are particularly associated with population growth in urbanised areas, where economic development is closely linked to increased municipal waste generation (Rattanawai et al., 2024). In developed countries, municipal waste collection is continually being made more efficient to maximise effectiveness while minimising costs, as it represents a significant portion of municipal expenditures (Filipiak et al., 2009).

Currently, with advances in information and communication technologies, artificial intelligence is becoming a key factor in waste collection decision-making. These decisions are increasingly based on multi-criteria decision models, which are often built on thoroughly developed management systems such as ISO 9001, ISO 14001, and OHSAS 18001 (Mendes et al., 2013; Krzywda & Krzywda, 2026). More broadly, the effective use of RFID-based operational data should also be considered in the context of digital sustainability maturity, as digital maturity, data-management capabilities, digital competencies and process integration create organisational conditions for transforming operational data into sustainability-oriented decision-support knowledge (Ingaldi & Ulewicz, 2026).

2. Literature review

The variability in the amount of waste generated is a fundamental characteristic and a widely discussed topic in the optimisation of municipal waste collection processes (Neto et al., 2024). Reducing this variability has a crucial impact on the predictability of waste amounts, which in turn helps optimise the cost side of municipal waste collection. Collection costs account for 70% to 85% of total waste management expenses (Apaydin & Gonullu, 2007; Hemidat et al., 2017). According to Neto et al. (2024), the determinants of optimisation can be divided into three groups:

1. Waste generation – Fill level sensors, Bin capacity, Bin location
2. Collection and transportation – Frequency, Route, Consume, Distance, Costs, Emissions
3. Disposal – Capacity, Location.

This optimisation process is called the Waste Collection Vehicle Routing Problem (WCVRP) (Han & Cueto, 2015). Numerous determinants influence the variability of waste amounts. One of these is seasonality, which encompasses several factors, including tourism, population mobility, socio-economic characteristics, holidays, and exceptional events (such as COVID-19) (Mendes et al., 2013; Richter et al., 2021). Another factor affecting the amount and variability of municipal waste and its collection can be regional differences, which are often overlooked in the literature but are very important (Lu et al., 2017; Przydatek & Ciągło, 2020). Additionally, factors such as temperature, the nature of the waste, biological risk, and similar factors play a role in municipal waste collection strategies (Rattanawai et al., 2024). Recent Slovak evidence also shows that household solid-waste eco-efficiency differs across NUTS-3 regions and that demographic and contextual characteristics can support the selection of predictors for AI-based waste-management models (Lacko et al., 2026).

Concepts and strategies focused on optimising municipal waste collection through modern hardware and software solutions have proven particularly effective. A fundamental prerequisite for these solutions is the Geographical Information System (GIS) (Abdallah et al., 2019). Nowadays, cities are implementing modern waste collection systems such as the Smart Collection System (SCS), which uses sensors (for example, RFID) to measure waste amounts and optimise collection routes based on many characteristics. According to Wu et al. (2020), the waste filling level is also important for optimising municipal waste collection—they recommend an optimal fill level of 60–80%. As Abdallah et al. (2019) report, in the United Arab Emirates, which is heavily influenced by tourism, the use of this concept reduced municipal waste collection costs by up to 19% and reduced the time vehicles spent collecting waste by 18% to 42%, contributing to a reduction in emissions by 5% to 22%. Significant savings were also achieved by introducing modern municipal waste collection management systems in the city of Trabzon, Turkey—route lengths were reduced by 4% to 59%, collection times by 14% to 65%, and costs by up to 24% (Apaydin & Gonullu, 2007). Similar benefits were found in a study from cities in Jordan, where municipal waste collection time was reduced by up to 30% (Hemidat et al., 2017). Moreover, investments in municipal waste collection management infrastructure have a relatively quick payback period (Erçin et al., 2021). The study by Das & Bhattacharyya (2015) also demonstrated that the use of appropriate mathematical methods can reduce route lengths by up to 30%. Methods inspired by insect behaviour, such as Ant Colony Optimisation, are also used (Xue & Cao, 2016). Neto et al. (2024) conducted a detailed analysis of the current state of scientific research, thoroughly outlining all the key advantages of active municipal waste collection management. Other important studies, such as the literature review by Beliën et al. (2014), should not be overlooked. In addition to optimisation methods, one way to reduce municipal waste collection costs is to deploy larger collection vehicles (Kinobe et al., 2015) or use waste transfer stations, which can also help reduce emissions in urbanised areas (Apaydin & Gonullu, 2008; Ghiani et al., 2021; Nematollahi et al., 2024; Sousa et al., 2024). However, when optimising routes, the "destination" of the collected waste by type is also crucial (Xue et al., 2015).

The availability of GPS sensors, RFID (Smart Trucks and Bins), advanced algorithms, GIS, software solutions such as MapInfo and Roadnet, IoT solutions, and machine learning leads to the creation of big data databases. These databases enable the extraction of rich contextual insights (Erçin et al., 2021; Shah et al., 2018; Lu et al., 2017) for decision-making in municipal waste collection, a task that can today largely be transferred from humans to AI. These measures and strategies are key to city development and long-term sustainability in waste management (Khan et al., 2024; Nematollahi et al., 2024). The use of advanced methods in municipal waste collection, enabled by the availability of reliable data (Barth et al., 2023), appears to be a significant determinant of reduced air pollution not only in urbanised areas. This approach is referred to as the cyber-physical waste collection model (Bányai et al., 2019).

In this study, the term data-driven analysis is used to describe the evaluation of RFID-based operational records from municipal waste collection, rather than big data in the strict technological sense. Based on the above-stated literature review, the objective of this study is to identify and assess the potential to reduce variability in waste amounts during the collection process in the selected town using statistical analysis.

3. Materials and methods

In this section, we will describe the methods used and the nature of the problem addressed in this case study.

3.1. Research object and methodology

The focus of this study is the waste management system of the town of Michalovce in eastern Slovakia. In 2024, the town began weighing municipal waste using modern collection vehicles and RFID chips. This innovation resulted in a large database that can be used to improve waste management processes in the town. The article is focused only on collective housing development areas. The population of these areas is 26,446 citizens. The total population of Michalovce is 34,310 inhabitants. Each waste container has a unique chip that is read when the container is emptied, and the weight of the waste removed is recorded. All waste containers are assigned to specific waste collection sites, and these collection sites are assigned to taxpayers. In this case, to individual apartment building entrances. Multiple entrances may be assigned to one waste stand (collection point). All households in a given entrance are always assigned to the same stand. There may be a different number of waste containers per waste stand. Typically, from 3 to 6 waste containers per waste stand. Michalovce provided us with this data for the project, in which it also participates as part of a research consortium. The location of Michalovce is shown in Figure 1. The data was fully anonymised.

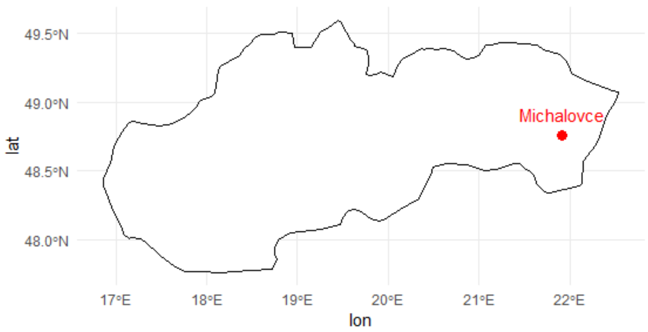

Fig. 1. Location of the town of Michalovce
Source: own processing in RStudio

3.2. Data

Based on the above, data were collected from 147 municipal waste collection sites. The municipal waste collection dataset originally contained 16,420 records. During preprocessing, the value 1 was treated as a technical error/status code generated during failed RFID-based weighing and was not interpreted as a valid measured mass of 1 kg. As a result, we analysed data from sites connected to complex apartment (flat) buildings, i.e., containers with a capacity of 1100 litres (individual houses with 240-litre containers did not participate in the measurement provided by Michalovce), covering a total of 15,273 waste collection events. The available dataset records waste mass per collection event, not the direct volumetric fill level. Therefore, the present analysis evaluates mass-based operational patterns rather than the true volumetric utilisation of 1100-litre containers. Cycle collection of mixed litter is carried out on the housing estate once, twice, or 3 times a week. The number of collections for each site is shown in Figure 2, which provides an overview of how collection frequency varies across all monitored locations. This visualisation helps to quickly identify sites with unusually high or low numbers of collections, setting the stage for deeper analysis of waste generation patterns and potential areas for operational optimisation.

