Journal of Sustainable Development of Transport and Logistics
journal home page: https://jsdtl.sciview.net

Tson, O., Vovk, Y., Rozhko, N., Lyashuk, O., & Plekan, U. (2026). Demand concentration and temporal imbalance in last-mile urban freight distribution. *Journal of Sustainable Development of Transport and Logistics, 11*(1), 153-171. doi:10.14254/jsdtl.2026.11-1.8.

Scientific Platform

ISSN 2520-2979

Demand concentration and temporal imbalance in last-mile urban freight distribution

Oleg Tson, Yuriy Vovk, Nataliya Rozhko, Oleg Lyashuk, Uliana Plekan

Department of Automotive Transport and Logistics, Ternopil Ivan Puluj National Technical University, Ruska str, 56, 46001 Ternopil, Ukraine
tson_oleg_@ukr.net

Article history:

Received: April 11, 2026
1st Revision: May 02, 2026
Accepted: June 20, 2026

DOI:
10.14254/jsdtl.2026.11-1.8

Abstract: *Purpose*. This paper identifies temporal, spatial, and organizational patterns in urban freight demand and develops a statistically grounded pre-optimization framework for sustainable last-mile logistics planning in a medium-sized city. *Methodology*. The empirical basis comprises 4,967 anonymized delivery records collected over 125 active service days in Ternopil. Demand intensity was normalized by active day and examined through Pareto analysis, Gini and Herfindahl-Hirschman concentration indices, and a 5,000-replication location bootstrap. Weekday and organizational differences were tested using Kruskal-Wallis tests with Holm-adjusted post-hoc comparisons, while a weekday-adjusted HC3 regression assessed the within-period trend and Pearson and Spearman correlations examined the consistency between load mass and volume. *Results*. Average daily order intensity increased by 25.9% and average daily load mass by 36.4% between January and May. Weekday imbalance was substantial, with Saturday significantly lower than every Monday-Friday group after correction. Destination demand was highly concentrated: the Gini coefficient reached 0.729 for order frequency and 0.832 for load mass, with a dual core of just 30 locations (7.1%) generating 65.0% of total mass. Load mass and recorded volume were strongly and positively associated (Spearman's rho = 0.817), while organizational divisions differed significantly in shipment-size profiles. *Theoretical contribution*. The framework integrates temporal imbalance, dual-capacity demand, and two-dimensional location concentration within a single decision architecture for urban freight distribution. *Practical implications*. The findings support weekday-specific capacity thresholds, fixed service templates for the demand core, sector-based consolidation for the middle segment, and dynamic routing for the dispersed long tail, with cost and emission effects formulated for subsequent operational validation rather than assumed.

Corresponding author: *Oleg Tson*
E-mail: tson_oleg_@ukr.net

Keywords: weekday capacity planning, location-based demand segmentation, Gini coefficient concentration analysis, organizational shipment heterogeneity, dynamic route allocation, sustainable urban logistics

Sustainable Development Goals (SDGs): SDG 9: Industry, Innovation and Infrastructure; **SDG 11:** Sustainable Cities and Communities; **SDG 12:** Responsible Consumption and Production; **SDG 13:** Climate Action

1. Introduction

Mid-mile logistics, linking regional hubs, distribution centers, and fulfillment points, has emerged as a key component of urban freight distribution. The sustainability challenge is not limited to selecting cleaner vehicles. It also concerns when deliveries occur, how demand is distributed among locations, and whether daily workloads are sufficiently predictable to support consolidation. Reviews of last-mile logistics consistently identify consolidation, information sharing, and data-driven planning as core mechanisms for improving operational and environmental performance (Bosona, 2020; Golinska-Dawson & Sethanan, 2023; Lyons & McDonald, 2023).

For medium-sized cities, the empirical evidence needed to implement these mechanisms is frequently incomplete. Public freight statistics are usually aggregated, while company-level operational records remain commercially sensitive and are rarely transformed into decision-oriented indicators. Recent research emphasizes that freight planning needs disaggregated demand data, collaboration between public and private actors, and methods that can identify priority delivery areas without exposing customer information (Bjørgen et al., 2025; Diana et al., 2020; Maxner et al., 2025). Establishment-level data are especially valuable because they reveal the frequency, magnitude, and temporal recurrence of deliveries that aggregate traffic counts cannot provide.

The operational problem addressed in this paper is the coexistence of three forms of heterogeneity. First, daily demand may vary systematically by weekday and trend over time. Second, locations may differ substantially in both visit frequency and delivered mass. Third, organizational divisions may generate different order-size profiles even when they share the same distribution infrastructure. Treating these flows as homogeneous can produce unnecessary route changes, weak consolidation, and mismatches between vehicle capacity and actual workload. In contrast, a combined temporal-spatial segmentation can support stable service frequencies for recurring demand and flexible routing for the long tail of occasional destinations.

The study uses anonymized records from a regional distribution operator serving Ternopil and nearby settlements. The observation period covers January 2 to May 30, 2026. The paper does not attempt to solve a full vehicle routing problem because the source system does not contain geographic coordinates, vehicle capacities, route distances, service times, or fuel consumption. Instead, it develops the empirical layer that must precede optimization: a reproducible diagnosis of demand intensity, concentration, and imbalance.

The purpose of the study is to identify temporal, spatial, and organizational patterns in urban freight demand and to translate those patterns into a decision framework for capacity allocation and route design. Four research questions are addressed: (RQ1) How strongly does daily demand differ across weekdays and over the observation period? (RQ2) How concentrated are order frequency and load mass across destinations? (RQ3) Do organizational divisions generate distinct shipment-size profiles? (RQ4) What capacity and segmentation rules follow from the observed demand structure? Five hypotheses are tested: H1, daily order, mass, and volume distributions differ by weekday; H2, the weekday-adjusted time coefficient is positive; H3, destination demand is materially concentrated; H4, order mass and volume are positively associated; and H5, order-mass distributions differ across organizational divisions.

2. Literature review

2.1. Urban freight demand as a planning input

Urban freight demand is more difficult to observe than passenger movement because heterogeneous establishments, shipment sizes, commercial schedules, and contractual arrangements generate freight trips. Demand models therefore combine establishment attributes, commodity quantities, deliveries, and vehicle movements at different levels of aggregation. Nuzzolo and Comi (2014) proposed a mixed quantity-delivery-vehicle framework, while Balla et al. (2023) showed that model choice must reflect the distributional properties of establishment-level freight generation. Middela and Ramadurai (2024) further demonstrated that measurement period and spatial dependence affect model accuracy. These studies support two methodological decisions in the present paper: active-day normalization and the use of non-parametric tests for highly skewed order data.

Demand-side observation can also identify the most important loading and unloading areas. Diana et al. (2020) combined fleet traces and spatial clustering to prioritize urban freight survey locations. Liu et al. (2018) showed that freight supply and demand can be simultaneously clustered yet spatially mismatched, which calls for differentiated management rather than a uniform citywide policy. The present study does not include coordinates, but frequency and mass concentration provide a non-spatial precursor to geographic clustering: they identify which anonymized locations merit fixed service design and which should remain within dynamic routes.

