Journal of Information System Exploration and Research Vol. 4, No. 2, April 2026, pp. 15-34

SHM PUBLISHER
SURYA HIJAU MANFAAT

Journal of Information System Exploration and Research

https://joiser.shmpublisher.com

p-ISSN 2964-1160 | e-ISSN 2963-6361

Risk-Aware Multi-Capacity Routing for Shift-Feasible Cosmetic Distribution

Ayub Prasetyo[1*], Firda Amalia[2], Sarifuddin Madenda[3], Ernastuti[4], Murni[5]
[1,2,3,4,5]*Department of Information Technology, Universitas Gunadarma, Indonesia*

DOI: https://doi.org/10.52465/joiser.v4i2.62
Received 02 June 2026; Accepted 19 June 2026; Available online 21 June 2026

Article Info

Keywords:
Risk-Aware Routing
Multi-Capacity
Shift-Feasible Routing
Soft Time Window
Cosmetic Distribution

Abstract

Urban cosmetic distribution requires routing decisions that satisfy delivery time windows, vehicle capacity, and working-hour constraints. Conventional routing models often rely on deterministic travel-time assumptions and single-capacity limits, which may obscure lateness risk and volume overload. This study proposes a risk-aware multi-capacity routing framework for shift-feasible cosmetic distribution. Customer and depot coordinates were used to estimate distances with the Haversine formula, adjusted by a circuity factor and converted into travel time. The model applies split-delivery preprocessing for oversized records, enforces weight and volume capacities, and distinguishes delivery trips from vehicle requirements within a 09:00–17:00 shift. Three routing models were evaluated: deterministic single-capacity, distance-based multi-capacity, and risk-aware multi-capacity routing. Results show that oversized records were split, eliminating unserved and infeasible customers. The single-capacity model produced multiple over-volume cases, while the multi-capacity model removed capacity violations without significantly affecting lateness or resource needs. The risk-aware model reduced lateness and late customers but increased travel time, delivery trips, and vehicle requirements. Sensitivity analysis confirmed stable conclusions across key assumptions. These findings highlight improved delivery reliability alongside trade-offs between service performance and resource usage.

1. Introduction

Urban distribution requires routing decisions that are not only efficient but also feasible under real operational constraints. In daily delivery operations, route performance is affected by customer location, service time, vehicle capacity, delivery time windows, travel duration, and driver working hours. The Vehicle Routing Problem with Time Windows has therefore become an important model for

[*] ***Corresponding Author:***

Ayub Prasetyo,
Department of Information Technology,
Universitas Gunadarma,
Jl. Margonda Raya No. 100, Kec. Pondok Cina, Kel. Beji, Kota Depok, Jawa Barat - Indonesia, 16424.
Email: ayub.prasetyo@staff.gunadarma.ac.id

logistics planning because customer service must be completed within specific time intervals [1], [2]. Classical vehicle routing studies emphasize distance, cost, and fleet assignment [3], [4], while recent studies show that urban delivery problems are increasingly multi-constrained, involving soft time windows, heterogeneous demand, vehicle capacity, travel-time variation, and service reliability [5], [6], [7], [8]. A route with a short distance may still be operationally weak if it produces late deliveries, exceeds vehicle capacity, or cannot be completed within the working shift. Therefore, routing evaluation should not rely only on distance minimization, but should also consider lateness risk, loading feasibility, and shift feasibility.

Cosmetic product distribution has a specific operational challenge because customer orders are heterogeneous in weight and volume. A delivery unit may remain feasible based on weight capacity but exceed the available vehicle space. Conversely, a vehicle may still have available space but carry a relatively heavy load. Previous studies on vehicle routing with loading constraints show that route feasibility can be affected by more than one loading dimension, especially when products have different physical characteristics [9], [10], [11], [12], [13]. This issue becomes more critical when customer-date demand exceeds single-vehicle capacity and requires split delivery or additional delivery units [14], [15], [16]. If routing evaluation only uses weight capacity, the resulting route may appear feasible mathematically while still being infeasible from a loading perspective. For this reason, cosmetic distribution requires a routing framework that evaluates weight capacity, volume capacity, and split-delivery feasibility simultaneously.

Travel time is another critical factor because it directly determines arrival time and lateness. Several studies have addressed travel-time variation through time-dependent routing, stochastic routing, and road-network-based approaches [7], [8], [17], [18]. However, direct road-network data or historical traffic observations are not always available, stable, or computationally practical for applied routing studies. In such conditions, coordinate-based distance approximation can provide a reproducible alternative for route evaluation. The Haversine formula is commonly used to estimate geographical distance between two latitude-longitude points [19], while circuity factors can approximate the deviation between direct geographical distance and operational travel distance in urban networks [20], [21]. Based on this rationale, this study uses distance-based travel-time scenarios to evaluate route vulnerability. The scenarios are not intended to predict actual traffic conditions or represent historical traffic percentiles, but to provide controlled travel-time levels for assessing lateness risk.

Heuristic and learning-based routing methods are widely used for complex vehicle routing variants because exact optimization becomes difficult when time windows, capacity constraints, uncertainty, and operational restrictions are considered simultaneously [5], [8] [22], [23], [24], [25]. However, many applied routing models still emphasize distance minimization, travel-time optimization, or single-capacity feasibility as separate concerns. Limited studies have jointly examined lateness-risk scenario escalation, split-delivery preprocessing, simultaneous weight-volume capacity feasibility, and physical vehicle utilization under a fixed working shift in a reproducible coordinate-based routing framework. This gap is important because delivery trips and physical vehicles are not always equivalent. A vehicle may perform more than one delivery trip within the same working day if it can return to the depot, reload, and depart again before the shift ends.

Previous studies have addressed vehicle routing with time windows, split delivery, loading constraints, time-dependent routing, and heuristic route construction. However, these components are commonly examined as separate modeling concerns. Many routing studies focus on reducing distance or travel time, while loading feasibility, lateness-risk escalation, and physical vehicle reuse under a fixed daily working shift are not always evaluated simultaneously. This separation may produce routing results that appear feasible from one perspective but remain operationally weak from another, for example routes that satisfy weight capacity but violate volume capacity, or routes that reduce travel distance but cannot be completed within the working shift.

This study addresses this gap by proposing an integrated routing evaluation framework for shift-feasible cosmetic distribution. The scientific contribution of this study is not the development of a new optimization algorithm, but the integration of four operational components into one reproducible framework: split-delivery preprocessing for oversized demand, simultaneous weight-volume capacity feasibility, lateness-risk-based travel-time scenario escalation, and multi-trip physical vehicle utilization within a 09:00–17:00 working shift. This integration enables the framework to evaluate hidden loading infeasibility, lateness exposure, and the trade-off between service reliability and operational resource requirements.

This study proposes a risk-aware multi-capacity routing framework for shift-feasible cosmetic distribution. The framework estimates geographical distance using the Haversine formula, adjusts it using a circuity factor, and converts it into travel time using a baseline speed. Oversized customer-date records are first handled using split-delivery preprocessing so that each delivery unit can satisfy single-vehicle capacity. The model then evaluates three routing configurations: deterministic single-capacity routing, distance-based multi-capacity routing, and risk-aware multi-capacity routing. The final framework distinguishes delivery trips from physical vehicle requirements by allowing vehicles to be reused within a 09:00–17:00 working shift. The new value of this study lies in the integration of distance-based travel-time scenarios, lateness-risk-based scenario escalation, split-delivery preprocessing, two-dimensional capacity feasibility, and shift-feasible vehicle utilization. This study does not aim to develop a new optimization algorithm, but to provide a transparent and reproducible routing evaluation framework for assessing delivery feasibility, lateness vulnerability, and operational resource trade-offs in urban cosmetic distribution.

2. Literature Review

To provide a more detailed understanding of the relevant concepts, the following subsection discusses the Vehicle Routing Problem with Time Windows.

2.1 Vehicle Routing with Time Windows

The Vehicle Routing Problem with Time Windows is widely used to model delivery operations in which customers must be served within specific time intervals. This model is relevant for urban logistics because route feasibility depends not only on distance and fleet capacity, but also on arrival time, service duration, and customer availability [1], [2]. Classical vehicle routing models provide the foundation for route planning and fleet assignment [3], [4], while recent studies show that modern delivery systems increasingly involve multiple constraints, such as soft time windows, demand uncertainty, heterogeneous vehicles, and service reliability [5], [6], [7], [8].