In Figure 2, the individual municipal waste collection sites are divided into four main groups. The first group consists of sites with around 50 collections, followed by a second group, by far the largest, with approximately 90 collections. The third group includes sites with nearly 140 collections, and the fourth group comprises those with collection numbers oscillating around 150. The number of collection frequencies was set by the data collected in 2023. The data, specifically the waste amounts for each site, were subjected to further analysis. Table 1 summarises the analytical dataset used in this study, including the observation period, the number of collection sites and collection events, and the preprocessing context under which the anonymised data were provided to the authors.

Table 1. Descriptive characteristics of the dataset

Item	Value
Number of collection points	147
Number of collection events in the analytical dataset	15,273
Original number of records before removing RFID errors	16,420
Number of removed records	1,147
Share of removed records (%)	6.99
Share of retained records (%)	93.01
Average number of containers per collection point	3.44
Minimum number of containers per collection point	1
Maximum number of containers per collection point	12
Date range	2024-01-02 to 2024-12-31
Minimum waste mass after cleaning	10
Maximum waste mass after cleaning	1,441

Source: own computation

3.3. Methods

To achieve the objectives of this case study, we primarily used descriptive statistical analysis. We applied indicators such as means, standard deviations, coefficients of variation, and graphical data presentations, including boxplots and correlation graphs. To identify highly variable municipal waste collection sites, we used the following mathematical formula:

$$CV > \overline{CV} + SD_{CV} \,, \tag{1}$$

Where CV is the coefficient of variation, $\overline{CV}$ is the average coefficient of variation and SD_{CV} is the standard deviation of the coefficient of variation.

This formula states that sites with high variability are those whose coefficient of variation (CV) exceeds the sum of the average coefficient of variation $(\overline{CV})$ and its standard deviation (SD_{CV}) (Abbasi, 2020).

Moreover, we have decided to identify inefficient collection sites by examining containers with higher or lower mass. To do this, we calculated, for each stand, the average amount of waste collected per container during each collection event. We have computed this indicator as:

$$W_{CC} = \frac{TWC}{N\_Col \times N\_Con} \quad (2)$$

where W_{CC} is waste per collection per container, TWC is the total mass of waste collected, N_Col is the number of collections, and N_Con is the number of containers.

By comparing this value across all collection sites, we were able to segment the sites into three categories:

- Too frequent collections: Sites where the average waste mass per collection per container is significantly lower than most others.
- Efficient collections: Sites where the average waste per collection per container falls within a typical range, suggesting a well-balanced collection schedule.
- Too rare collections: Sites where the average waste per collection per container is much higher than most others, indicating that containers contain a higher amount of mass before being collected.

Because the dataset records waste mass rather than volumetric fill level, the categories should be interpreted as relative mass-based operational segments rather than direct evidence of underfilled or overfilled containers. To assess operational collection efficiency, we calculated the average amount of waste collected per container at each collection site during each collection event. This metric was computed as the total waste collected at a site divided by the product of the number of collections and the number of containers. By comparing this value across all collection sites, we segmented the sites into three categories. Sites below the 20th percentile were classified as having too frequent collections, indicating that containers were often not sufficiently filled at the time of collection. Sites above the 80th percentile were classified as having collections that were too rare, indicating a higher likelihood of higher mass in containers prior to collection. Sites between these thresholds were considered operationally efficient. This measure captures operational efficiency in terms of alignment between waste generation, container capacity, and collection frequency; however, it does not represent full-system efficiency because the available dataset does not include route costs, labour time, fuel consumption, service quality, or reliability indicators.

To complement the descriptive, variability, and correlation-based analyses, we estimated two simplified linear regression models at the collection-event level, where the dependent variable was the waste mass collected during a single collection event. The aim of these models was to assess whether waste mass per collection event was systematically associated with operational and temporal factors.

In the first model – Model A, waste mass was explained by the number of containers and the day of the week. This specification was designed to capture short-cycle operational regularities in the waste collection process. In the second model – Model B - waste mass was explained by the number of containers and month effects to capture broader seasonal variation over the calendar year. The reference category for the weekday effect is Friday, and for the month effect, January.

The models were estimated using ordinary least squares. Since the response variable is measured at the event level, the regression results should be interpreted as supportive inferential evidence rather than as strict causal estimates. Models do not include random effects for collection sites, number of residents per site, route-level characteristics, labour time, fuel consumption, or service-quality indicators.

4. Results and discussion

In this section, we present the results of the analyses used to identify waste collection sites where optimisation is most necessary. Figure 3 displays a histogram of the absolute frequencies of municipal waste collections by amount of waste collected.

The data on waste amounts collected from different sites shows a high variation. On average, each site produces about 244.77 kilograms of waste, but the typical (median) site produces 210 kilograms. This difference between the mean and median indicates that some sites with much higher waste amounts are pulling the average up. There's also a lot of spread in the data: the standard deviation is 167.16 kilograms. This means the amount of waste collected can vary greatly from one site to another. The smallest site produces just 10 kilograms, while the largest produces 1441 kilograms. Looking at the shape of the data, we see that it's right-skewed (skewness of 1.29), so most sites have lower waste amounts, but a few outliers generate much more. The high kurtosis (2.36) tells us that there are more extreme values, both very low and very high, in a typical bell-shaped (normal) distribution. Some sites are outliers and may need special attention, while most are clustered at lower waste amounts. This kind of variability suggests that a tailored approach to waste management would be much more effective than a one-size-fits-all solution. In Figure 4, we present the relationship between Waste collected and the number of containers.

The correlation between waste collected and the number of containers is significant (0.86, $p < 0.001$, 95% CI: 0.807–0.894), as expected, and proves that some optimisation has been undertaken in recent years. Using equation (2), we have identified 87 efficient collection sites, 30 with too frequent collections, and 30 with too infrequent collections. The efficiency threshold could be adjusted to the specific needs of the town of Michalovce. This approach allowed us to visually and quantitatively identify which collection sites could benefit from adjusting their collection frequency, either by increasing or decreasing the number of collections to better match the actual waste generation patterns. In Figure 5, we can see the relationship between the Number of collections of the waste and the Amount of Waste grouped by individual collection sites.

The analysis reveals a moderate positive relationship between the number of waste collections and the total amount of waste generated at each location, as indicated by a correlation coefficient of approximately 0.71 ($p < 0.001$, 95% CI: 0.623–0.784). This means that sites that are serviced more frequently tend to produce more waste. The scatterplot supports this finding: each point represents a site, and the red regression line, along with its confidence band, clearly illustrates the upward trend. This relationship suggests that changes in collection frequency could significantly impact the total volume of waste managed. In practical terms, this insight is valuable for optimising waste collection schedules and determining the appropriate container sizes for different locations, ultimately leading to more efficient and effective waste management operations. However, the chart also shows relatively significant deviations from the linear axis, which may indicate increased variability among the observations. In other words, not all locations follow the main trend - some sites generate more or less waste than would be expected based on the number of collections. These deviations suggest that additional factors influence the amount of waste produced, so it is important to consider the specific characteristics of each location when planning waste management strategies. Correlation analysis was conducted at the collection-point level (N = 147). In the following figures, we will begin analysing the variability among the collection sites. Figure 6 presents box plots of 20 collection sites with the highest variability.