2.2. Consolidation, dynamic routing, and sustainable logistics

The literature distinguishes between structural interventions, such as consolidation centers and micro-depots, and operational interventions, such as route optimization, delivery windows, and real-time dispatching. Fontaine et al. (2023) connected city logistics design with consolidation and regulation, whereas Calabrò et al. (2023) used spatial agent-based modeling to compare fragmented door-to-door delivery with consolidation-based alternatives. Golinska-Dawson and Sethanan (2023) concluded that micro-depots, parcel lockers, mobile depots, and data-enabled planning can reduce energy demand by improving vehicle load factors and limiting redundant trips. These benefits are conditional: consolidation is useful only when sufficient demand density exists.

Dynamic vehicle routing research addresses uncertainty and the arrival of new information during operations. Mardešić et al. (2024) emphasized that dynamic routing requires policies that respond to changing demand rather than fixed route lists. This insight is directly relevant to the long tail of low-frequency locations. A hybrid operating model may therefore be preferable: recurrent high-volume locations can be scheduled through stable route templates, while dispersed low-volume requests can be assigned dynamically after daily order cut-off.

Sustainable last-mile strategies also involve curb management, stakeholder coordination, and planning integration. Castrellon and Sanchez-Diaz (2024) found that curbside interventions can support sustainable cities but may create trade-offs across stakeholders. Plazier et al. (2024) highlighted the interdependence of cooperation, regulation, and innovation, and Maxner et al. (2025) documented that last-mile freight is still insufficiently represented in city sustainability plans. This reinforces the value of transparent, firm-level indicators that can later be connected to vehicle kilometers, emissions, and urban policy.

2.3. Research gap and conceptual contribution

Existing studies offer sophisticated routing models and broad sustainability taxonomies, but a practical gap remains between raw enterprise records and optimization-ready demand profiles. Many operators possess order histories yet lack a procedure for combining temporal imbalance, location concentration, organizational heterogeneity, and data quality in one diagnostic workflow. This paper contributes a pre-optimization framework in which demand is first normalized by active day, then tested for temporal differences, measured for concentration, and finally translated into operating segments. The contribution is deliberately intermediate: it does not claim optimized routes, but provides a statistically grounded basis for designing and testing them.

3. Research methods

3.1. Study design and data source

The study follows an observational case-study design based on operational delivery records. The original spreadsheet contained 4,968 records. One exact duplicate was identified and removed, resulting in an analytical sample of 4,967 records. The records represented 125 active service days, 420 destination labels, 20 route-label variants, and three organizational divisions. The observation period extended from January 2 to May 30, 2026. Customer names were not present in the analytical output, and rank-based location codes replaced destination labels for concentration analysis.

Each row represented one dispatch record and included route label, organizational division, arrival date, destination, load mass, recorded volume, and weekday. The source field records load mass in kilograms, as confirmed by the data owner. Volume is reported in cubic meters. Because volume was missing in 899 records (18.1%), load mass was selected as the primary workload measure and volume-based results were treated as secondary available-case evidence.

Table 1 summarizes the analytical framework applied throughout the study, linking each dimension of demand heterogeneity to its corresponding statistical method and interpretive purpose.

Table 1: Analytical framework and principal indicators

Dimension	Indicator/method	Interpretation
Data quality	Duplicate count; missingness by variable; route-label variants	Defines the reliability and limits of subsequent inference.
Temporal intensity	Records, load mass, and volume per active service day	Separates demand growth from differences in the number of operating days.
Temporal imbalance	Weekday averages; Kruskal-Wallis tests; epsilon-squared effect size	Tests whether workload distributions differ across service days.
Spatial concentration	Pareto shares; Gini coefficient; Herfindahl-Hirschman Index (HHI)	Measures whether a small subset of locations dominates service frequency or mass.
Organizational heterogeneity	Division shares, medians, and Kruskal-Wallis test	Identifies differences in shipment-size profiles among internal demand sources.
Mass-volume consistency	Pearson correlation and Spearman rank correlation on complete cases	Assesses linear and monotonic agreement between two workload dimensions.
Time trend	OLS with weekday indicators and HC3 robust standard errors	Estimates within-period change without interpreting it as causal.

Source: Authors' methodological design.

3.2. Data preparation and descriptive indicators

Dates were converted from spreadsheet serial values to calendar dates. Leading and trailing spaces were removed from text fields, while exact duplicate detection used all analytical fields jointly. No missing volume value was replaced by zero because a missing observation does not imply a zero-volume shipment. The principal descriptive indicators were the arithmetic mean, median, standard deviation, coefficient of variation, monthly totals, and averages per active service day.

For month m, active-day-normalized demand was calculated as:

$$I(m) = Q(m)/D(m) \tag{1}$$

where I_m is monthly demand intensity, Q_m is the monthly total of records, mass, or volume, and D_m is the number of active service days in month m. This normalization avoids overstating demand in months containing more operating days.

3.3. Concentration measures

The Gini coefficient was used to assess inequality across destination-level demand values:

$$G = \Sigma_i \Sigma_j \, |x(i) - x(j)| \, / \, [2n^2 \bar{x}] \tag{2}$$

where x_i and x_j denote location-level order counts or load mass, n is the number of locations, and $\bar{x}$ is the corresponding mean. A value near zero indicates an even distribution, while a value near one indicates strong concentration. The Herfindahl-Hirschman Index was calculated as:

$$HHI = \Sigma_i \, s(i)^2 \tag{3}$$

where s_i is the share of location i in total order frequency or load mass. Pareto curves were additionally used because they are readily interpretable for operational segmentation.

3.4. Statistical testing and trend model

The order-level mass and volume distributions were strongly right-skewed; therefore, weekday and division comparisons used the Kruskal-Wallis test. Effect size was summarized by epsilon squared, $\varepsilon^2 = (H - k + 1)/(N - k)$, where H is the Kruskal-Wallis statistic, k is the number of groups, and N is the total number of observations. Significance was assessed at $\alpha = 0.05$. Correlations between mass and volume were calculated on complete cases using Pearson's r and Spearman's rho.

A weekday-adjusted regression assessed the direction of the within-period trend:

$$Y(d) = \beta_0 + \beta_1 t(d) + \Sigma[j = 2 \ldots 6] \, \beta(j) D(j,d) + \varepsilon(d) \tag{4}$$

where Y(d) is the daily number of records, load mass, or recorded volume; t(d) is calendar time; D(j,d) are weekday indicators with Monday as the reference; and ε(d) is the disturbance term. HC3 heteroskedasticity-robust standard errors were used. The coefficient β_1 is interpreted as an association over the five months, not a causal growth estimate.

3.5 Formal hypotheses and inferential logic

For metric k ∈ {N, M, V}, let F(j,k) denote the weekday-specific distribution for weekday j. The temporal null hypothesis is $H_{0,1}$: F(1,k) = ... = F(6,k), against the alternative that at least one distribution differs. H2 is expressed as $H_{0,2}$: $\beta_1 \leq 0$ against $H_{a,2}$: $\beta_1 > 0$ in the weekday-adjusted model. Concentration H3 is evaluated through Gini/HHI magnitude, cumulative demand shares, and location-level bootstrap intervals. H4 tests $H_{0,4}$: $\rho(M, V) = 0$ against a positive association. H5 tests equality of division-specific order-mass distributions.

When a Kruskal-Wallis test was significant, pairwise weekday comparisons were conducted using two-sided Mann-Whitney tests with Holm family-wise error correction. Rank-biserial correlation summarized pairwise effect size. Non-parametric 95% confidence intervals for Gini coefficients were estimated from 5,000 bootstrap resamples of the 420 location-level observations. Hypotheses were considered supported when the statistical decision and the operational effect measure were directionally consistent; statistical significance alone was not treated as operational relevance.