In practical delivery operations, time-window violations do not always make a route impossible, but they indicate a decline in service quality. Therefore, soft time-window modeling is often used to allow late delivery while measuring lateness as a penalty or reliability indicator [5], [6]. This study adopts this perspective by treating lateness as an evaluative outcome rather than an absolute route rejection criterion. The performance of each routing model is therefore assessed using late customers, lateness rate, total lateness, and maximum lateness.

2.2 Multi-Capacity and Loading Feasibility

Vehicle capacity is often represented using a single dimension, commonly weight. However, this simplification may be insufficient for product distribution involving heterogeneous physical characteristics. Cosmetic products may differ in packaging size, volume, and weight, making it possible for a vehicle to satisfy weight capacity while exceeding volume capacity. Studies on vehicle routing with loading constraints show that route feasibility can be affected by loading dimensions and product arrangement, especially when demand cannot be represented by weight alone [9], [10], [11], [12], [13].

Two-dimensional loading and multi-capacity routing provide a more realistic representation of vehicle feasibility because they consider more than one capacity dimension. In this study, capacity feasibility is evaluated using both weight utilization and volume utilization. This approach differs from single-capacity routing models because a route is considered feasible only when both capacity dimensions remain within vehicle limits. The distinction is important for cosmetic distribution, where route plans that appear feasible based on weight may still be operationally infeasible due to volume overload.

2.3 Split Delivery for Oversized Demand

Split delivery is relevant when customer demand exceeds the capacity of a single vehicle or when dividing demand can improve routing feasibility. Previous studies have examined split delivery in relation to loading constraints, heterogeneous fleets, and delivery options [14], [15], [16]. In practical logistics, split delivery allows a large customer order to be served through more than one delivery unit instead of being treated as infeasible.

This study applies split-delivery preprocessing to customer-date records whose aggregated demand exceeds single-vehicle capacity. The purpose is not to artificially improve route performance, but to ensure that oversized delivery demand is transformed into feasible delivery units before route construction. This preprocessing step is necessary because a routing model cannot realistically serve a customer-date demand that exceeds both vehicle capacity and loading feasibility in one trip. Therefore, split delivery functions as a feasibility correction before comparing the routing models.

2.4 Distance-Based Travel-Time Scenario

Travel time is a critical component of vehicle routing with time windows because it directly affects arrival time and lateness. Some routing studies model travel-time variation using time-dependent routing, stochastic routing, historical traffic data, or road-network-based computation [7], [8], [17], [18]. These approaches are useful when reliable traffic or network data are available. However, such data may not always be accessible, stable, or computationally practical for applied distribution studies.

In situations where direct road-network data are not used, coordinate-based distance approximation can be applied as a reproducible alternative. The Haversine formula is commonly used to estimate geographical distance between two latitude-longitude points [19]. Since straight-line distance does not fully represent actual travel distance, a circuity factor can be used to approximate the difference between geographical distance and operational travel distance [20], [21]. Based on this logic, this study estimates travel time from adjusted coordinate-based distance and evaluates three travel-time scenarios: normal, moderate-risk, and high-risk conditions. These scenarios are used for sensitivity-based lateness evaluation, not for claiming actual traffic prediction.

2.5 Heuristic Routing and Risk-Aware Evaluation

Vehicle routing problems become computationally challenging when time windows, capacity constraints, travel-time variation, and operational restrictions are considered simultaneously. For this reason, heuristic and learning-based approaches are widely used to obtain practical routing solutions within reasonable computational effort [5], [8], [22],[23], [24], [25]. In applied routing studies, heuristic methods are especially useful when the objective is not to prove mathematical optimality, but to evaluate routing feasibility and operational trade-offs.

This study uses a Parallel Nearest Neighbor with Soft Time Windows heuristic as the route construction engine. The heuristic selects customers based on travel time, waiting time, predicted lateness, due-time urgency, and capacity feasibility. However, the contribution of this study does not lie in proposing a new optimization algorithm. Instead, the contribution lies in embedding route construction into a risk-aware multi-capacity evaluation framework. The model first evaluates route performance under normal travel-time conditions and then uses the initial lateness rate to determine whether a more conservative travel-time scenario is required.

2.6 Shift-Feasible Vehicle Utilization

In many routing studies, the number of routes is often interpreted as the number of vehicles. This assumption may be too simplistic for daily distribution operations because one physical vehicle can perform more than one delivery trip within the same working shift. A route or trip represents one depot-customer-depot cycle, while a physical vehicle represents an operational resource that may be reused if sufficient working time remains.

This study distinguishes delivery trips from physical vehicles required. A vehicle is allowed to perform multiple trips if it can return to the depot, complete reload time, and start another trip within the 09:00–17:00 working shift. This distinction is important because multi-capacity routing may increase the number of delivery trips, but not necessarily in the same proportion as physical vehicle requirements. Therefore, shift-feasible vehicle utilization provides a more realistic basis for interpreting routing results and operational resource trade-offs.

2.7 Research Position

Based on the reviewed literature, previous studies have addressed vehicle routing with time windows, soft time windows, time-dependent routing, loading constraints, split delivery, and heuristic route construction. However, these components are often discussed separately. Many studies emphasize route distance, travel-time optimization, or capacity feasibility without jointly evaluating lateness-risk escalation, split-delivery preprocessing, simultaneous weight-volume capacity, and physical vehicle utilization under a fixed working shift.

This study fills this gap by proposing a risk-aware multi-capacity routing evaluation framework for shift-feasible cosmetic distribution. The framework integrates distance-based travel-time scenarios, split-delivery preprocessing, soft time-window lateness evaluation, weight-volume capacity feasibility, and multi-trip physical vehicle utilization. The contribution of this study lies in showing how these components interact in a daily delivery setting, particularly in identifying hidden volume overload, evaluating lateness vulnerability, and distinguishing delivery trips from physical vehicles. Therefore, the proposed framework provides a more realistic basis for interpreting routing feasibility and operational resource trade-offs in urban cosmetic distribution.

3. Method

This section describes the methodology used in this study. The following subsection outlines the research design.

3.1 Research Design

This study used a computational routing experiment to evaluate urban cosmetic distribution under time-window, capacity, travel-time risk, and working-shift constraints. The research design compared three routing configurations: deterministic single-capacity routing, distance-based multi-capacity routing, and risk-aware multi-capacity routing. The comparison was intended to examine how routing performance changes when volume feasibility, split delivery, and shift-feasible vehicle utilization are considered.

The research procedure consisted of six stages: data acquisition, data preprocessing, split-delivery preprocessing, distance and travel-time estimation, route construction, and performance evaluation. The overall research framework is shown in Figure 1. The methodological design was supported by previous studies on vehicle routing with time windows, soft time windows, loading constraints, split delivery, distance approximation, and heuristic routing [1], [25].