The boxplot provides a clear visual summary of how waste amounts differ across various locations. For most sites, the median waste amount is below 100 units, indicating that typical waste generation is relatively modest. However, the presence of numerous extreme outliers indicates that some locations produce much more waste than is typical for their group. These outliers stand out sharply from the rest of the data, suggesting that a few sites have unique circumstances or demands. Additionally, the sizes of the boxes, which represent the interquartile range (the middle 50% of values), vary widely from one location to another. This means that some sites have very consistent waste generation, while others experience much more fluctuation. Together, these patterns highlight the diversity in waste production across locations and suggest that both typical and „not typical" cases need to be considered when planning waste management strategies in the case of the town of Michalovce. In Figure 7, we can see a relationship between Mean waste and the coefficient of variation grouped by collection sites.

This figure shows the relationship between the average waste amount at each collection site and the coefficient of variation. We can observe that, unlike observations with a normal or slightly elevated coefficient of variation, the average waste amounts for sites with high or very high coefficients of variation are concentrated on the left side of the chart, indicating lower average waste values. This observation was further examined using a boxplot of individual observations, in which we merged the categories into two groups: "normal CV" (including the normal and above-average categories) and "high CV" (including the high and very high categories from Figure 6). This boxplot is displayed in Figure 8.

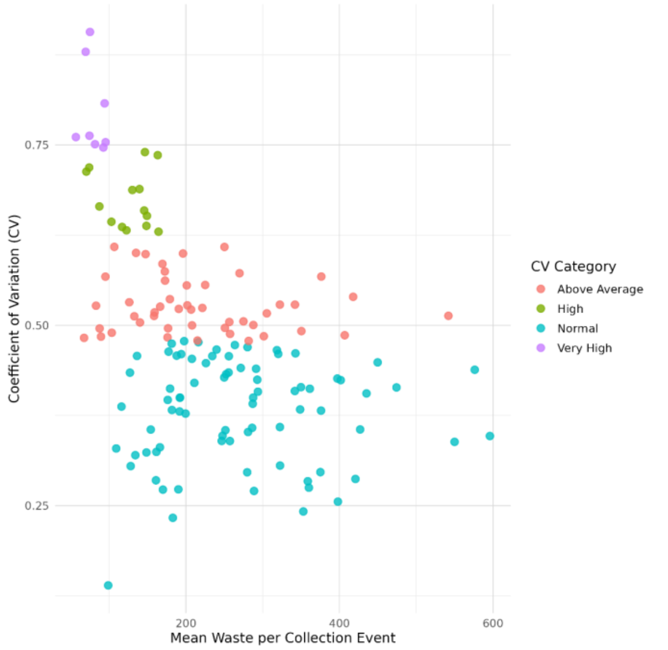

Fig. 7. Plot of Mean waste and Coefficient of variation
Source: Database of waste collection provided by the town of Michalovce

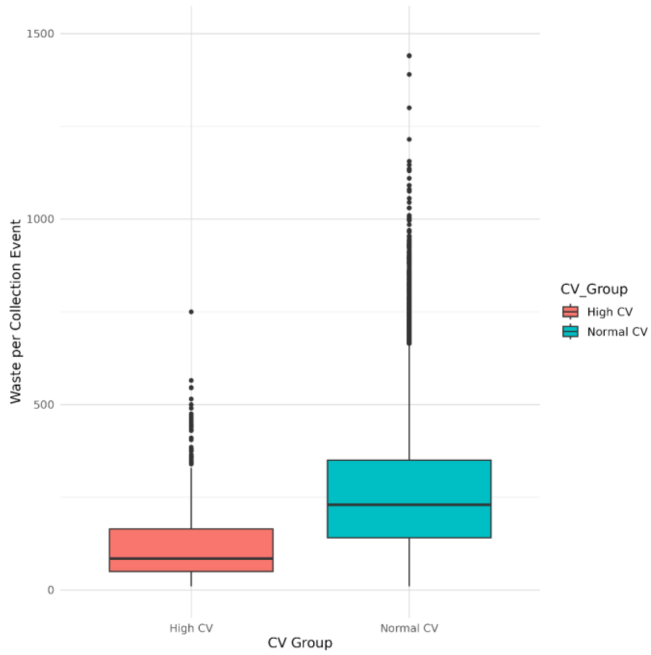

Fig.8. Box plot of the relationship among waste amount by the Coefficient of Variation
Source: Database of waste collection provided by the town of Michalovce.

Municipal waste collections that exhibited high variability (N=2093, 15,88%) during the observed period had lower average (117,01 kg) and median (85 kg) waste amounts compared to sites with normal variability (average 265.06 kg and median 230 kg). This suggests opportunities to reduce costs by optimising collection schedules for these highly variable locations. In the next analysis, we also included the standard deviation and sample size—that is, the number of collections at each municipal waste site. This is illustrated in Figure 9, which visualises the interaction between mean waste, variability, and collection frequency across all sites.

There is considerable spread, indicating that some sites with similar average waste amounts can have very different levels of variability. This pattern underscores the importance of considering both the average and the variability when planning waste management strategies, as sites with high variability may require more flexible or adaptive collection schedules compared to those with more consistent waste generation.

To complement the variability analysis, we also examined the temporal development of municipal waste collection during 2024. The results show that waste generation was not evenly distributed throughout the year, suggesting seasonal variation in both waste volume and stability.

Figure 10 presents the total monthly amount of waste collected. The highest monthly total was recorded in August 2024, when the system collected 356,296 kg of waste, and the lowest was in October 2024, with 236,420 kg. This means that the difference between the highest and lowest monthly total reached 119,876 kg, indicating substantial temporal fluctuation in the overall burden on the collection system.

Figure 11 shows the monthly average waste per collection event. The highest average value was recorded in April 2024, when the mean waste per collection event reached 262.06 kg, and the lowest was in December 2024 at 219.10 kg. The difference of nearly 43 kg per collection event suggests that individual collections were noticeably heavier in some parts of the year than in others. This means that seasonality affects not only the total amount of waste collected but also the average intensity of a single collection event.

Figure 12 illustrates the monthly coefficient of variation of waste, which captures how stable or unstable waste amounts were over time. The highest coefficient of variation was observed in January 2024 at 70.62%, while the lowest was in December 2024 at 64.97%. Although the variation in CV across months is smaller than the fluctuation in total waste volume, the results still indicate that the predictability of waste generation also changes during the year. In other words, some months are associated not only with more or less waste but also with more or less stable waste collection patterns.

Overall, the seasonality analysis confirms that temporal factors should be taken into account when evaluating waste collection performance. The observed monthly differences in total waste, mean waste per collection event, and waste variability suggest that a fixed collection schedule may not be equally suitable throughout the entire year. Therefore, incorporating seasonal patterns into waste collection planning may improve the alignment between collection frequency, container capacity, and actual waste generation.

To strengthen the empirical analysis beyond descriptive statistics, two simplified regression models were estimated, with waste mass per collection event as the dependent variable.

Diagnostic checks were performed on both simplified regression models to assess whether the main assumptions of linear regression were reasonably satisfied. In both models, the correlation between fitted values and residuals was effectively zero, suggesting that no major systematic linear bias remained. The Shapiro–Wilk test was statistically significant in both cases, indicating that the residuals were not perfectly normally distributed. However, given the large sample size ($N = 15{,}273$), this result is not unexpected and should be interpreted cautiously.