3.6. Operational indices and pre-optimization model

A three-dimensional vector represents the demand state of active day d:

$$q(d) = [N(d), M(d), V(d)] \tag{5}$$

where N(d) is the number of dispatch records, M(d) is total load mass in kilograms, and V(d) is recorded volume in cubic meters. For weekday j and dimension k, a relative capacity index was calculated as:

$$K(j,k) = \bar{q}(j,k) / \{J^{-1}\Sigma[h = 1 \dots J]\bar{q}(h,k)\} \quad (6)$$

K(j,k) = 1 denotes average weekly intensity, K(j,k) > 1 indicates above-average capacity pressure, and K(j,k) < 1 indicates below-average pressure. The system-level weekday imbalance coefficient is:

$$B(k) = sd_{j}[\bar{q}(j,k)] / mean_{j}[\bar{q}(j,k)] \quad (7)$$

A two-dimensional destination priority score can be defined from relative frequency f(i) and relative mass m(i):

$$S(i;\, \alpha) = \alpha f(i)/\Sigma_{h} f(h) + (1-\alpha)m(i)/\Sigma_{h} m(h),\;\; 0 \leq \alpha \leq 1 \quad (8)$$

For the empirical segmentation, no arbitrary α was imposed. Segment A was defined as the union of the top 20 frequency and top 20 mass rankings; Segment B comprised the additional locations contained in the union of the top 50 rankings; Segment C contained the remainder. This dual-ranking rule prevents frequent light destinations or infrequent heavy destinations from being overlooked.

Future route assignment should respect both mass and volume capacities. For route r on day d, the binding utilization factor is:

$$LF(r,d) = max\{\Sigma_{i} M(i,d)x(i,r,d)/Cmass(r), \Sigma_{i} V(i,d)x(i,r,d)/Cvol(r)\} \leq 1 \quad (9)$$

where x(i,r,d) equals one when destination i is assigned to route r, and Cmass(r) and Cvol(r) are the vehicle mass and volume capacities. The number of vehicles planned for weekday j can be parameterized from the 90th percentile of demand:

$$n*(j) = max\{\lceil Q_{0.90}(M|j)/(\eta mass\; Cmass)\rceil, \lceil Q_{0.90}(V|j)/(\eta vol\; Cvol)\rceil, \lceil Q_{0.90}(N|j)/Smax\rceil\} \quad (10)$$

Here ηmass and ηvol are target utilization ratios, and Smax is the operational stop capacity per vehicle. Once distances are available, route redesign can minimize travel cost, utilization imbalance, and unserved-demand penalties:

$$min\, Z = \Sigma_{r}\Sigma_{i}\Sigma_{h} c(i,h)x(i,h,r,d) + \lambda\Sigma_{r}|LF(r,d) - \bar{L}F(d)| + \pi\Sigma_{i} u(i,d) \quad (11)$$

The present paper estimates the demand-side parameters of this model but does not solve the distance-dependent routing problem because geographic coordinates, vehicle capacities, and trip sequences are absent.

3.7. Ethical and confidentiality considerations

The study used secondary operational records and did not involve direct interaction with human participants. Exact customer addresses were excluded from tables and figures. The operator and customers remain anonymous in the review manuscript. The analysis workbook contains only coded location ranks and aggregated temporal indicators. Before publication, the authors should document permission to use the commercial dataset and confirm whether institutional ethical review or an exemption statement is required.

4. Research results

4.1. Data quality and overall demand profile

Table 2 reports the resulting data quality and descriptive profile of the analytical sample.

Table 2: Dataset quality and descriptive profile

Indicator	Result
Observation period	January 2-May 30, 2026
Analytical records after duplicate removal	4,967
Active service days	125
Unique destination labels	420
Route-label variants	20
Missing mass records	13 (0.26%)
Missing volume records	899 (18.1%)
Total load mass	191,373.64
Total recorded volume	708.34 m^3
Mean/median order mass	38.63 / 8.18
Mean/median recorded volume	0.174 / 0.040 m^3

Note. Load mass is reported in kilograms.

The difference between the mean and median order mass was substantial: 38.63 compared with 8.18. The coefficient of variation for mass was approximately 3.29, confirming a long-tailed distribution with a relatively small number of very large dispatches. Recorded volume displayed a similar pattern, with a mean of 0.174 m^3, a median of 0.040 m^3, and a coefficient of variation close to 3.00. These properties justify the use of medians, concentration measures, and rank-based tests alongside conventional averages.

Most records were generated within Ternopil: 4,703 dispatches, or 94.7% of the analytical sample. These dispatches accounted for 87.8% of total load mass and 90.7% of recorded volume. Deliveries outside the city were less frequent but heavier on average, which indicates that urban frequency and peripheral load intensity should be treated as separate planning dimensions.

4.2. Monthly dynamics

Table 3 presents the active-day-normalized monthly intensity of records, load mass, and recorded volume, calculated using Equation (1).

Table 3: Monthly urban distribution intensity per active service day

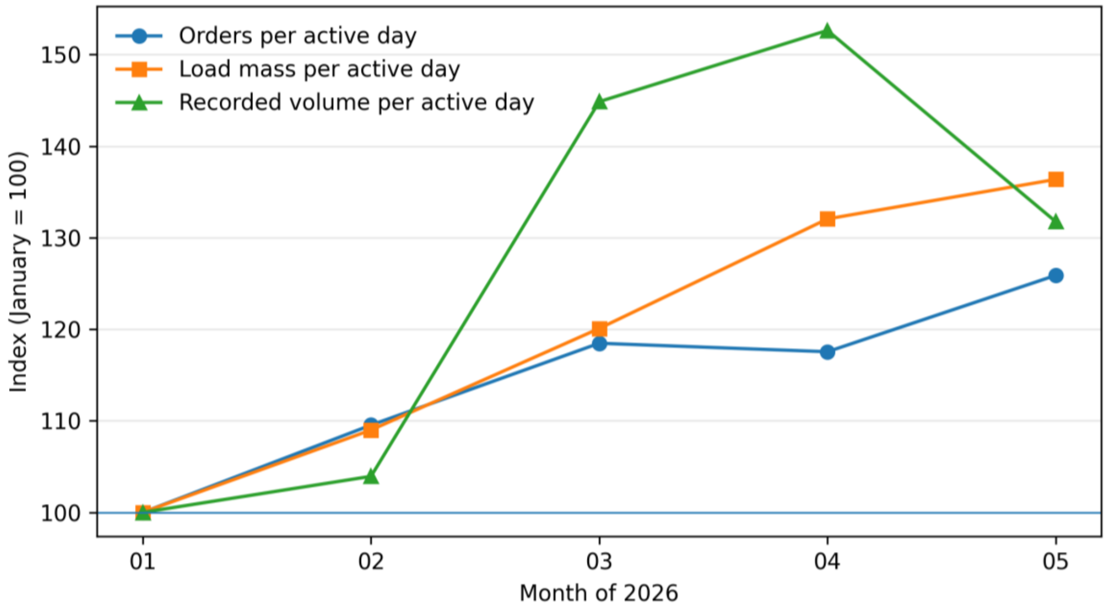
Month	Active days	Records/day	Mass/day	Volume/day, m^3
January	26	34.77	1,282.57	4.48
February	24	38.08	1,397.97	4.66
March	26	41.19	1,540.21	6.50
April	23	40.87	1,693.66	6.84
May	26	43.77	1,749.07	5.91

Source: Authors' calculations. Volume indicators use available recorded values.