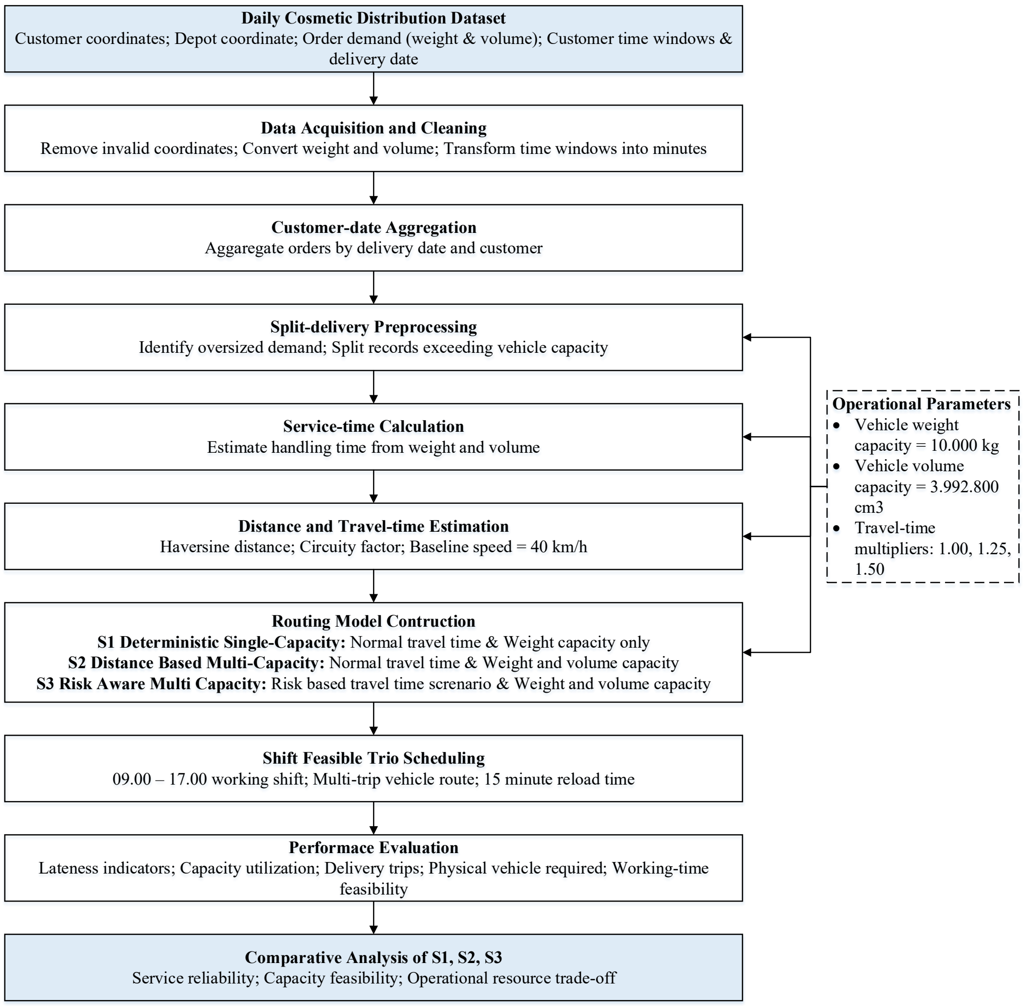

Figure 1. Research framework of the proposed risk-aware multi-capacity routing model

This study does not aim to develop a new optimization algorithm. Instead, it provides a reproducible routing evaluation framework that integrates distance-based travel-time scenarios, split-delivery preprocessing, two-dimensional capacity feasibility, lateness-risk-based scenario escalation, and shift-feasible physical vehicle utilization.

3.2 Data Acquisition

The dataset used in this study represents daily cosmetic product delivery operations. The data include customer identity, customer coordinates, delivery date, total product demand, total weight, total volume, customer time window, and depot coordinate. The depot is treated as the starting and ending point of each delivery trip. Each delivery date is treated as one routing instance. The main input variables and operational parameters are presented in Table 1 and 2.

Table 1. Main variables and operational parameters

Variable / Parameter	Description	Value / Source
Customer coordinate	Latitude and longitude of customer location	Dataset
Depot coordinate	Latitude and longitude of depot	Dataset
Delivery date	Daily routing instance	Dataset
Weight demand, $(\mathrm{w_i})$	Total product weight per customer	Aggregated order data
Volume demand, $(\mathrm{u_i})$	Total product volume per customer	Aggregated order data
Time window, $(\mathrm{e_i}, \mathrm{l_i})$	Earliest and latest service time	Dataset
Service time, (s_i)	Customer handling time	Calculated from demand

Table 2. Main variables and operational parameters

Variable / Parameter	Description	Value / Source
Weight capacity, (W)	Maximum vehicle weight load	10,000 kg
Volume capacity, (U)	Maximum vehicle volume load	3,992,800 cm^3
Baseline speed, (S)	Assumed average delivery speed	40 km/h
Circuity factor, (ρ)	Adjustment from geographical distance to operational distance	1.35
Working shift	Daily operational period	09:00–17:00
Reload time	Depot handling time between trips	15 minutes
Travel-time scenarios	Normal, moderate-risk, high-risk	1.00, 1.25, 1.50

3.3 Operational Parameter Justification

The operational parameters used in this study were treated as controlled scenario assumptions for routing evaluation rather than direct traffic prediction parameters. The circuity factor of 1.35 was used to adjust straight-line Haversine distance into an approximate operational travel distance because actual urban travel paths are generally longer than geographical distance. The baseline speed of 40 km/h was used as a controlled average-speed assumption for urban delivery evaluation. The travel-time multipliers of 1.00, 1.25, and 1.50 represent normal, moderate-risk, and high-risk travel-time scenarios, respectively. These multipliers were used to assess routing vulnerability under progressively more conservative travel-time conditions. The reload time of 15 minutes represents depot handling time required before a physical vehicle can start another delivery trip.

The lateness-risk thresholds were defined as operational risk categories for scenario escalation. A lateness rate below 10% was categorized as normal because only a small proportion of served customers experienced late arrival. A lateness rate between 10% and 30% was categorized as moderate-risk because lateness affected a noticeable portion of daily deliveries and required a more conservative travel-time scenario. A lateness rate of 30% or higher was categorized as high-risk because lateness affected a substantial share of the delivery operation. These thresholds were used as decision rules for risk escalation, not as universal traffic-risk standards.

The service-time equation was defined as an operational handling assumption consisting of a fixed handling component and demand-dependent handling components. The fixed term represents minimum customer handling activity, while the weight and volume coefficients represent additional handling effort associated with larger delivery quantities. Since these coefficients were not estimated from direct time-motion observation, this assumption is acknowledged as a limitation of the current study and should be calibrated using observed service-time data in future research.

3.4 Data Preprocessing

The preprocessing stage was conducted to prepare the delivery data for daily routing evaluation. First, records with missing customer coordinates were removed because distance calculation requires valid latitude and longitude values. Second, weight and volume values were converted into numerical format. Third, time-window values were transformed into minutes from midnight to support arrival-time, waiting-time, and lateness calculations.

Order records were then aggregated by delivery date and customer. This aggregation was required because one customer may have more than one order record on the same delivery date. After aggregation, weight and volume were summed for each customer-date pair. The customer location was retained from the coordinate record, while the customer time window was represented using the available start and end times.

3.5 Notation

Let (C) denote the set of customers and node (0) denote the depot. The symbols (i) and (j) represent origin and destination nodes. Each customer $(i \in C)$ has weight demand (w_i), volume demand (u_i), earliest service time (e_i), latest service time (l_i), and service time (s_i). The vehicle weight capacity is denoted by (W), while the vehicle volume capacity is denoted by (U). The arrival time and departure time at customer (i) are denoted by (a_i) and (d_i), respectively. The lateness at customer (i) is denoted by (L_i), while the waiting time is denoted by (WT_i). The travel-time risk scenario is denoted by $(r \in \{normal, moderate, high\})$.

3.6 Split-Delivery Preprocessing

After customer-date aggregation, some records may exceed single-vehicle capacity. Such records cannot be served by one vehicle trip without split delivery or a larger vehicle. Therefore, this study applies split-delivery preprocessing to oversized customer-date records before route construction. The number of split units for customer (i) is calculated using equation (1).

$$m_i = \max\left(\left\lceil\left(\frac{w_i}{W}\right)\right\rceil, \left\lceil\left(\frac{u_i}{U}\right)\right\rceil\right) \tag{1}$$

Where (m_i) is the number of split delivery units, (w_i) is the aggregated weight demand, (u_i) is the aggregated volume demand, (W) is vehicle weight capacity, and (U) is vehicle volume capacity. If $(m_i > 1)$, the customer-date demand is divided into (m_i) delivery units. The weight and volume of each split unit are calculated using equation (2) and (3).

$$w_i' = \frac{\{w_i\}}{\{m_i\}} \tag{2}$$

$$u_i' = \frac{\{u_i\}}{\{m_i\}} \tag{3}$$

Where (w_i') and (u_i') denote the weight and volume of each split unit. This preprocessing step ensures that each delivery unit is feasible under single-vehicle capacity before routing is performed. Split delivery is applied only to customer-date records whose demand exceeds vehicle capacity.