A more practically relevant issue was the presence of heteroskedasticity. The correlation between fitted values and absolute residuals was 0.329 in Model A and 0.376 in Model B, indicating that the residual variance increased with higher fitted values. This pattern is consistent with the skewed and heavy-tailed nature of waste mass data. Influential observations were also present, with 1,001 events exceeding the Cook's distance threshold in Model A and 901 events in Model B. These findings do not invalidate the use of the regression models, but they do imply that the models should be interpreted as supportive analytical tools rather than as fully assumption-free causal models.

In Model A, waste mass was explained by the number of containers and weekday effects. The results showed that the number of containers had a strong positive and statistically significant effect ($\beta = 48.514$, $p < 0.001$), indicating that each additional container was associated with an increase of approximately 48.5 units of waste per collection event. The model explained 37.4% of the observed variation in waste mass ($R^2 = 0.374$). Weekday effects were also statistically significant. Compared with the reference weekday, waste mass was significantly higher on Monday ($\beta = 44.188$, $p < 0.001$), Tuesday ($\beta = 39.415$, $p < 0.001$), and Thursday ($\beta = 66.728$, $p < 0.001$), while it was significantly lower on Wednesday ($\beta = -37.197$, $p < 0.001$) and especially Saturday ($\beta = -79.900$, $p < 0.001$).

In Model B, waste mass was explained by the number of containers and the monthly effects. The coefficient for the number of containers remained strong, positive, and statistically significant ($\beta = 47.553$, $p < 0.001$). This model explained 32.5% of the variation in waste mass ($R^2 = 0.325$). Monthly effects indicated temporal variation. Compared with the reference month, waste mass per collection event was significantly higher in March ($\beta = 10.972$, $p < 0.05$) and April ($\beta = 15.080$, $p < 0.01$), while significantly lower values were observed in May ($\beta = -11.983$, $p < 0.05$), June ($\beta = -18.895$, $p < 0.001$), October ($\beta = -12.162$, $p < 0.05$), and December ($\beta = -29.308$, $p < 0.001$).

The regression analysis confirms that waste mass per collection event is strongly associated with the number of containers and also varies systematically over time. The weekday model captures short-cycle operational regularities, whereas the month model confirms broader seasonal effects. These results strengthen the interpretation of the descriptive findings and provide a more rigorous empirical basis for the claim that municipal waste collection is shaped by both infrastructure-related and temporal factors.

5. Summary and conclusion

The issue of waste-management governance is widely discussed across national economies, municipalities, and the scientific community. Mathematical optimisation models have also found their place in waste management, a trend accelerated by the rapid development of ICT. Recently, a new paradigm has emerged: artificial intelligence (AI).

The project that produced this case study aims to develop predictive models to manage selected waste-management processes under specific conditions. Because these models are complex, it is crucial to identify shortcomings in the quality of individual waste-management processes. In this context, the primary factors indicating process quality must be determined based on real-world data.

This case study has demonstrated relationships between the variability of municipal waste collections and both the average and the absolute quantities of waste. A limitation of the paper, however, is that it abstracts from the determinants of this variability. These determinants will be examined in subsequent studies, setting the direction for future research. On the one hand, our research showed a correlation between the amount of waste and the number of collections, though there is considerable variability around the linear regression line. On the other hand, reducing municipal waste remains the most effective way to optimise collection frequency. The practical question, therefore, is how to influence those who overproduce waste. One possibility, already used in the town of Michalovce, is pay-as-you-throw pricing. Exploring the effects of such incentives could be a focus of future research. Our analysis confirms that additional determinants influence waste quantities.

Waste-management optimisation must consider each city and municipality as a whole, but priority should be given to rapidly resolving collection points with high variability in pick-ups. Our analysis shows that these sites have relatively low average waste volumes yet occasionally experience extremely high peaks, so the mean remains low despite severe outliers. High variability complicates optimal collection scheduling. We believe that incorporating AI, together with suitable technical and managerial measures, can markedly reduce variability. As indicated in the literature review, this can lower costs for waste operators and residents, reduce greenhouse gas emissions, and improve both environmental quality and residents' quality of life beyond local boundaries.

Overall, the seasonality analysis confirms that temporal factors should be taken into account when evaluating waste collection performance. The observed monthly differences in total waste, mean waste per collection event, and waste variability suggest that a fixed collection schedule may not be equally suitable throughout the entire year. Therefore, incorporating seasonal patterns into waste collection planning may improve the alignment between collection frequency, container capacity, and actual waste generation.

Although the analysis is primarily based on descriptive and comparative statistical methods, the results provide important operational insights into the mechanisms behind inefficient waste collection. In particular, sites with lower mean waste per collection event often exhibited higher relative variability, suggesting unstable filling patterns and a higher risk of mismatched collection frequency. Conversely, sites with more stable and consistently higher waste volumes were more likely to follow predictable collection patterns, making them easier to optimise operationally.

A large database of collection records for a specific town has revealed great potential for a wide range of analyses. Data-driven processes of this sort should be supported throughout Slovakia and in other EU regions. Targeted assistance from EU structures could have a highly positive effect not only on specific waste-management practices but also on scientific research, connecting economic practice with academia. The study identifies collection sites with comparatively unstable waste-generation patterns and operational mismatches between waste volume, container capacity, and collection frequency. These findings can support municipal decision-making and the design of improved collection schedules. However, because the analysis does not include an intervention or post-implementation evaluation, the results should be interpreted as evidence of optimisation potential rather than as direct proof of realised optimisation effects.

Acknowledgements

Funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09105-03-V02-00056 AI4WasteManagement: Research on prediction and optimization of waste management processes.

During the preparation of this work, the authors used Grammarly and DeepL to improve the study's language and clarity. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Reference

Abbasi, S. A., 2020. Efficient Control Charts for Monitoring Process CV Using Auxiliary Information. IEEE Access, 8, 46176–46192, DOI: 10.1109/ACCESS.2020.2977833

Abdallah, M., Adghim, M., Maraqa, M., Aldahab, E., 2019. Simulation and optimization of dynamic waste collection routes. Waste Management & Research: The Journal for a Sustainable Circular Economy, 37(8), 793–802, DOI: 10.1177/0734242X19833152

Apaydin, O., Gonullu, M. T., 2007. Route optimization for solid waste collection: Trabzon (Turkey) case study. Global NEST Journal, 9(1), 6–11, DOI: 10.30955/gnj.000388

Apaydin, O., Gonullu, M. T., 2008. Emission control with route optimization in solid waste collection process: A case study. Sadhana, 33(2), 71–82, DOI: 10.1007/s12046-008-0007-4

Bányai, T., Tamás, P., Illés, B., Stankevičiūtė, Ž., Bányai, Á., 2019. Optimization of Municipal Waste Collection Routing: Impact of Industry 4.0 Technologies on Environmental Awareness and Sustainability. International Journal of Environmental Research and Public Health, 16(4), 634, DOI: 10.3390/ijerph16040634

Barth, L., Schweiger, L., Benedech, R., Ehrat, M., 2023. From data to value in smart waste management: Optimizing solid waste collection with a digital twin-based decision support system. Decision Analytics Journal, 9, 100347, DOI: 10.1016/j.dajour.2023.100347

Beliën, J., De Boeck, L., Van Ackere, J., 2014. Municipal Solid Waste Collection and Management Problems: A Literature Review. Transportation Science, 48(1), 78–102, DOI: 10.1287/trsc.1120.0448

Das, S., Bhattacharyya, B. Kr., 2015. Optimization of municipal solid waste collection and transportation routes. Waste Management, 43, 9–18, DOI: 10.1016/j.wasman.2015.06.033

Erçin, M., Altıntaş, U., Köse, M., Kös, R., Atasoy, A., 2021. Route Optimization for Waste Collection Process Through IoT Supported Waste Management System. ResearchGate preprint, DOI: 10.13140/RG.2.2.23182.38720