Average daily order intensity rose from 34.77 records in January to 43.77 in May, an increase of 25.9%. Average daily load mass increased from 1,282.57 to 1,749.07, or 36.4%. Recorded volume per active day increased by 31.8% between January and May, although its highest monthly value occurred in April. The normalized indicator trajectories are shown in Figure 1.

The weekday-adjusted regression confirmed a positive within-period association. Per 30 calendar days, the estimated increase was approximately 2.31 daily records ($p < 0.001$), 137.9 units of daily load mass ($p = 0.010$), and 0.57 m^3 of recorded daily volume ($p = 0.004$). These results describe the observation window only and should not be extrapolated as a long-term forecast without additional years of data.

Figure 1: Monthly change in average daily distribution demand



Source: Authors' calculations based on the anonymized operational dataset.

4.3. Weekday load imbalance

Table 4 disaggregates the same demand measures by weekday to examine whether the observed growth is evenly distributed across the operating week.

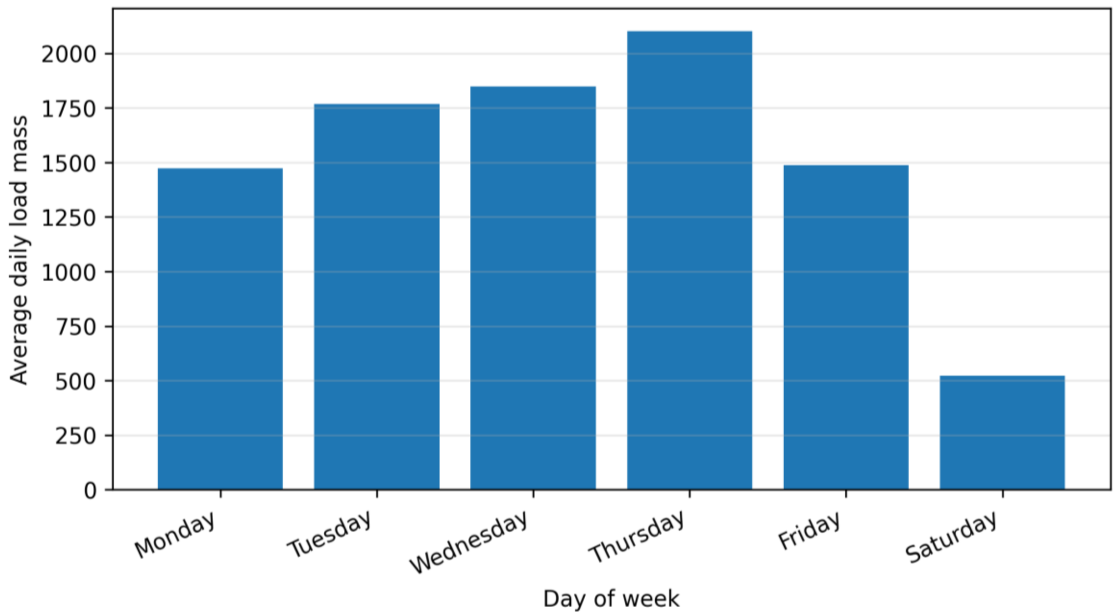
Table 4: Average daily demand by weekday

Weekday	Days	Records/day	Mass/day	Volume/day, m^3
Monday	20	34.75	1,472.72	5.87
Tuesday	20	48.05	1,767.27	6.60
Wednesday	21	45.62	1,847.09	6.49
Thursday	21	46.10	2,100.55	6.99
Friday	22	42.14	1,486.84	6.22
Saturday	21	21.81	522.04	1.86

Source: Authors' calculations. Sunday was not an active service day in the source data.

Tuesday generated the highest average number of dispatch records, at 48.05 per active day, while Thursday had the highest average daily load mass, at 2,100.55. Saturday was structurally different: it averaged 21.81 records and 522.04 units of load mass, far below every weekday from Monday to Friday. Figure 2 visualizes this imbalance.

Figure 2: Average daily load mass by weekday



Source: Authors' calculations based on the anonymized operational dataset.

Table 5 reports the results of the Kruskal-Wallis tests used to evaluate these weekday and divisional differences formally.

Table 5: Non-parametric tests of temporal and organizational differences

Comparison	Kruskal-Wallis statistic	p-value	Effect size
Daily order records by weekday	H = 66.34	< 0.001	$\varepsilon^2 = 0.515$
Daily load mass by weekday	H = 47.65	< 0.001	$\varepsilon^2 = 0.358$
Daily recorded volume by weekday	H = 41.69	< 0.001	$\varepsilon^2 = 0.308$
Order mass by organizational division	H = 202.21	< 0.001	$\varepsilon^2 = 0.040$

Note. Epsilon squared is reported as an effect-size measure. All tests reject equality of group distributions.

The weekday effect was statistically significant for all three daily demand measures. The largest effect was observed for daily record counts ($\varepsilon^2 = 0.515$), followed by load mass ($\varepsilon^2 = 0.358$) and recorded volume ($\varepsilon^2 = 0.308$). Weekday mean dispersion was also operationally substantial: B(N) = 0.251, B(M) = 0.358, and B(V) = 0.336. The peak-to-trough ratio equaled 2.20 for records, 4.02 for mass, and 3.76 for volume. These results support H1 and show that a uniform six-day capacity assumption is not defensible.

Holm-adjusted post-hoc tests refined the omnibus result. Saturday was significantly lower than each Monday-Friday group for daily mass (adjusted $p \leq 0.000278$; absolute rank-biserial effect 0.771-1.000) and volume (adjusted $p \leq 0.000085$; effect 0.819-0.964). No Monday-Friday mass pair remained significant after correction, indicating that the overall mass effect is driven primarily by the structural Saturday reduction rather than by stable separation among all weekdays. For order counts, Saturday differed from every weekday, while Monday was also lower than Tuesday-Friday. This distinction is important for planning: Thursday has the highest mean mass, but the clearest statistical regime change is between Saturday and the working-week pattern.

4.4. Organizational heterogeneity

Table 6 disaggregates record and mass shares by organizational division to test H5.

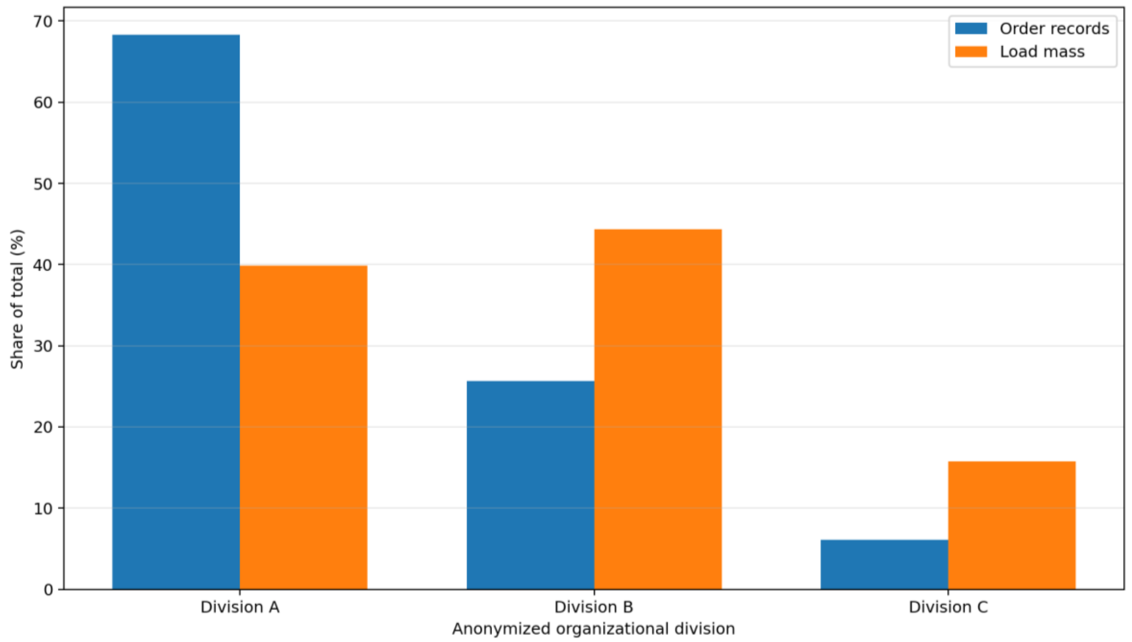
Table 6: Demand structure by organizational division