3.7 Distance and Travel-Time Estimation

This study uses a coordinate-based distance approximation because direct road-network and historical traffic data are not used. The geographical distance between two nodes is calculated using the Haversine formula [19]. Since straight-line distance does not fully represent operational travel distance, the geographical distance is adjusted using a circuity factor [20], [21]. The estimated operational distance was calculated using equation (4).

$$D_{\{ij\}} = \rho d_{\{ij\}}^{\{geo\}} \tag{4}$$

Where $(D_{\{ij\}})$ denotes the estimated operational distance from node (i) to node (j), $\left(d_{\{ij\}}^{\{geo\}}\right)$ denotes the geographical distance, and (ρ) denotes the circuity factor. In this study, $(\rho = 1.35)$. The baseline travel time is calculated using equation (5).

$$T_{\{ij\}}^{\{base\}} = \frac{\{D_{\{ij\}}\}}{\{S\}} \times 60 \tag{5}$$

Where $T_{\{ij\}}^{\{base\}}$ is the baseline travel time in minutes and (S) is the baseline speed in km/h. In this study, (S = 40) km/h. To represent travel-time risk, three scenario multipliers are calculated using equation (6)

$$T_{\{ij\}}^{\{r\}} = \gamma_r T_{\{ij\}}^{\{base\}} \tag{6}$$

Where $T_{\{ij\}}^{\{r\}}$ is the travel time under risk scenario (r), and (γ_r) is the scenario multiplier. The multiplier is defined using equation (7)

$$\gamma_r = \begin{matrix} 1.00, & r = normal \\ 1.25, & r = moderate \\ 1.50, & r = high \end{matrix} \tag{7}$$

These scenarios are used as sensitivity-based travel-time levels for evaluating lateness vulnerability. They are not intended to represent historical traffic percentiles or actual traffic prediction.

3.8 Service Time, Arrival Time, and Lateness

Service time at customer (i) is calculated using equation (8)

$$s_i = 8 + 0.05w_i + 0.00001u_i \tag{8}$$

where (s_i) is the service time in minutes, (w_i) is the weight demand, and (u_i) is the volume demand of customer (i). The equation represents fixed handling time and demand-based handling time. If a vehicle travels from node (i) to node (j), the arrival time at node (j) is calculated using equation (9)

$$a_j = d_i + T_{\{ij\}}^{\{r\}} \tag{9}$$

The departure time from customer (i) is calculated using equation (10)

$$d_i = max(a_i, e_i) + s_i \tag{10}$$

where $\left(max(a_i, e_i)\right)$ indicates that the vehicle waits if it arrives before the earliest service time.
Waiting time is calculated using equation (11)

$$WT_i = max(0, e_i - a_i) \tag{11}$$

Lateness is calculated using equation (12)

$$L_i = max(0, a_i - l_i) \tag{12}$$

Total lateness is calculated using equation (13)

$$TL = \sum_{\{i \in C\}} L_i \tag{13}$$

The lateness rate is calculated using equation (14)

$$LR = \frac{\{N_{\{late\}}\}}{\{N_{\{served\}}\}} x100\% \tag{14}$$

Where (LR) is the lateness rate and $\frac{\{N_{\{late\}}\}}{\{N_{\{served\}}\}}$ is the number of late customers.

3.9 Capacity Feasibility

For each delivery trip (p), let (C_p) denote the set of customers or delivery units served in that trip. Weight utilization is calculated using equation (15)

$$WU_p = \left\{\frac{\Sigma\{i \in C_p\} w_i}{W}\right\} x100\% \tag{15}$$

Volume utilization is calculated using equation (16)

$$VU_p = \left\{\frac{\Sigma\{i \in C_p\} u_i}{U}\right\} x100\% \tag{16}$$

A delivery trip is considered capacity-feasible if both conditions are calculated using equation 17 and 18.

$$WU_p \leq 100\% \tag{17}$$

$$VU_p \leq 100\% \tag{18}$$

In the deterministic single-capacity baseline, only weight capacity is enforced during route construction, while volume utilization is evaluated after routing. In the multi-capacity models, both weight and volume constraints are enforced during route construction.

3.10 Routing Configurations

This study evaluates three routing configurations to compare the effect of capacity representation and travel-time risk on delivery performance. The first configuration is the deterministic single-capacity baseline, denoted as S1. This configuration uses the normal travel-time scenario and enforces only vehicle weight capacity during route construction. Volume utilization is not used as a hard constraint in S1, but it is evaluated after routing to identify whether a route that appears feasible based on weight may still violate volume capacity.

The second configuration is distance-based multi-capacity routing, denoted as S2. This configuration also uses the normal travel-time scenario, but it enforces both weight and volume capacity during route construction. A customer or delivery unit can only be inserted into a delivery trip if the remaining vehicle capacity is feasible in both dimensions. Therefore, S2 is used to evaluate whether simultaneous weight-volume constraints can reduce hidden capacity violations compared with the single-capacity baseline.

The third configuration is risk-aware multi-capacity routing, denoted as S3. This configuration extends S2 by incorporating lateness-risk-based travel-time scenario escalation. The model first evaluates the initial lateness rate under the normal travel-time scenario. Based on this lateness rate, the travel-time scenario is then selected as normal, moderate-risk, or high-risk. The selected scenario is used to construct the final risk-aware routing result. This configuration is designed to evaluate how delivery performance changes when routing decisions are assessed under more conservative travel-time conditions. The risk scenario is selected according to equation (19).

$$r(LR) = \begin{cases} normal, & LR < 10\% \\ moderate, & 10\% \leq LR < 30\% \\ high, & LR \geq 30\% \end{cases} \tag{19}$$

Where (LR) is the initial lateness rate. The lateness-risk thresholds were defined as operational risk categories for scenario escalation. A lateness rate below 10% was categorized as normal because only a small proportion of served customers experienced late arrival. A lateness rate between 10% and 30% was categorized as moderate-risk because lateness affected a noticeable portion of daily deliveries. A lateness rate of 30% or higher was categorized as high-risk because lateness affected a substantial share of the delivery operation. The lateness-risk-based scenario escalation mechanism is illustrated in Figure 2.

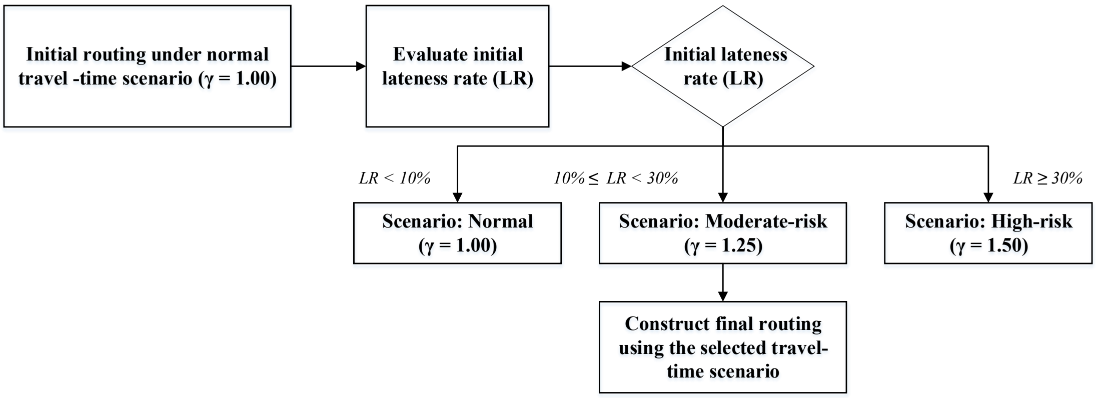

Figure 2. Lateness-risk-based travel-time scenario escalation mechanism

3.11 Route Construction Heuristic

Route construction is performed using a Parallel Nearest Neighbor with Soft Time Windows heuristic. The heuristic was selected because it provides a transparent and reproducible procedure for applied vehicle routing studies [22], [23], [24], [25]. The next customer is selected using a composite score that considers travel time, predicted lateness, waiting time, and due-time urgency. For candidate customer (j), the score is calculated using equation (20).

$$Score_j = T_{\{ij\}}^{\{r\}} + \omega_1 L_j + \omega_2 WT_j + \omega_3 DU_j \quad (20)$$

Where $T_{\{ij\}}^{\{r\}}$ is the travel time from current node (i) to candidate customer (j), (L_j) is predicted lateness, (WT_j) is waiting time, and (DU_j) is due-time urgency. The coefficients (ω_1), (ω_2), and (ω_3) determine the relative contribution of lateness, waiting time, and urgency. In this study, the score weights were set as $\omega_1 = 10.00$, $\omega_2 = 0.15$, and $\omega_3 = 0.05$. These values indicate that predicted lateness was prioritized over waiting time and due-time urgency during customer selection. The customer with the lowest score is selected as the next visit if capacity and shift feasibility remain satisfied.

3.12 Shift-Feasible Physical Vehicle Utilization

This study distinguishes delivery trips from physical vehicle requirements. A delivery trip represents one depot-customer-depot cycle, while a physical vehicle may perform more than one trip within the same working shift. The working shift is defined from 09:00 to 17:00.