Filipiak, K. A., Abdel-Malek, L., Hsieh, H.-N., Meegoda, J. N., 2009. Optimization of Municipal Solid Waste Collection System: Case Study. Practice Periodical of Hazardous, Toxic, and Radioactive Waste Management, 13(3), 210–216, DOI: 10.1061/(ASCE)1090-025X(2009)13:3(210)

Ghiani, G., Manni, A., Manni, E., Moretto, V., 2021. Optimizing a waste collection system with solid waste transfer stations. Computers & Industrial Engineering, 161, 107618, DOI: 10.1016/j.cie.2021.107618

Han, H., Cueto, E. P., 2015. Waste Collection Vehicle Routing Problem: Literature Review. PROMET - Traffic&Transportation, 27(4), 345–358, DOI: 10.7307/ptt.v27i4.1616

Hemidat, S., Oelgemöller, D., Nassour, A., Nelles, M., 2017. Evaluation of Key Indicators of Waste Collection Using GIS Techniques as a Planning and Control Tool for Route Optimization. Waste and Biomass Valorization, 8(5), 1533–1554, DOI: 10.1007/s12649-017-9938-5

Ingaldi, M., Ulewicz, R., 2026. A Digital Sustainability Maturity Framework for Assessing Industry 4.0 and ESG Integration. Production Engineering Archives, 32(2), 262-277, DOI: 10.30657/pea.2026.32.21

Shah, P. J., Anagnostopoulos, T., Zaslavsky, A., Behdad, S., 2018. A stochastic optimization framework for planning of waste collection and value recovery operations in smart and sustainable cities. Waste Management, 78, 104–114, DOI: 10.1016/j.wasman.2018.05.019

Khan, M. A., Khan, R., Al-Zghoul, T. M., Khan, A., Hussain, A., Baarimah, A. O., Arshad, M. A., 2024. Optimizing municipal solid waste management in urban Peshawar: A linear mathematical modeling and GIS approach for efficiency and sustainability. Case Studies in Chemical and Environmental Engineering, 9, 100704, DOI: 10.1016/j.cscee.2024.100704

Kinobe, J. R., Bosona, T., Gebresenbet, G., Niwagaba, C. B., Vinnerås, B., 2015. Optimization of waste collection and disposal in Kampala city. Habitat International, 49, 126–137, DOI: 10.1016/j.habitatint.2015.05.025

Krzywda, D., Krzywda, J., 2026. Integrating circular economy and reverse logistics into the quality management systems: Evidence from Polish SMEs. System Safety: Human - Technical Facility - Environment, 8(1), 189-201, DOI: 10.2478/czoto-2026-0018

Lacko, R., Kuchta, M., Hajduová, Z., Dula, R., 2026. Household Solid Waste Eco-Efficiency in Slovak NUTS-3 Regions: A Two-Stage DEA Analysis. Production Engineering Archives, 32(2), 288-301, DOI: 10.30657/pea.2026.32.23

Lu, J.-W., Chang, N.-B., Liao, L., Liao, M.-Y., 2017. Smart and Green Urban Solid Waste Collection Systems: Advances, Challenges, and Perspectives. IEEE Systems Journal, 11(4), 2804–2817, DOI: 10.1109/JSYST.2015.2469544

Mendes, P., Santos, A. C., Nunes, L. M., Teixeira, M. R., 2013. Evaluating municipal solid waste management performance in regions with strong seasonal variability. Ecological Indicators, 30, 170–177, DOI: 10.1016/j.ecolind.2013.02.017

Nematollahi, H., Gitipour, S., Mehrdadi, N., 2024. Comparative life cycle assessment and route optimization modeling of smart versus conventional municipal waste collection: Environmental impact analysis in an urban context. Results in Engineering, 24, 103408, DOI: 10.1016/j.rineng.2024.103408

Neto, A. B. P. S., Simões, C. L., Simoes, R., 2024. Optimization of municipal solid waste collection system: Systematic review with bibliometric literature analysis. Journal of Material Cycles and Waste Management, 26(4), 1906–1917, DOI: 10.1007/s10163-024-01966-y

Przydatek, G., Ciągło, K., 2020. Factors of variability in the accumulation of waste in a mountain region of southern Poland. Environmental Monitoring and Assessment, 192(2), 153, DOI: 10.1007/s10661-020-8103-y

Rattanawai, N., Arunyanart, S., Pathumnakul, S., 2024. Optimizing municipal solid waste collection vehicle routing with a priority on infectious waste in a mountainous city landscape context. Transportation Research Interdisciplinary Perspectives, 24, 101066, DOI: 10.1016/j.trip.2024.101066

Richter, A., Ng, K. T. W., Vu, H. L., Kabir, G., 2021. Waste disposal characteristics and data variability in a mid-sized Canadian city during COVID-19. Waste Management, 122, 49–54, DOI: 10.1016/j.wasman.2021.01.004

Sousa, V., Drumond, A., Meireles, I., 2024. Fuel consumption rate and emissions variability in waste collection services routes: Case study of Cascais Ambiente. Environmental Science and Pollution Research, 31(12), 17732–17747, DOI: 10.1007/s11356-023-29045-z

Wu, H., Tao, F., Yang, B., 2020. Optimization of Vehicle Routing for Waste Collection and Transportation. International Journal of Environmental Research and Public Health, 17(14), 4963. DOI: 10.3390/ijerph17144963

Xue, W., Cao, K., 2016. Optimal routing for waste collection: A case study in Singapore. International Journal of Geographical Information Science, 30(3), 554–572, DOI: 10.1080/13658816.2015.1103374

Xue, W., Cao, K., Li, W., 2015. Municipal solid waste collection optimization in Singapore. Applied Geography, 62, 182–190, DOI: 10.1016/j.apgeog.2015.04.002

Appendix

Appendix A

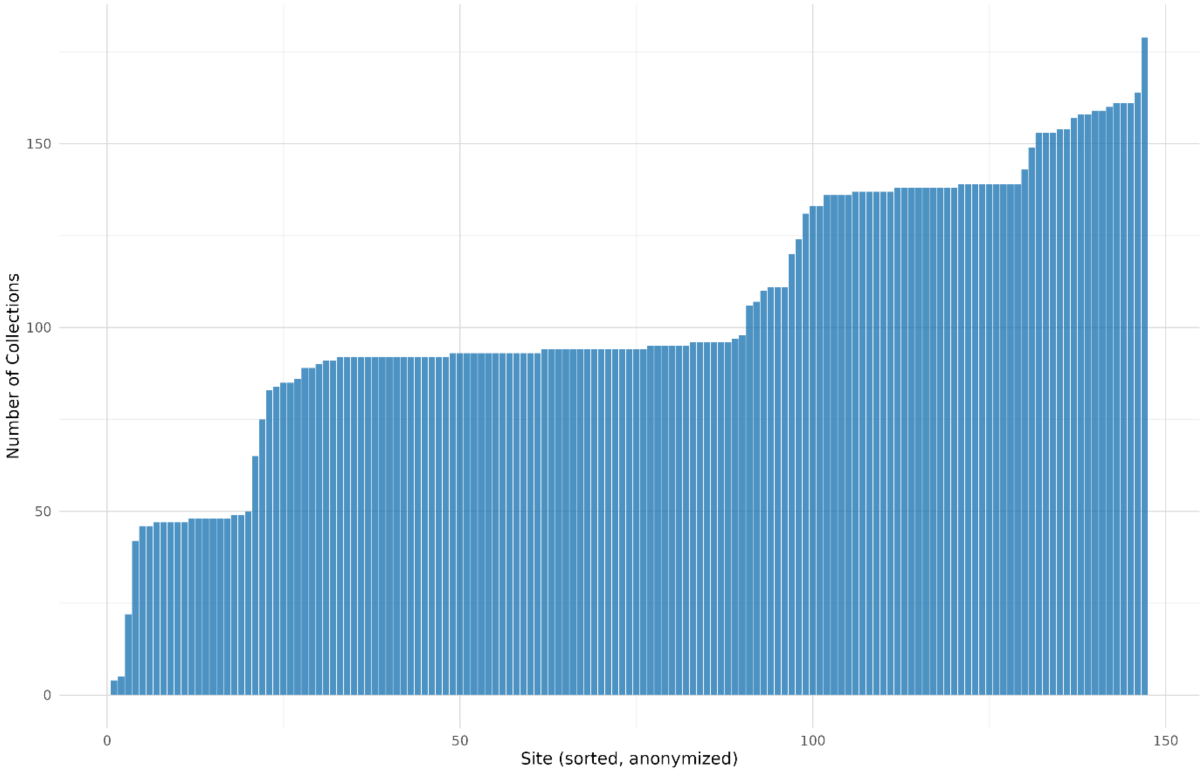

Fig. 2. Number of collections per location (sorted ascending)
Source: Database of waste collection provided by the town of Michalovce