Division	Record share	Mass share	Mean mass	Median mass	Missing volume
Division A	68.3%	39.9%	22.56	6.60	17.8%
Division B	25.7%	44.4%	66.86	15.52	16.5%
Division C	6.1%	15.7%	100.08	12.60	27.9%

Source: Authors' calculations. Division codes are anonymized operational categories from the source system.

Division A generated 68.3% of all records but only 39.9% of total load mass. Division B generated 25.7% of records and 44.4% of mass, whereas Division C generated 6.1% of records and 15.7% of mass. The mean order mass in Division C was more than four times the mean in Division A. The overall difference among division-level mass distributions was statistically significant ($H = 202.21$, $p < 0.001$), supporting H5; however, $\varepsilon^2 = 0.040$ indicates a modest distribution-level effect. The managerial relevance therefore arises from the large contrast in mean and median shipment sizes and the aggregate mass shares, not from the p-value alone. Figure 3 compares record and load-mass shares.

Figure 3: Distribution structure by organizational division



Source: Authors' calculations based on the anonymized operational dataset.

4.5. Destination concentration and Pareto structure

Table 7 reports the cumulative frequency and mass shares, together with the Gini and HHI coefficients, used to evaluate destination-level concentration under H3.

Table 7: Location-level concentration indicators

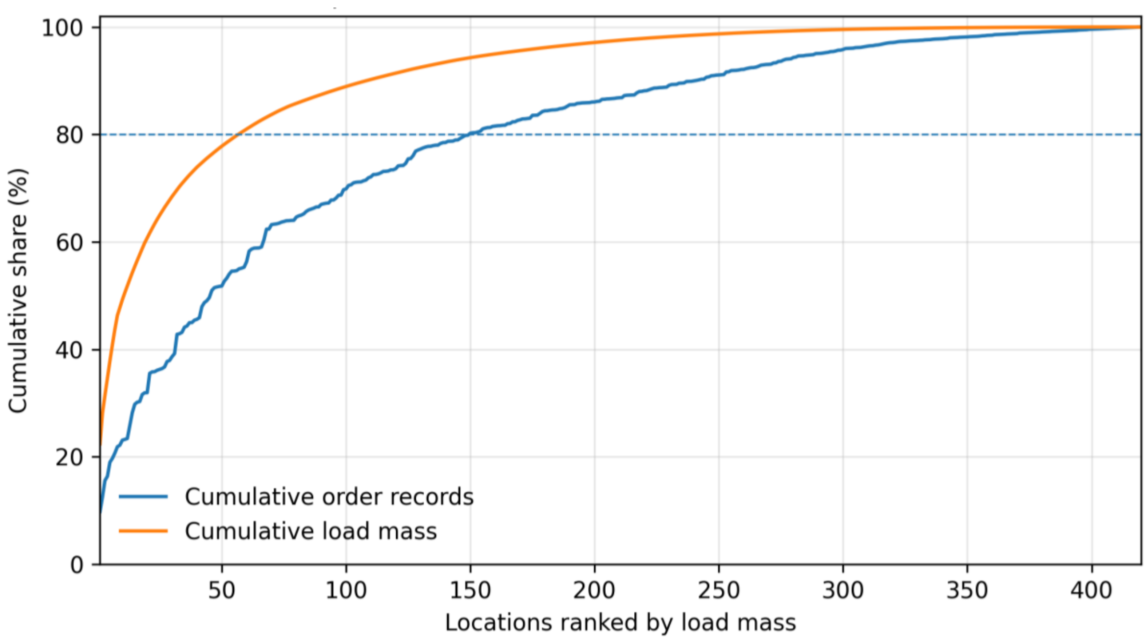
Location group/indicator	Order-frequency result	Load-mass result
Top 1 location by service frequency	9.8%	22.4%*
Top 10 locations by service frequency	34.2%	39.3%
Top 20 locations by service frequency	46.9%	49.0%
Top 50 locations by service frequency	65.9%	64.0%
Gini coefficient	0.729	0.832
HHI	0.020	0.062

Note: *The load-mass share in the first row refers to the single location with the highest load mass; the top-frequency and top-mass location is the same in the source data. Other top-n mass shares use the locations ranked by service frequency unless otherwise stated.

Demand concentration was high under both metrics. The Gini coefficient reached 0.729 for order frequency and 0.832 for load mass. The 5,000-replication location bootstrap produced 95% intervals of 0.673-0.776 and 0.755-0.883, respectively, confirming that the conclusion is not driven by one point estimate. The ten most frequently served locations accounted for 34.2% of all records and 39.3% of total mass, while the ten locations ranked directly by mass accounted for 49.2% of mass. Only 24 of 420 locations were needed to exceed 50% of order records, and 11 locations exceeded 50% of total mass. These findings support H3.

Figure 4 illustrates this concentration through the cumulative Pareto curve for order frequency and load mass.

Figure 4: Spatial concentration of distribution demand



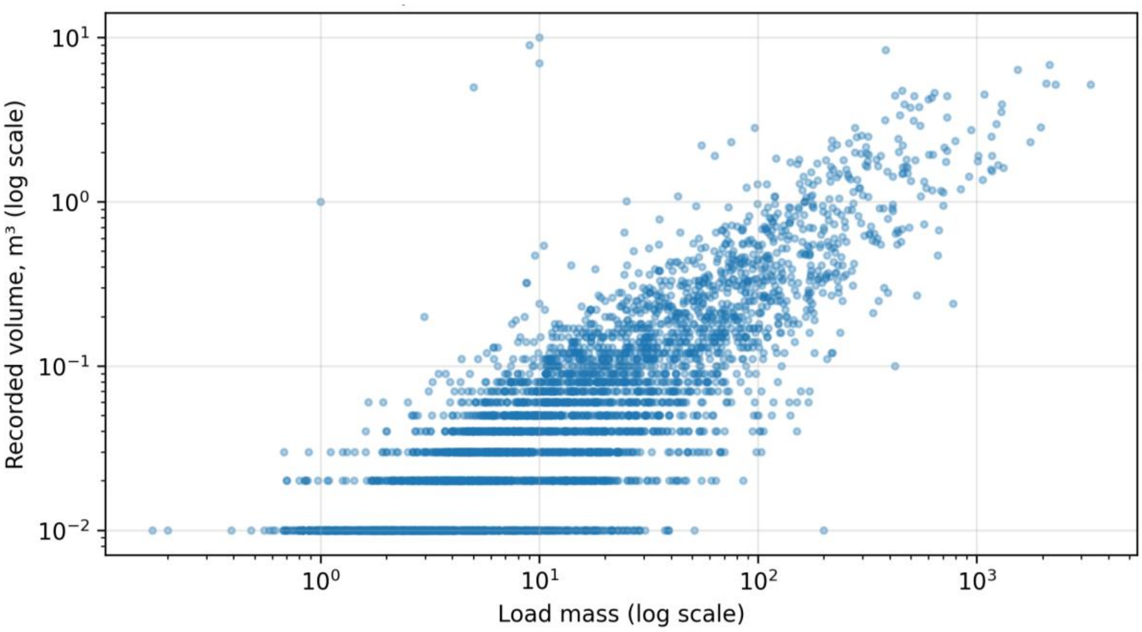
Source: Authors' calculations based on the anonymized operational dataset.