A trip can be assigned to a physical vehicle if the vehicle can depart from the depot, serve customers, return to the depot, and remain within the end of the working shift. The shift feasibility condition is calculated using equation (21).

$$t_p^{\{end\}} \leq H^{\{end\}} \quad (21)$$

where $t_p^{\{end\}}$ is the return time of delivery trip (p), and $H^{\{end\}}$ is the end of the working shift. After a vehicle returns to the depot, it can be reused for another trip if the following condition is calculated using equation (22).

$$t_p^{\{end\}} + t^{\{reload\}} \leq H^{\{end\}} \quad (22)$$

where $t^{\{reload\}}$ is the reload time at the depot. In this study, $t^{\{reload\}}$=15) minutes. If an existing vehicle cannot perform the next trip within the shift, a new physical vehicle is opened. Therefore, the number of delivery trips and the number of physical vehicles are reported separately.

3.13 Algorithm Procedure

The complete procedure of the proposed framework is summarized in Algorithm 1. The process begins with data acquisition and customer-date aggregation, followed by split-delivery preprocessing for oversized demand. Distance and travel time are then estimated using Haversine distance, a circuity factor, and scenario-based travel-time multipliers. Three routing configurations are constructed sequentially: S1 as the deterministic single-capacity baseline, S2 as the distance-based multi-capacity model, and S3 as the risk-aware multi-capacity model using lateness-based scenario escalation. Each generated trip is evaluated to ensure capacity feasibility and the ability to return to the depot within the 09:00–17:00 working shift. Finally, feasible trips are assigned to physical vehicles using multi-trip utilization, and the three configurations are compared using lateness, capacity, travel-time, delivery-trip, and physical-vehicle indicators.

Algorithm 1. Risk-Aware Multi-Capacity Routing with Split Delivery and Shift Feasibility

Input:
- Customer coordinates and depot coordinates
- Customer weight demand and volume demand
- Customer time windows
- Vehicle weight and volume capacity
- Working shift and reload time
- Baseline speed and circuity factor
- Travel-time scenario multipliers
- Lateness-risk thresholds

Output:
- Delivery trips
- Physical vehicles required
- Lateness indicators
- Capacity utilization
- Working-time feasibility status

Step 1: Load customer, depot, demand, coordinate, and time-window data.
Step 2: Aggregate order records by delivery date and customer.
Step 3: Convert time windows into minutes from midnight.
Step 4: Identify oversized customer-date records.
Step 5: Apply split-delivery preprocessing when weight or volume exceeds vehicle capacity.
Step 6: Calculate service time for each delivery unit.
Step 7: Estimate geographical distance using the Haversine formula.
Step 8: Adjust geographical distance using the circuity factor.
Step 9: Convert operational distance into baseline travel time.
Step 10: Construct S1 using normal travel time and weight-only capacity.
Step 11: Construct S2 using normal travel time and weight-volume capacity.
Step 12: Calculate the initial lateness rate under S2.
Step 13: Select risk scenario using the lateness-rate threshold rule.
Step 14: Construct S3 using the selected travel-time scenario.
Step 15: Ensure each trip can return to the depot before the end of the working shift.
Step 16: Assign feasible trips to physical vehicles using multi-trip utilization.
Step 17: Calculate lateness, travel time, distance, capacity utilization, delivery trips, and physical vehicles required.
Step 18: Compare S1, S2, and S3 using aggregate performance indicators.

3.14 Model Testing and Evaluation Metrics

The models were tested using all daily delivery instances in the dataset. Each delivery date was treated as one routing instance. The results were reported using mean and standard deviation across delivery dates, with mean and SD presented in separate columns.

The evaluation metrics include served customers, unserved customers, infeasible customers, late customers, lateness rate, total lateness, maximum lateness, total distance, total travel time, delivery trips, physical vehicles required, average trips per vehicle, vehicle workload, weight utilization, volume utilization, over-capacity cases, and working-time violation cases.

Relative changes from the deterministic single-capacity baseline were also calculated to evaluate how the multi-capacity and risk-aware configurations differ from the simplified baseline. The analysis focuses on three aspects: service reliability, capacity feasibility, and operational resource trade-off.

3.15 Statistical Test and Sensitivity Analysis

Statistical testing was conducted to evaluate whether the observed differences among S1, S2, and S3 were statistically meaningful. Since the three routing configurations were evaluated on the same 63 delivery dates, paired statistical testing was applied. The Shapiro-Wilk test was first used to examine the normality of paired differences. Because several paired differences were not normally distributed, the Wilcoxon signed-rank test was used as the main non-parametric test.

Sensitivity analysis was also conducted by rerunning the routing framework under alternative operational assumptions. The tested parameters included baseline speed, circuity factor, reload time, lateness-risk threshold, travel-time multiplier, and service-time scale. This analysis was conducted to examine whether the main conclusions remained stable under different operational settings.

4. Results and Discussion

The results of this analysis are presented and discussed in the following section.

4.1 Split-Delivery Preprocessing and Routing Feasibility

The routing evaluation was conducted using 63 daily delivery instances. Before route construction, customer-date records whose aggregated demand exceeded single-vehicle capacity were processed using split-delivery preprocessing. As shown in Table 3, six oversized customer-date records were identified and transformed into seven additional delivery units. The maximum split count was three, indicating that the largest oversized demand required three feasible delivery units.

Table 3. Split-delivery preprocessing summary

Indicator	Value
Oversized records split	6
Additional delivery units	7
Maximum split count	3

After split-delivery preprocessing, all routing configurations produced zero unserved customers and zero infeasible customers. In addition, all delivery trips satisfied the 09:00–17:00 working-shift constraint, as indicated by zero working-time violation cases. This confirms that the split-delivery and shift-feasibility mechanisms successfully transformed the delivery dataset into feasible routing instances before the three routing configurations were compared.

4.2 Aggregate Operational Performance

Table 4 presents the aggregate operational performance of the three routing configurations. The values are reported as mean and standard deviation (SD) across all delivery dates. The deterministic single-capacity baseline (S1), the distance-based multi-capacity routing model (S2), and the risk-aware multi-capacity routing model (S3) served the same average number of customers or delivery units, namely 62.44 per day. This indicates that the comparison was conducted under equal service coverage.

Table 4. Aggregate operational performance

Indicator	S1		S2		S3	
	Mean	SD	Mean	SD	Mean	SD
Served customers	62.44	19.68	62.44	19.68	62.44	19.68
Unserved customers	0.00	0.00	0.00	0.00	0.00	0.00
Infeasible customers	0.00	0.00	0.00	0.00	0.00	0.00
Late customers	8.25	4.35	8.22	3.97	7.13	3.79
Lateness rate (%)	12.88	5.13	12.89	4.67	11.05	4.83
Total lateness (min)	671.54	424.56	689.96	436.95	636.85	424.70
Maximum lateness (min)	206.03	115.59	210.33	114.33	227.84	121.21

Total distance (km)	420.95	130.96	424.06	133.28	429.02	133.86
Total travel time (min)	631.43	196.44	636.10	199.92	765.79	259.82
Delivery trips	3.05	0.96	3.10	1.03	3.40	1.11
Physical vehicles required	3.05	0.96	3.06	0.97	3.37	1.07
Working-time violation cases	0.00	0.00	0.00	0.00	0.00	0.00

The S2 model produced lateness performance that was close to the S1 baseline. The lateness rate changed only from 12.88% in S1 to 12.89% in S2, while total lateness increased slightly from 671.54 minutes to 689.96 minutes. Therefore, the main contribution of S2 is not lateness reduction, but capacity-feasibility correction. By enforcing both weight and volume capacity, S2 prevents hidden volume overload that cannot be detected by a weight-only baseline.

The S3 model produced the lowest average number of late customers and the lowest lateness rate. The mean number of late customers decreased from 8.25 in S1 to 7.13 in S3, while the lateness rate decreased from 12.88% to 11.05%. Total lateness also decreased from 671.54 minutes to 636.85 minutes. However, this improvement was achieved with additional operational resources. S3 required an average of 3.40 delivery trips and 3.37 physical vehicles, compared with 3.05 delivery trips and 3.05 physical vehicles in S1. This indicates that risk-aware routing improves service reliability but requires higher resource allocation.

As illustrated in Figure 3. the difference between delivery trips and physical vehicles becomes visible after shift-feasible vehicle utilization is applied. S3 required more delivery trips and physical vehicles than S1 and S2, indicating that lateness reduction was achieved through additional operational resource allocation. This confirms that delivery trips should not be interpreted directly as physical vehicles, because one vehicle may perform multiple trips within the working shift if sufficient time remains.