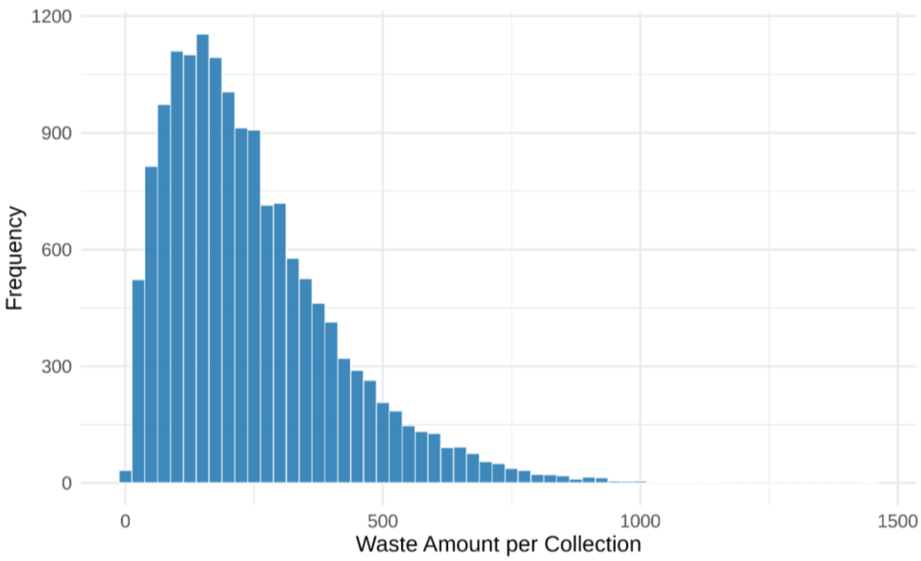

Fig. 3. Distribution of waste amounts
Source: Database of waste collection provided by the town of Michalovce

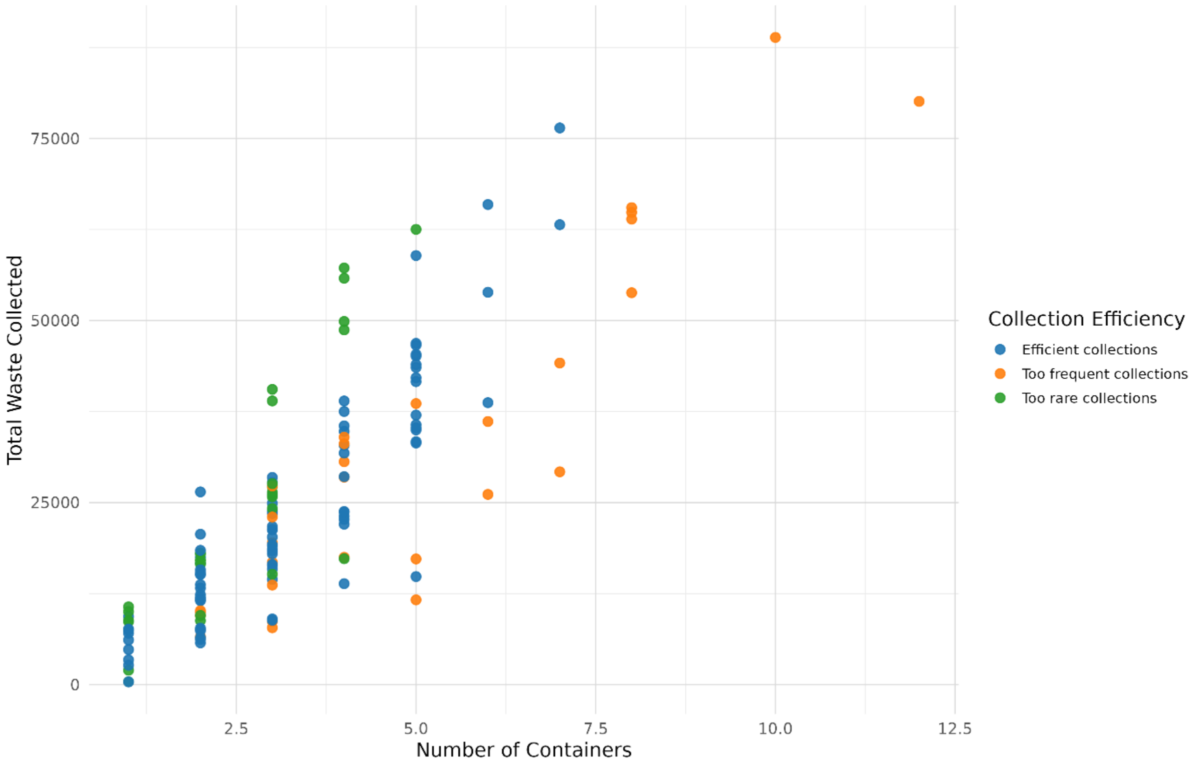

Fig. 4. Relationship between collected waste and number of containers
Source: Database of waste collection provided by the town of Michalovce