The Pareto curve indicates a dual operating system. A compact core of high-demand locations can support stable service frequencies and preassigned capacity. The remaining locations form a long tail in which fixed routing would likely create low utilization unless stops are consolidated geographically or dynamically assigned.

4.6. Relationship between mass and recorded volume

Complete mass-volume pairs were available for 4,065 records. Pearson's correlation was $r = 0.681$, while Spearman's rank correlation was rho = 0.817; both were statistically significant at $p < 0.001$ and support H4. The higher rank correlation indicates a strong monotonic relationship combined with nonlinearity and outliers. Figure 5 presents the relationship on logarithmic axes.

Figure 5: Relationship between load mass and recorded volume



Source: Authors' calculations based on the anonymized operational dataset.

The association is strong enough to support cross-checking and anomaly detection, but not sufficient to replace missing volume observations automatically. Different product densities can produce similar mass with substantially different space requirements. Vehicle assignment should therefore use both indicators when complete data are available, and the source system should require volume entry for bulky shipments.

4.7. Proposed demand-segmentation framework

The empirical results were translated into a three-segment operating framework. The segments are not universal thresholds; they are decision rules to be recalibrated as more data become available. The initial classification can use cumulative frequency and mass shares, followed by geographic clustering once coordinates are added.

Table 8 translates these empirical concentration results into a three-segment operating logic with corresponding planning actions.

Table 8: Proposed operating logic for location-demand segments

Segment	Empirical rule	Planning logic	Operational action	Expected sustainability mechanism
A: dual-core locations	Union of the top 20 locations by order frequency and the top 20 by load mass.	Demand is recurrent, heavy, or both.	Use fixed service templates, protected capacity, and location-specific delivery windows.	Fewer emergency reallocations; higher load stability and service reliability.
B: consolidatable middle	Additional locations contained in the union of the top 50 frequency and top 50 mass rankings.	Demand is meaningful but less stable than Segment A.	Batch by sector and time window after order cut-off; permit controlled route rotation.	Higher stop density and fewer fragmented low-load departures.
C: dynamic long tail	All remaining locations outside the dual top-50 set.	Demand is individually sparse and highly dispersed.	Apply dynamic insertion, designated delivery days, pickup options, or minimum-	Lower probability of dedicated low-utilization trips.

			shipment rules where contracts allow.	
Temporal overlay	Weekday-specific capacity index and 90th-percentile demand threshold.	The same segment requires different resources by day.	Protect stop capacity on Tuesday and mass capacity on Wednesday-Thursday; reduce the Saturday core fleet.	Better weekly balancing and lower peak-day overflow risk.

Source: Authors' framework derived from the statistical results. Actual environmental effects require route, vehicle, and fuel data.

Applying the dual-ranking rule produced an empirically compact core. Segment A contains 30 locations, only 7.1% of the destination set, but generates 50.8% of records and 65.0% of load mass. Segment B contains 44 locations and contributes a further 19.7% of records and 17.3% of mass. Segment C contains 346 locations, or 82.4% of all destinations, yet accounts for only 17.7% of mass. This structure supports different service rules while preserving both stop-frequency and mass constraints.

Table 9 quantifies the resulting segment composition in terms of location count, record share, and mass share.

Table 9: Observed composition of the dual-criteria demand segments

Segment	Locations	Location share	Record share	Mass share
A: Dual core	30	7.1%	50.8%	65.0%
B: Consolidatable middle	44	10.5%	19.7%	17.3%
C: Dynamic long tail	346	82.4%	29.5%	17.7%

Note. Segment A is the union of the top 20 locations by frequency and mass. Segment B is the additional union of ranks 21-50; Segment C is the remainder.

4.8. Hypothesis evaluation

Table 10 consolidates the inferential decisions and separates statistical evidence from operational magnitude.

Table 10: Summary of statistical hypotheses and operational interpretation

Hypothesis	Statistical evidence	Decision	Operational interpretation
H1: weekday distributions differ	H = 41.69-66.34; all $p < 0.001$; $\varepsilon^2 = 0.308\text{-}0.515$	Supported	Large weekday effects; post-hoc separation is dominated by Saturday.
H2: positive weekday-adjusted time trend	30-day effect: +2.31 records, +137.9 kg, +0.57 m^3; $p \le 0.010$	Supported within the study window	Association is not a causal or long-term forecast.
H3: destination demand is concentrated	Gini = 0.729 records and 0.832 mass; bootstrap lower limits 0.673 and 0.755	Supported	A 7.1% dual core generates 65.0% of mass.
H4: mass and volume are positively related	r = 0.681; rho = 0.817; $p < 0.001$	Supported	Strong for validation, insufficient for automatic volume imputation.
H5: division mass distributions differ	H = 202.21; $p < 0.001$; $\varepsilon^2 = 0.040$	Supported	Statistically clear but modest rank-based effect.

Source: Authors' calculations. Hypotheses are interpreted jointly with effect sizes and decision relevance.

4.9. Weekday capacity parameters

The relative mass-capacity index K(j, M) ranges from 0.341 on Saturday to 1.370 on Thursday. Consequently, a plan based on the six-day mean would understate Thursday mass pressure by 37.0% and overstate Saturday pressure by almost threefold. Table 11 provides 90th-percentile planning thresholds. These values are not vehicle counts; they are demand thresholds to be converted into vehicles through Equation (10) after capacity and target-utilization parameters are supplied.

Table 11 converts the weekday capacity index K(j,k) from Equation (6) into 90th-percentile planning thresholds for records, mass, and volume.

Table 11: Weekday capacity indices and 90th-percentile planning thresholds

Weekday	Mass index	P90 records	P90 mass, kg	P90 volume, m^3	Planning role
Monday	0.961	42.3	2,434	9.67	Standard capacity; protect contingency for bulky orders.
Tuesday	1.153	63.1	2,415	12.58	Highest stop-handling capacity; early route release.
Wednesday	1.205	54.0	2,698	10.01	High mass capacity; reserve second-wave flexibility.
Thursday	1.370	54.0	3,402	12.68	Peak mass day; maximum fleet and loading-bay readiness.
Friday	0.970	51.0	2,322	8.61	Standard capacity with end-of-week exception buffer.
Saturday	0.341	27.0	735	2.92	Reduced core fleet; use spare time for deferred or reverse flows.

Source: Authors' calculations. P90 is the empirical 90th percentile of active-day demand for the corresponding weekday.

Stop-handling and payload constraints should be planned jointly. Tuesday has the highest P90 record count (63.1), while Thursday has the highest P90 mass (3,402 kg) and volume (12.68 m^3). Thus, the same vehicle count may be inadequate for different reasons on different days: route duration and service capacity dominate Tuesday, whereas payload and cubic capacity dominate Thursday.