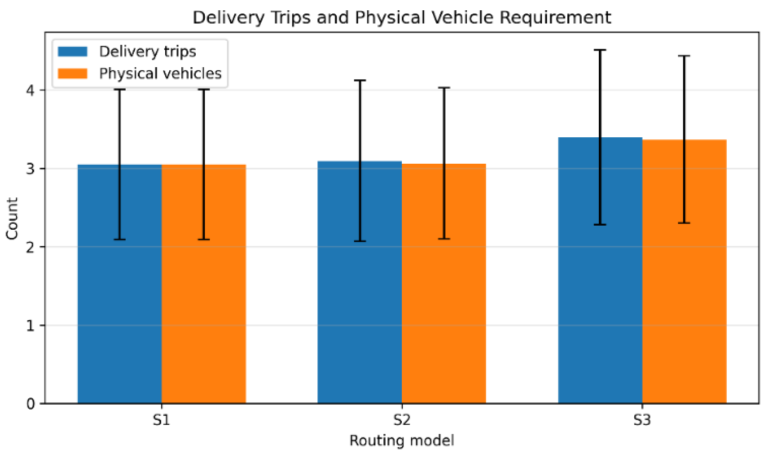

Figure 3. Delivery trips and physical vehicle requirements across routing models

4.3 Relative Change from the Baseline

Table 5 presents the relative performance changes of S2 and S3 compared with the deterministic single-capacity baseline. S2 reduced late customers by only 0.38%, while its lateness rate increased slightly by 0.09% and total lateness increased by 2.74%. These changes are relatively small, indicating that enforcing multi-capacity constraints does not substantially worsen punctuality. However, S2 increased delivery trips by 1.56% and physical vehicle requirements by 0.52%, showing that a minor operational adjustment was needed to achieve loading feasibility.

Table 5. Relative change from s1 baseline

Indicator	S2 vs S1 (%)	S3 vs S1 (%)
Late customers	-0.38	-13.65
Lateness rate	0.09	-14.17
Total lateness	2.74	-5.17

Maximum lateness	2.09	10.58
Total distance	0.74	1.92
Total travel time	0.74	21.28
Delivery trips	1.56	11.46
Physical vehicles required	0.52	10.42

Compared with S1, S3 reduced late customers by 13.65%, lateness rate by 14.17%, and total lateness by 5.17%. However, the statistical test later shows that the reduction in total lateness was descriptive and not statistically significant. These results show that risk-aware routing can reduce overall lateness exposure. However, S3 increased total travel time by 21.28%, delivery trips by 11.46%, and physical vehicles by 10.42%. Therefore, S3 should not be interpreted as the most resource-efficient configuration. Instead, it represents a service-reliability-oriented configuration that reduces lateness at the cost of additional operational resources.

An important observation is that S3 increased maximum lateness by 10.58%, although it reduced late customers, lateness rate, and total lateness. This indicates that risk-aware routing reduced lateness more broadly across delivery instances, but some individual trips still experienced higher extreme lateness. This finding suggests that future improvements should include stronger control of maximum lateness, such as adding a hard lateness limit or a stronger penalty for highly delayed customers.

Figure 4 further shows the difference between weight and volume utilization across routing configurations. Although weight utilization remained low in all models, volume utilization was substantially higher and became the critical capacity dimension. This explains why the single-capacity baseline produced over-volume cases even though it did not violate weight capacity. The figure supports the need for multi-capacity routing in cosmetic distribution, where vehicle space can become more restrictive than vehicle weight capacity.

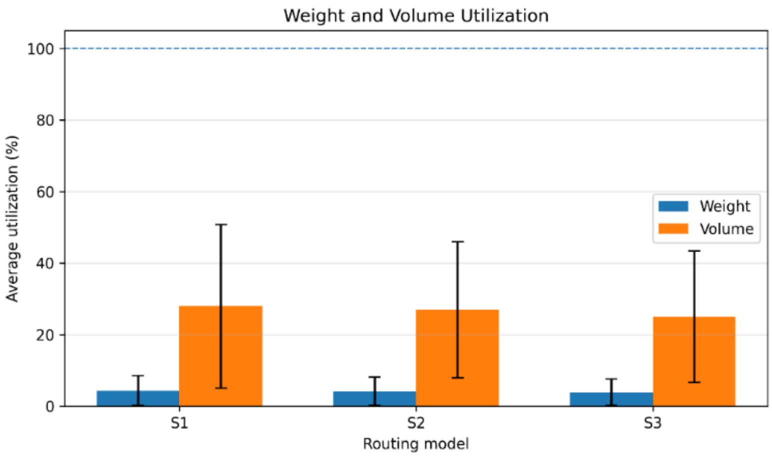

Figure 4. Weight and volume utilization across routing models

4.4 Statistical Test of Model Differences

To verify whether the observed differences among routing configurations were statistically meaningful, paired Wilcoxon signed-rank tests were conducted across the 63 daily delivery instances. The test was selected because the three routing configurations were evaluated on the same delivery dates, and several paired differences did not follow normal distribution. The results are presented in Table 6.

Table 6. Paired Wilcoxon signed-rank test summary

Metric	Comparison	Mean Difference	Wilcoxon p-value	Interpretation
Lateness rate (%)	S2 vs S1	0.012	0.889	Not significant
Lateness rate (%)	S3 vs S1	-1.825	0.005	Significant
Total lateness (min)	S2 vs S1	18.416	0.314	Not significant
Total lateness (min)	S3 vs S1	-34.694	0.351	Not significant

Late customers	S2 vs S1	-0.032	0.778	Not significant
Late customers	S3 vs S1	-1.127	0.008	Significant
Delivery trips	S2 vs S1	0.048	0.180	Not significant
Delivery trips	S3 vs S1	0.349	<0.001	Significant
Physical vehicles required	S2 vs S1	0.016	0.317	Not significant
Physical vehicles required	S3 vs S1	0.317	<0.001	Significant
Total travel time (min)	S2 vs S1	4.670	0.139	Not significant
Total travel time (min)	S3 vs S1	134.360	<0.001	Significant
Maximum lateness (min)	S2 vs S1	4.300	0.123	Not significant
Maximum lateness (min)	S3 vs S1	21.805	0.157	Not significant

The statistical results show that S2 did not significantly change lateness, travel time, delivery trips, or physical vehicle requirements compared with S1. This confirms that S2 mainly functions as a capacity-feasibility correction model. In contrast, S3 significantly reduced lateness rate and late customers compared with S1, but it also significantly increased total travel time, delivery trips, and physical vehicle requirements. Therefore, S3 should be interpreted as a service-reliability-oriented configuration rather than a resource-minimization configuration. The reduction in total lateness was not statistically significant, indicating that the improvement of S3 is more evident in reducing lateness frequency than reducing accumulated lateness duration.

4.5 Capacity and Shift Feasibility

Table 7 shows the capacity and shift-feasibility results. The deterministic single-capacity baseline did not violate weight capacity, but it produced nine over-volume cases. Its maximum volume utilization reached 242.15%, meaning that several routes that appeared feasible by weight were actually infeasible by volume. This finding confirms that weight-only routing can hide substantial loading infeasibility in cosmetic distribution.

Table 7. Capacity and shift feasibility

Scenario	Weight Utilization Mean (%)	Weight Utilization SD (%)	Max Weight Utilization (%)	Volume Utilization Mean (%)	Volume Utilization SD (%)	Max Volume Utilization (%)	Over-Weight Cases	Over-Volume Cases	Working-Time Violation Cases
S1	4.34	4.16	35.57	27.95	22.85	242.15	0	9	0
S2	4.21	3.89	32.81	26.94	19.05	100.00	0	0	0
S3	3.90	3.70	32.75	24.97	18.37	99.94	0	0	0

The S2 and S3 models eliminated all capacity violations. Their maximum volume utilization values were 100.00% and 99.94%, respectively. This confirms that enforcing volume capacity during route construction successfully prevented hidden loading infeasibility. These results support the argument that cosmetic distribution requires multi-capacity routing because product volume can become a binding constraint even when weight capacity is still available.

All configurations produced zero working-time violations. This confirms that the shift-feasible routing mechanism successfully ensured that every delivery trip could return to the depot before the end of the 09:00–17:00 working shift. This is important because delivery trips and physical vehicles are not equivalent. By distinguishing delivery trips from physical vehicle requirements, the framework provides a more realistic interpretation of operational resource needs.