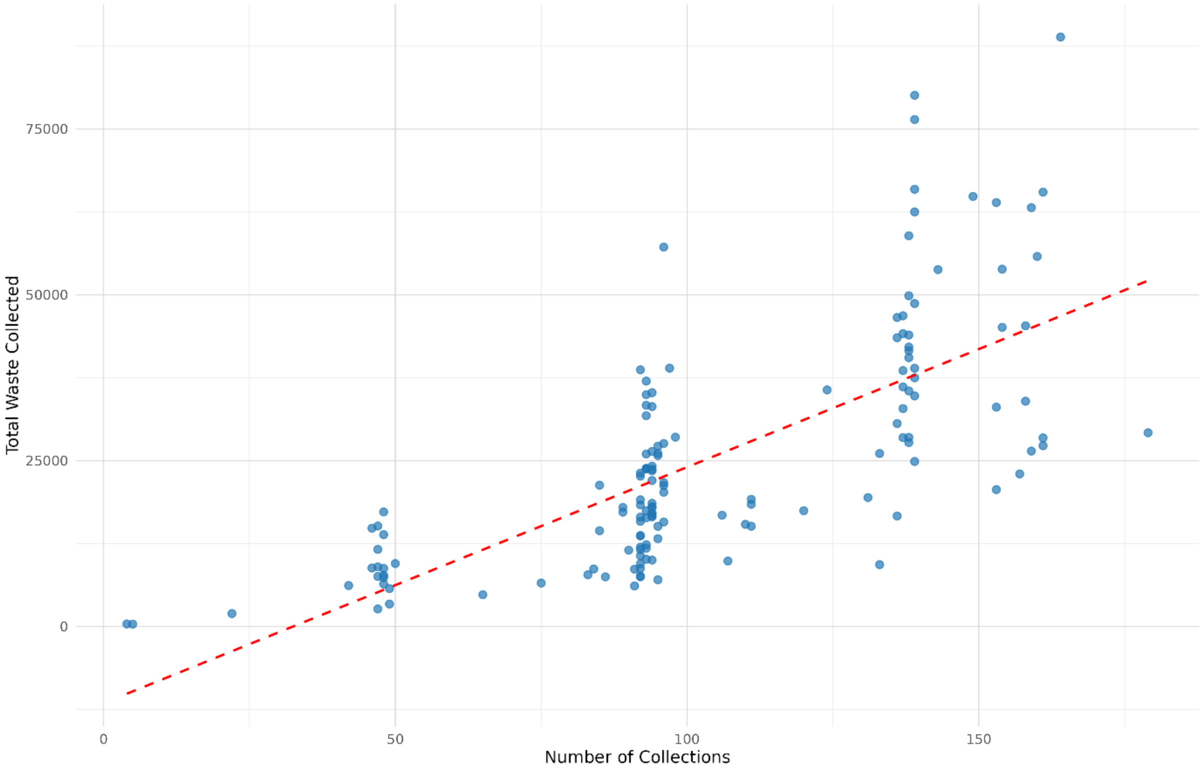

Fig. 5. Correlation between Number of Waste Collections and Total MSW (in kg)
Source: Database of waste collection provided by the town of Michalovce

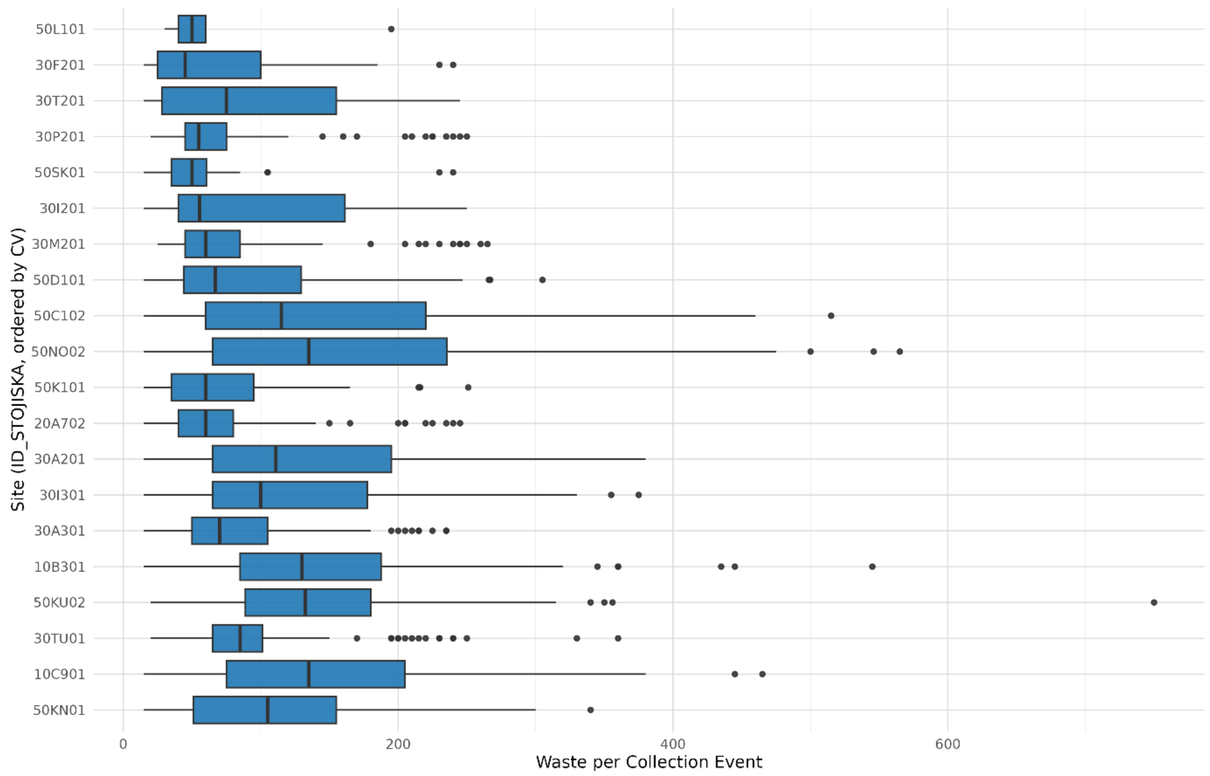

Fig. 6. Box plot for locations with the highest CV
Source: Database of waste collection provided by the town of Michalovce

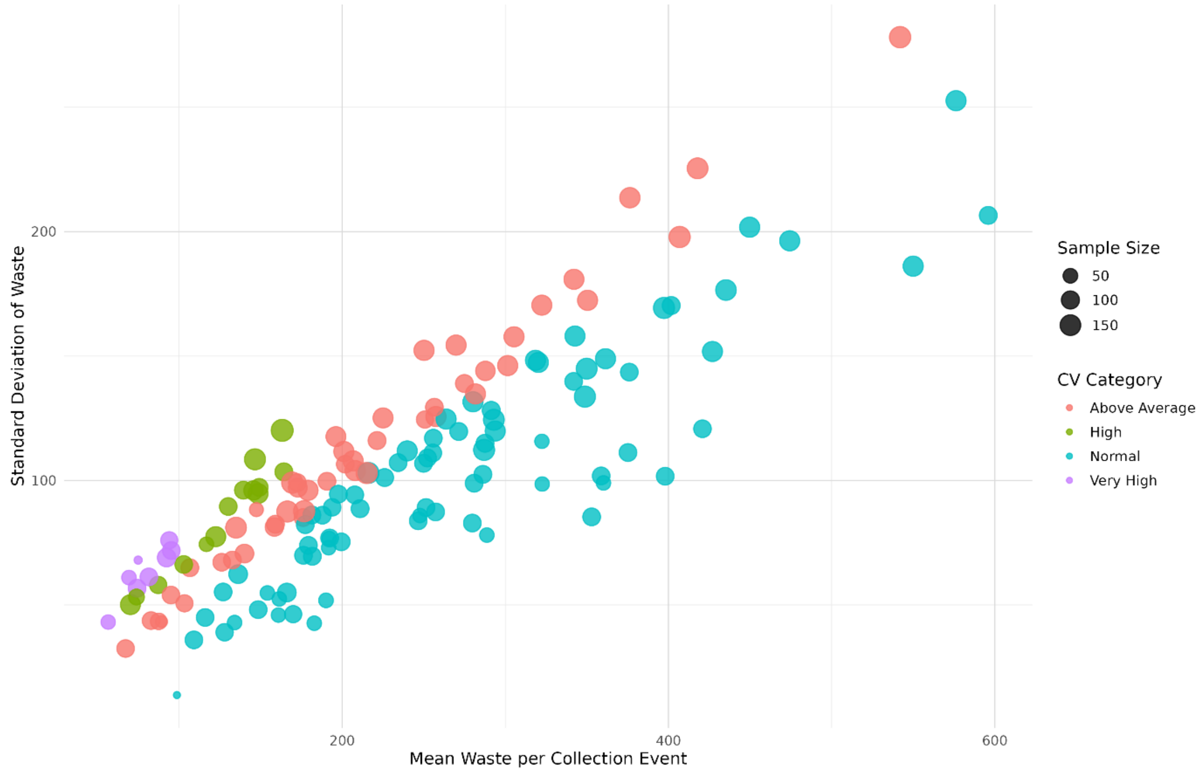

Fig. 9. Scatter plot of relations between mean, standard deviation, and sample size related to the Coefficient of variation
Source: Database of waste collection provided by the town of Michalovce