5. Discussion of the Results

5.1. Interpretation of temporal imbalance

The weekday pattern is one of the most actionable findings. Tuesday through Thursday combine high order intensity with high load mass, while Saturday operates at a distinctly lower level. This means that a uniform six-day capacity plan would either underprovide capacity during the midweek peak or leave capacity underused on Saturday. A better approach is to forecast and allocate capacity by service day, not by a weekly average. The result is consistent with the broader routing literature, which treats demand as time-dependent and argues for flexible policies when order information changes during the planning horizon (Mardešić et al., 2024).

The positive five-month trend reinforces the need for periodic recalibration. The observed increase may reflect business growth, seasonality, changes in customer behavior, or recording practices; the available data cannot distinguish among these explanations. Therefore, the regression should be viewed as an early-warning indicator for capacity planning rather than a forecast. A full seasonal model would require at least one complete year and preferably multiple years.

5.2. Interpretation of concentration and organizational heterogeneity

The concentration results show that order frequency and mass are related but not interchangeable. A location can be visited frequently with small consignments or less frequently with heavy consignments. Route design based only on stop counts can therefore overload vehicles, while design based only on mass can ignore service-frequency constraints. The combined use of both metrics is a practical extension of establishment-level freight demand approaches that distinguish shipment generation from vehicle movement (Balla et al., 2023; Nuzzolo & Comi, 2014).

The organizational results provide an additional layer. Division A produces many small records, whereas Divisions B and C generate fewer but heavier consignments. A single order-handling and dispatch rule is unlikely to serve these profiles efficiently. Organizational origin can be used as a prior variable for vehicle assignment, cut-off times, and consolidation: high-frequency small orders may be batched, while heavy orders require earlier capacity reservation.

5.3. Contribution to sustainable urban logistics

The sustainability mechanism proposed here is operational rather than technological. Better segmentation can improve load factors and reduce redundant service, thereby creating conditions for lower vehicle kilometers, fuel use, emissions, and curb occupancy. Literature on smart urban freight similarly emphasizes that data analytics and consolidation can improve energy efficiency through better planning and infrastructure sharing (Golinska-Dawson & Sethanan, 2023; Pan et al., 2021). However, this paper does not quantify those benefits because the dataset lacks route distance, travel time, vehicle type, fuel consumption, and emission factors.

This distinction is important for methodological credibility. It would be inappropriate to convert a reduction in the number of planned stops directly into an emissions estimate without knowing whether vehicles are rerouted, how far they travel, and how loaded they are. The present results define where potential savings are most likely: the recurrent core is suitable for route templates and consolidated time windows; the long tail is suitable for dynamic insertion or service-day consolidation. A follow-up experiment can compare baseline and redesigned routes using vehicle-kilometer, load-factor, fuel, and service-reliability indicators.

5.4. Managerial and policy implications

For the operator, the immediate priority is to convert the statistical results into a controlled planning cycle. First, capacity should be assigned by weekday using the maximum of stop, mass, and volume requirements in Equation (10), rather than by one weekly average. Second, the 30 Segment A locations should receive stable service templates and protected capacity; the 44 Segment B locations should be released in geographic batches after order cut-off; and the 346 Segment C locations should be dynamically inserted or shifted to designated consolidation days where contractual service levels permit. Third, organizational source should be included in dispatch rules because Division A is stop-intensive, whereas Divisions B and C are payload-intensive.

For urban planners, anonymized concentration indicators can support freight policy without disclosing commercial relationships. High-frequency zones may justify designated loading space, coordinated delivery windows, micro-consolidation facilities, or data-sharing pilots. Diana et al. (2020) showed the value of identifying priority loading areas, and Castrellon and Sanchez-Diaz (2024) emphasized that curb interventions should be assessed across stakeholder and sustainability outcomes. The proposed framework can therefore operate at both company and city levels once location coordinates are available.

A pilot should use a baseline-versus-redesign comparison rather than assume savings. Let L, T, and F denote vehicle-kilometers, route time, and fuel use. The direct operating-cost effect is:

$$\Delta C = cL[L(0) - L(1)] + cT[T(0) - T(1)] + cF[F(0) - F(1)] \tag{12}$$

where subscripts 0 and 1 denote the baseline and redesigned systems. The corresponding tank-to-wheel emission change is:

$$\Delta E = EFF[F(0) - F(1)] \tag{13}$$

where EFF is the verified emission factor for the actual fuel and fleet; positive savings should be reported only when service reliability is maintained and confidence intervals from repeated operating weeks exclude zero.

Table 12 sequences these validation steps into a phased implementation roadmap with measurable decision rules.

Table 12: Phased implementation and validation roadmap

Horizon	Action	Core KPI	Decision rule
0-1 month	Standardize route labels and destination identifiers; enforce kg and m^3 metadata.	Duplicate rate; missing volume share; valid address-code share.	Duplicate rate <0.1%; missing volume <5%; 100% coded destinations.
1-3 months	Introduce weekday-specific P90 capacity calendar and exception dashboard.	Peak overflow; unused capacity; on-time departure.	Reduce peak-day emergency reallocations without increasing missed service.
3-6 months	Pilot A/B/C service logic on matched operating weeks.	Vehicle-km, stops/route, kg/route, m^3/route, on-time delivery.	Adopt only if distance or time falls while service reliability is non-inferior.
6-12 months	Geocode destinations and solve the bi-capacity routing model with telemetry.	Cost/order, fuel/order, CO_2e/order, load factor, route-time variability.	Scale the design after statistically and operationally significant improvement.

Source: Authors' recommendations based on the estimated demand structure. Thresholds should be adapted to contractual service levels.

5.5. Limitations and future research

The study has seven main limitations. First, it covers one operator and five months, limiting external and seasonal validity. Second, the recorded kilogram values were not independently verified against calibrated weighing records. Third, volume missingness is substantial and may not be random. Fourth, destination labels are not geocoded, preventing distance-based clustering. Fifth, the source system records dispatch demand rather than completed trip trajectories, so failed deliveries, route sequence, service time, and empty mileage are unknown. Sixth, vehicle capacities and fleet composition are unavailable, so Equation (10) cannot yet be converted into a required vehicle count. Seventh, the observational design cannot establish that the proposed segmentation causes cost or environmental improvement.

Future research should extend the period to at least twelve months, geocode anonymized destinations, and integrate fleet telemetry. A baseline vehicle-routing model can then be compared with a hybrid model containing fixed A-segment routes and dynamically generated B-C routes. Performance should be evaluated using total distance, travel time, load factor by mass and volume, on-time delivery, number of trips, fuel consumption, CO_2-equivalent emissions, and robustness under daily demand uncertainty. Scenario analysis may also test micro-depots, pickup points, cargo bikes, or electric vans where demand density is sufficient (Calabrò et al., 2023; Mantecchini et al., 2025).

6. Conclusions

This paper transformed 4,967 anonymized delivery records into a statistically tested pre-optimization model of urban freight demand in Ternopil, directly addressing all four research questions posed at the outset. Weekday distributions differed materially (RQ1), destination demand was highly concentrated (RQ2), organizational divisions generated distinct shipment-size profiles (RQ3), and the resulting capacity and segmentation rules translate directly into an operational decision framework (RQ4). All five hypotheses were supported: weekday distributions differed, the short-period trend was positive, destination demand was concentrated, mass and volume were strongly associated, and division-level shipment profiles diverged. The post-hoc analysis showed that Saturday constitutes the clearest separate operating regime, while midweek differences are better interpreted through capacity indices than through claims of universal pairwise separation.