The risk distribution is visualized in Figure 5. Most delivery dates were categorized as moderate-risk, while no delivery date was categorized as high-risk. This indicates that the shift-feasible routing mechanism reduced extreme lateness exposure, but the delivery system still remained sensitive to increased travel-time assumptions. Therefore, deterministic travel-time evaluation alone may underestimate lateness vulnerability in daily delivery planning.

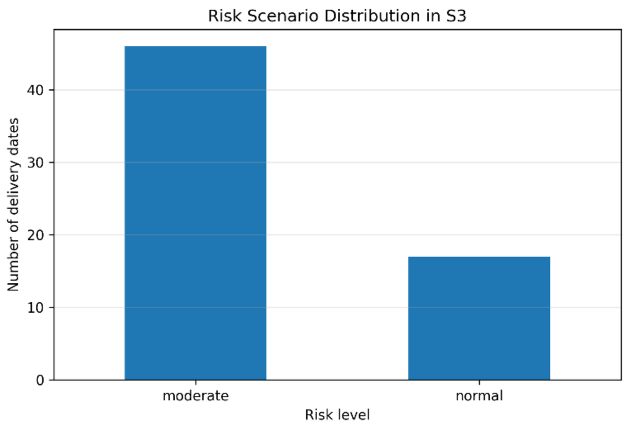

Figure 5. Risk scenario distribution in the risk-aware routing model

4.6 Risk Scenario Distribution

The risk-aware configuration selected the travel-time scenario based on the initial lateness rate. As shown in Table 8, 46 delivery dates were classified as moderate-risk, representing 73.02% of all evaluated dates. The remaining 17 delivery dates were classified as normal-risk, representing 26.98%. No delivery date was classified as high-risk.

Table 8. Risk scenario distribution in S3

Selected Risk Level	Interpretation	Number of Delivery Instances	Percentage (%)
Normal	Low lateness risk	17	26.98
Moderate	Moderate lateness risk	46	73.02
High	High lateness risk	0	0.00

The dominance of moderate-risk delivery dates indicates that the delivery system is not dominated by extreme lateness risk, but most delivery days remain sensitive to travel-time increases. The absence of high-risk instances suggests that the combination of split delivery and shift-feasible routing makes the delivery system more controlled. However, the large proportion of moderate-risk days still indicates that deterministic travel-time assumptions may be too optimistic for daily urban delivery planning.

The operational trade-off is summarized in Figure 6. S3 achieved a lower lateness rate than S1 and S2, but it required higher travel time and more physical vehicles. This confirms that risk-aware routing should be interpreted as a service-reliability-oriented configuration rather than the most resource-efficient option. In contrast, S2 provides a more balanced configuration when the primary objective is to eliminate capacity violations without substantially increasing operational resources.

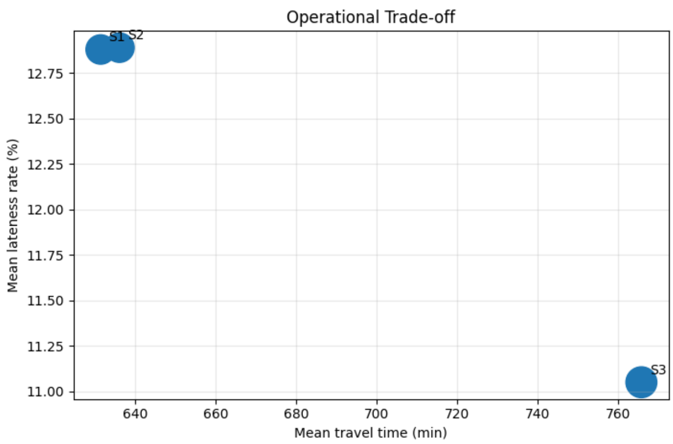

Figure 6. Operational trade-off between travel time, lateness rate, and physical vehicle requirement

4.7 Sensitivity Analysis

Sensitivity analysis was conducted by rerunning the routing framework under alternative baseline speeds, circuity factors, reload times, risk thresholds, travel-time multipliers, and service-time scales. The objective was to examine whether the main conclusions remained stable under different operational assumptions. The results are summarized in Table 9.

Table 9. Sensitivity analysis summary

Scenario group	Scenario values	S3 lateness rate change vs S1 (%)	S3 physical vehicle change vs S1 (%)	S2/S3 over-volume cases	Working-time violations
Speed	30, 40, 50 km/h	-14.173 to -9.643	10.405 to 12.273	0	0
Circuity factor	1.20, 1.35, 1.50	-14.173 to -8.246	7.882 to 11.050	0	0
Reload time	15, 30, 45 min	-14.321 to -14.173	10.417 to 10.938	0	0
Risk threshold	5–20, 10–30, 15–35	-14.173 to -7.106	4.688 to 13.021	0	0
Travel-time multiplier	1.00–1.15–1.30, 1.00–1.25–1.50, 1.00–1.35–1.70	-14.173 to -6.501	8.333 to 13.542	0	0
Service-time scale	0.75, 1.00, 1.25	-14.173 to -9.888	10.138 to 11.176	0	0

The sensitivity results show that the main conclusions are robust across alternative parameter settings. In all tested scenarios, S2 and S3 eliminated over-volume violations, and no working-time violations occurred. S3 also consistently reduced the lateness rate compared with S1, although the magnitude of improvement varied across assumptions. The largest lateness-rate reduction occurred under the baseline setting and reload-time scenarios, while the smallest reduction occurred under the lower travel-time multiplier setting. However, S3 consistently required more physical vehicles than S1, confirming the trade-off between service reliability and operational resource usage.

4.8 Operational Trade-Off Discussion

The final results provide three main operational insights. First, split-delivery preprocessing is necessary to make oversized customer-date demand feasible before route construction. Without this step, several customer-date records whose volume exceeded single-vehicle capacity could not be realistically served in one trip. Therefore, split delivery improves the validity of the routing evaluation rather than artificially improving the result.

Second, multi-capacity routing is necessary for cosmetic distribution. The S1 baseline showed no weight-capacity violation, but it produced nine over-volume cases. This means that using weight as the only capacity dimension can lead to misleading feasibility conclusions. S2 corrected this issue by enforcing both weight and volume capacity, eliminating all over-volume cases while maintaining lateness performance comparable to S1.

Third, risk-aware routing improves service reliability but introduces resource trade-offs. S3 reduced the average number of late customers and lateness rate compared with S1, while the reduction in total lateness was descriptive and not statistically significant. However, it increased total travel time, delivery trips, and physical vehicle requirements. This trade-off is expected because risk-aware routing applies more conservative travel-time assumptions and constructs routes that reduce lateness exposure. Therefore, S3 is suitable when service reliability is prioritized, while S2 provides a more balanced option when capacity feasibility is the main objective and additional vehicle use must be limited.

Overall, the proposed framework provides a more realistic routing evaluation than deterministic single-capacity routing. It identifies hidden volume overload, handles oversized demand through split delivery, enforces working-shift feasibility, and distinguishes delivery trips from physical vehicles. These capabilities are important for urban cosmetic distribution, where operational feasibility depends not only on distance and weight capacity, but also on volume constraints, service time windows, travel-time risk, and daily vehicle utilization.

The increase in maximum lateness under S3 requires careful interpretation. Although S3 reduced the overall lateness rate and the number of late customers, maximum lateness increased by 10.58% compared with S1. This indicates that the risk-aware model reduced lateness more broadly across daily delivery instances, but did not fully control extreme lateness in a small number of trips. In operational terms, S3 improves general service reliability but may still require an additional maximum-lateness control mechanism. Future model development should therefore include a stronger penalty or hard constraint for extreme lateness, so that reliability improvement does not occur at the expense of highly delayed individual deliveries.

5. Conclusion

This study proposed a risk-aware multi-capacity routing evaluation framework for shift-feasible cosmetic distribution. The framework integrates split-delivery preprocessing, simultaneous weight-volume capacity feasibility, lateness-risk-based travel-time scenarios, and physical vehicle reuse within a 09:00–17:00 working shift. The main contribution of this study is the integration of these operational components into one reproducible framework that can identify hidden capacity violations and evaluate service-reliability trade-offs.