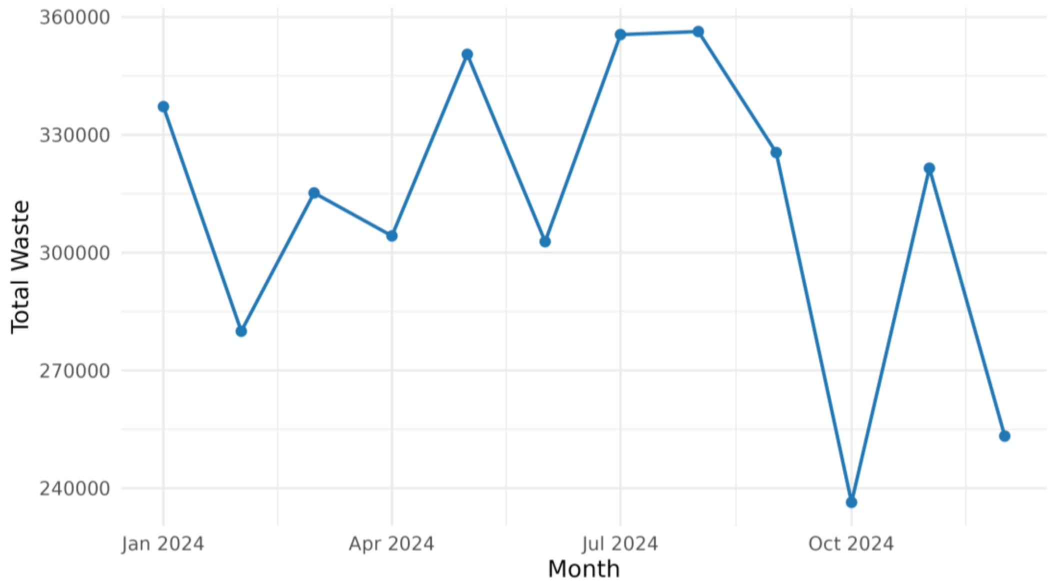

Fig. 10. Monthly total waste collected
Source: Database of waste collection provided by the town of Michalovce

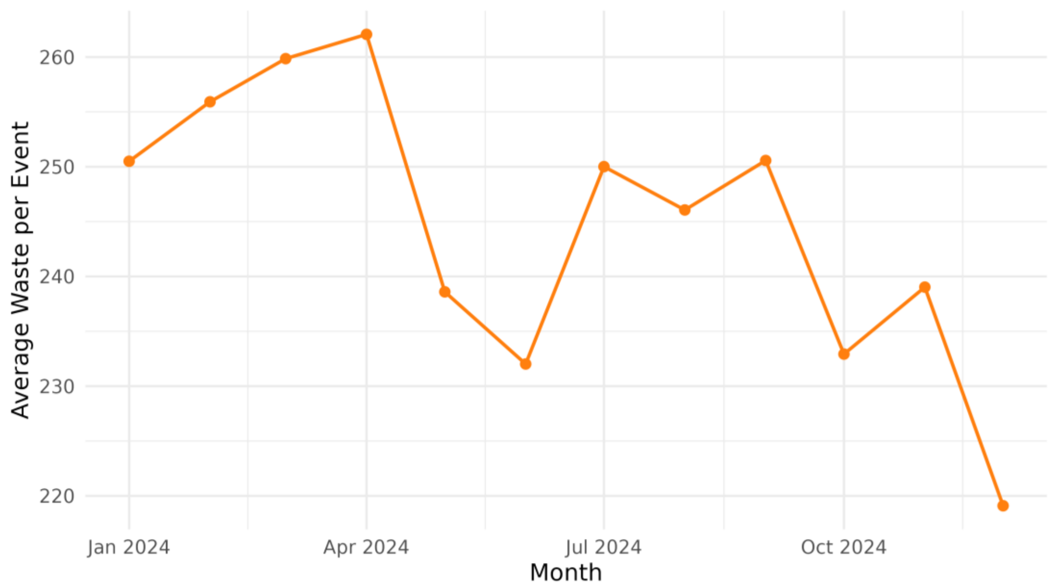

Fig. 11. Monthly average waste per collection event
Source: Database of waste collection provided by the town of Michalovce

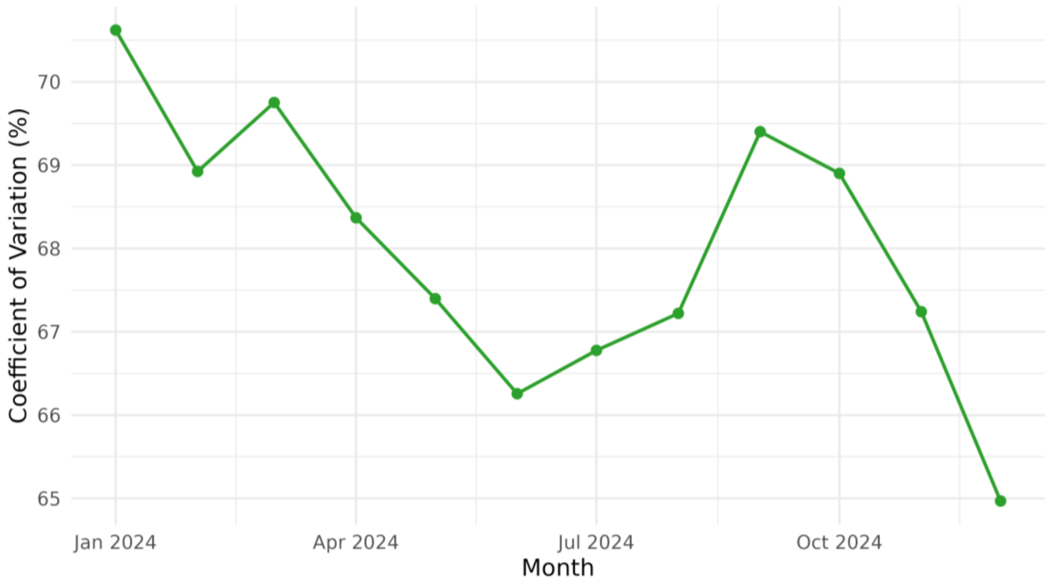

Fig. 12. Monthly coefficient of Variation of waste amount
Source: Database of waste collection provided by the town of Michalovce

Table 2: Regression analysis results

Model	Var	estimate	std_error	t_value	p_value	significance
Model A	(Intercept)	44.006	2.743	16.042	<.001	***
Model A	n_containers	48.514	0.571	84.933	<.001	***
Model A	Monday	44.188	2.976	14.846	<.001	***
Model A	Saturday	-79.9	4.447	-17.969	<.001	***
Model A	Thursday	66.728	3.691	18.079	<.001	***
Model A	Tuesday	39.415	3.076	12.814	<.001	***
Model A	Wednesday	-37.197	4.718	-7.885	<.001	***
Model B	(Intercept)	68.978	4.311	16.001	<.001	***
Model B	n_containers	47.553	0.559	85.035	<.001	***
Model B	Feb	7.161	5.593	1.28	0.2004	
Model B	Mar	10.972	5.44	2.017	0.0437	*
Model B	Apr	15.08	5.504	2.74	<.001	**
Model B	May	-11.982	5.184	-2.311	0.0208	*
Model B	Jun	-18.895	5.338	-3.54	<.001	***
Model B	Jul	2.926	5.225	0.56	0.5755	
Model B	Aug	-1.438	5.202	-0.276	0.7823	
Model B	Sep	-0.346	5.344	-0.065	0.9484	
Model B	Oct	-12.161	5.712	-2.129	0.0333	*
Model B	Nov	-5.912	5.298	-1.116	0.2644	
Model B	Dec	-29.308	5.51	-5.319	<.001	***

Source: Database of waste collection provided by the town of Michalovce