From a scientific standpoint, the theoretical contribution is a unified demand-side architecture linking the daily vector q(d), weekday capacity index K(j,k), imbalance coefficient B(k), dual-criteria location segmentation, and mass-volume route feasibility. This extends prior establishment-level freight-generation models (Balla et al., 2023; Nuzzolo & Comi, 2014) and responds to calls for disaggregated, decision-oriented demand data (Bjørgen et al., 2025; Diana et al., 2020) by showing how temporal, spatial, and organizational heterogeneity can be combined within one framework rather than treated separately. The approach also complements dynamic routing research (Mardešić et al., 2024) by identifying, ex ante, which locations warrant fixed templates and which require flexible assignment.

From a practical standpoint, the principal recommendation is a quantified hybrid operating model. Thirty dual-core locations should receive fixed service templates because they generate 50.8% of records and 65.0% of mass; forty-four middle-segment locations should be geographically consolidated; the remaining 346 locations should be dynamically managed. Capacity should be day-specific, since relative mass pressure ranges from 0.341 on Saturday to 1.370 on Thursday, with stop pressure peaking on Tuesday. A route redesign should be accepted only after matched-week testing demonstrates lower distance, time, fuel, cost, or emissions without deterioration in delivery reliability.

The study is limited by its single-operator, five-month scope, unverified weighing records, substantial volume missingness, the absence of geocoded destinations, and the lack of vehicle-capacity and fleet data, which together prevent the demand-side model from being converted into a fully solved routing solution. Future research should extend the observation period to at least twelve months, geocode destinations, and integrate fleet telemetry to test the proposed bi-capacity routing model against a baseline using distance, load factor, fuel, and emissions outcomes. Taken together, the findings show that sustainable last-mile performance depends as much on recognizing when and where demand concentrates as on the vehicles used to serve it.

Funding

This research received no external funding

Conflicts of interest

The authors declare no conflict of interest

Data availability statement

The operational dataset is not publicly available because it contains commercially sensitive delivery information. An anonymized aggregate dataset and the statistical code may be made available by the corresponding author upon reasonable request and subject to data-owner approval.

Declaration of generative AI and AI-assisted technologies

During manuscript preparation, OpenAI ChatGPT was used for document structuring, English-language editing, statistical workflow checking, and preparation of the working Word layout. All numerical results were calculated from the supplied dataset, and the authors must independently verify the analysis, references, interpretations, and final wording. The authors retain full responsibility for the submitted manuscript.

Citation information

Tson, O., Vovk, Y., Rozhko, N., Lyashuk, O., & Plekan, U. (2026). Demand concentration and temporal imbalance in last-mile urban freight distribution. *Journal of Sustainable Development of Transport and Logistics, 11*(1), 153-171. doi:10.14254/jsdtl.2026.11-1.8

References

Balla, B. S., Sahu, P. K., Pani, A., Sharma, S., & Majumdar, B. B. (2023). Comparison of parametric and non-parametric methods for modeling establishment-level freight generation. *Transportation Research Record, 2677*(2), 154-172. https://doi.org/10.1177/03611981221116369

Bjørgen, A., Pitera, K., Karlsson, H., Babri, S., & Strømmen, K. (2025). Knowledge and strategies for facilitating urban freight and last-mile deliveries amid digital transitions. *Research in Transportation Economics, 113*, 101625. https://doi.org/10.1016/j.retrec.2025.101625

Bosona, T. (2020). Urban freight last mile logistics-Challenges and opportunities to improve sustainability: A literature review. *Sustainability, 12*(21), 8769. https://doi.org/10.3390/su12218769

Calabrò, G., Le Pira, M., Giuffrida, N., Fazio, M., Inturri, G., & Ignaccolo, M. (2023). A spatial agent-based model of e-commerce last-mile logistics towards a delivery-oriented development. *Transportation Research Interdisciplinary Perspectives, 21*, 100895. https://doi.org/10.1016/j.trip.2023.100895

Castrellon, J. P., & Sanchez-Diaz, I. (2024). Effects of freight curbside management on sustainable cities: Evidence and paths forward. *Transportation Research Part D: Transport and Environment, 131*, 104165. https://doi.org/10.1016/j.trd.2024.104165

Diana, M., Pirra, M., & Woodcock, A. (2020). Freight distribution in urban areas: A method to select the most important loading and unloading areas and a survey tool to investigate related demand patterns. *European Transport Research Review, 12*, 40. https://doi.org/10.1186/s12544-020-00430-w

Fontaine, P., Minner, S., & Schiffer, M. (2023). Smart and sustainable city logistics: Design, consolidation, and regulation. *European Journal of Operational Research, 307*(3), 1071-1084. https://doi.org/10.1016/j.ejor.2022.09.022

Golinska-Dawson, P., & Sethanan, K. (2023). Sustainable urban freight for energy-efficient smart cities-Systematic literature review. *Energies, 16*(6), 2617. https://doi.org/10.3390/en16062617

Liu, S., Xu, J., Shi, X., Li, G., & Liu, D. (2018). Sustainable distribution organization based on the supply-demand coordination in large Chinese cities. *Sustainability, 10*(9), 3042. https://doi.org/10.3390/su10093042

Lyons, T., & McDonald, N. C. (2023). Last-mile strategies for urban freight delivery: A systematic review. *Transportation Research Record, 2677*(1), 1141-1156. https://doi.org/10.1177/03611981221103596

Mantecchini, L., Nanni Costa, F. P., & Rizzello, V. (2025). Last mile urban freight distribution: A modelling framework to estimate e-cargo bike freight attraction demand share. *Future Transportation, 5*(1), 31. https://doi.org/10.3390/futuretransp5010031

Mardešić, N., Erdelić, T., Carić, T., & Đurasević, M. (2024). Review of stochastic dynamic vehicle routing in the evolving urban logistics environment. *Mathematics, 12*(1), 28. https://doi.org/10.3390/math12010028

Maxner, T., Dalla Chiara, G., & Goodchild, A. (2025). The state of sustainable urban last-mile freight planning in the United States. *Journal of the American Planning Association, 91*(1), 88-101. https://doi.org/10.1080/01944363.2024.2324096

Middela, M. S., & Ramadurai, G. (2024). Effect of the measurement period and spatial dependence on the accuracy of urban freight trip generation models. *Transportation Research Part A: Policy and Practice, 179*, 103884. https://doi.org/10.1016/j.tra.2023.103884

Nuzzolo, A., & Comi, A. (2014). Urban freight demand forecasting: A mixed quantity/delivery/vehicle-based model. *Transportation Research Part E: Logistics and Transportation Review, 65*, 84-98. https://doi.org/10.1016/j.tre.2013.12.014

Pan, S., Zhou, W., Piramuthu, S., Giannikas, V., & Chen, C. (2021). Smart city for sustainable urban freight logistics. *International Journal of Production Research, 59*(7), 2079-2089. https://doi.org/10.1080/00207543.2021.1893970

Plazier, P., Rauws, W., Neef, R., & Buijs, P. (2024). Towards sustainable last-mile logistics? Investigating the role of cooperation, regulation, and innovation in scenarios for 2035. *Research in Transportation Business & Management, 56,* 101198. https://doi.org/10.1016/j.rtbm.2024.101198