The results show that split-delivery preprocessing handled oversized customer-date demand and eliminated unserved and infeasible customers. The deterministic single-capacity baseline did not violate weight capacity, but it produced over-volume cases, confirming that weight-only routing may hide loading infeasibility in cosmetic distribution. The multi-capacity model eliminated all capacity violations without significantly changing lateness or resource requirements. The risk-aware model significantly reduced lateness rate and the number of late customers, but significantly increased travel time, delivery trips, and physical vehicle requirements. These findings confirm that improving service reliability requires additional operational resources.

Sensitivity analysis further shows that the main conclusions remain stable across alternative speed, circuity factor, reload time, risk-threshold, travel-time multiplier, and service-time assumptions. The practical implication is that distribution planners should not evaluate routing performance only through distance or weight capacity. Volume feasibility, lateness risk, working-shift feasibility, and physical vehicle utilization must also be considered when designing daily delivery operations.

This study is limited by the use of coordinate-based travel-time estimation, fixed operational parameters, and a case-based dataset of 63 daily delivery instances. Future research should incorporate road-network travel time, dynamic traffic data, calibrated service-time observations, larger logistics networks, and more advanced optimization or metaheuristic comparisons.

Data Availability

The data used in this study are available from the corresponding author upon reasonable request. The processed routing outputs, tables, and figures were generated from the final shift-feasible routing evaluation.

References

[1] X. Liu, Y.-L. Chen, L. Y. Por, and C. S. Ku, “A Systematic Literature Review of Vehicle Routing Problems with Time Windows,” *Sustainability*, vol. 15, no. 15, p. 12004, Aug. 2023, doi: 10.3390/su151512004.

[2] M. M. Solomon, “Algorithms for the Vehicle Routing and Scheduling Problems with Time Window Constraints,” *Oper. Res.*, vol. 35, no. 2, pp. 254–265, Apr. 1987, doi: 10.1287/opre.35.2.254.

[3] G. B. Dantzig and J. H. Ramser, “The Truck Dispatching Problem,” *Manage. Sci.*, vol. 6, no. 1, pp. 80–91, Oct. 1959, doi: 10.1287/mnsc.6.1.80.

[4] D. Toth, Paolo; Vigo, *Vehicle Routing*, 2nd ed. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2014. doi: 10.1137/1.9781611973594.

[5] K. Zhang, F. He, Z. Zhang, X. Lin, and M. Li, “Multi-vehicle routing problems with soft time windows: A multi-agent reinforcement learning approach,” *Transp. Res. Part C Emerg. Technol.*, vol. 121, p. 102861, Dec. 2020, doi: 10.1016/j.trc.2020.102861.

[6] T. Yang, W. Wang, and Q. Wu, “Fuzzy Demand Vehicle Routing Problem with Soft Time Windows,” *Sustainability*, vol. 14, no. 9, p. 5658, May 2022, doi: 10.3390/su14095658.

[7] T. Adamo, M. Gendreau, G. Ghiani, and E. Guerriero, “A review of recent advances in time-dependent vehicle routing,” *Eur. J. Oper. Res.*, vol. 319, no. 1, pp. 1–15, Nov. 2024, doi: 10.1016/j.ejor.2024.06.016.

[8] M. Gendreau, G. Ghiani, and E. Guerriero, “Time-dependent routing problems: A review,” *Comput. Oper. Res.*, vol. 64, pp. 189–197, Dec. 2015, doi: 10.1016/j.cor.2015.06.001.

[9] X. Zhang, L. Chen, M. Gendreau, and A. Langevin, “A branch-and-cut algorithm for the vehicle routing problem with two-dimensional loading constraints,” *Eur. J. Oper. Res.*, vol. 302, no. 1, pp. 259–269, Oct. 2022, doi: 10.1016/j.ejor.2021.12.050.

[10] J. T. Soman and R. J. Patil, “Two-dimensional loading in vehicle routing problem with release and due dates,” *Expert Syst. Appl.*, vol. 232, p. 120862, Dec. 2023, doi: 10.1016/j.eswa.2023.120862.

[11] L. P. Fava *et al.*, “A Multi-Start Algorithm for Solving the Capacitated Vehicle Routing Problem with Two-Dimensional Loading Constraints,” *Symmetry (Basel).*, vol. 13, no. 9, p. 1697, Sep. 2021, doi: 10.3390/sym13091697.

[12] X. Zhang, L. Chen, M. Gendreau, and A. Langevin, “A Branch-and-Price-and-Cut Algorithm for the Vehicle Routing Problem with Two-Dimensional Loading Constraints,” *Transp. Sci.*, vol. 56, no. 6, pp. 1618–1635, Nov. 2022, doi: 10.1287/trsc.2022.1135.

[13] K. M. Ferreira, T. A. de Queiroz, and F. M. B. Toledo, “An exact approach for the green vehicle routing problem with two-dimensional loading constraints and split delivery,” *Comput. Oper. Res.*, vol. 136, p. 105452, Dec. 2021, doi: 10.1016/j.cor.2021.105452.

[14] K. M. Ferreira, T. A. de Queiroz, P. Munari, and F. M. B. Toledo, “A variable neighborhood search for the green vehicle routing problem with two-dimensional loading constraints and split delivery,” *Eur. J. Oper. Res.*, vol. 316, no. 2, pp. 597–616, Jul. 2024, doi: 10.1016/j.ejor.2024.01.049.

[15] V. F. Yu, P. Jodiawan, and A. A. N. P. Redi, “Crowd-shipping problem with time windows, transshipment nodes, and delivery options,” *Transp. Res. Part E Logist. Transp. Rev.*, vol. 157, p. 102545, Jan. 2022, doi: 10.1016/j.tre.2021.102545.

[16] M. Rajaei, G. Moslehi, and M. Reisi-Nafchi, “The split heterogeneous vehicle routing problem with three-dimensional loading constraints on a large scale,” *Eur. J. Oper. Res.*, vol. 299, no. 2, pp. 706–721, Jun. 2022, doi: 10.1016/j.ejor.2021.08.025.

[17] M. Gmira, M. Gendreau, A. Lodi, and J.-Y. Potvin, “Tabu search for the time-dependent vehicle routing problem with time windows on a road network,” *Eur. J. Oper. Res.*, vol. 288, no. 1, pp. 129–140, Jan. 2021, doi: 10.1016/j.ejor.2020.05.041.

[18] D. Merchán, M. Winkenbach, and A. Snoeck, “Quantifying the impact of urban road networks on the efficiency of local trips,” *Transp. Res. Part A Policy Pract.*, vol. 135, pp. 38–62, May 2020, doi: 10.1016/j.tra.2020.02.015.

[19] E. Maria, E. Budiman, Haviluddin, and M. Taruk, “Measure distance locating nearest public facilities using Haversine and Euclidean Methods,” *J. Phys. Conf. Ser.*, vol. 1450, no. 1, p. 012080, Feb. 2020, doi: 10.1088/1742-6596/1450/1/012080.

[20] G. Boeing, “The Morphology and Circuity of Walkable and Drivable Street Networks,” 2019, pp. 271–287. doi: 10.1007/978-3-030-12381-9_12.

[21] R. Jazemi, E. Alidadiani, K. Ahn, and J. Jang, “A Review of Literature on Vehicle Routing Problems of Last-Mile Delivery in Urban Areas,” *Appl. Sci.*, vol. 13, no. 24, p. 13015, Dec. 2023, doi: 10.3390/app132413015.

[22] W. Cook, S. Held, and K. Helsgaun, “Constrained Local Search for Last-Mile Routing,” *Transp. Sci.*, vol. 58, no. 1, pp. 12–26, Jan. 2024, doi: 10.1287/trsc.2022.1185.

[23] F. Liu, C. Lu, L. Gui, Q. Zhang, X. Tong, and M. Yuan, “Heuristics for Vehicle Routing Problem: A Survey and Recent Advances,” Mar. 2023, [Online]. Available: http://arxiv.org/abs/2303.04147

[24] T. Vidal, T. G. Crainic, M. Gendreau, and C. Prins, “Heuristics for multi-attribute vehicle routing

problems: A survey and synthesis," *Eur. J. Oper. Res.*, vol. 231, no. 1, pp. 1–21, Nov. 2013, doi: 10.1016/j.ejor.2013.02.053.

[25] B. Serrano, A. M. Florio, S. Minner, M. Schiffer, and T. Vidal, "Contextual Stochastic Vehicle Routing with Time Windows," Feb. 2024, [Online]. Available: http://arxiv.org/abs/2402.06968