# A Traffic-Aware Time-Dependent Vehicle Routing Problem for Middle-Mile Deliveries under Uncertain Travel Time

Pegah Bagheri[1], Muhammad Ahsanul Habib[2], and Hamid Afshari[3]

[1]PhD Student, Industrial Engineering Department, Dalhousie University, Halifax, Canada
[2]Professor, School of Planning, and Department of Civil and Resource Engineering, Dalhousie University, Halifax, Canada
[3]Assistant Professor, Industrial Engineering Department, Dalhousie University, Halifax, Canada

## Short summary

Urban freight distribution faces highly variable traffic conditions, where recurring congestion and unexpected disruptions can undermine delivery reliability and efficiency. This study develops a traffic-aware optimization framework for middle-mile urban logistics through a Stochastic Time-Dependent Vehicle Routing Problem with Time Windows (STDVRPTW). The model integrates time-varying traffic conditions, stochastic travel-time disruptions, and customer service requirements within a unified framework. It is formulated as a two-stage stochastic program in which first-stage routing and scheduling decisions are evaluated across multiple traffic scenarios. To solve the large-scale problem efficiently, we propose an improved Benders decomposition algorithm with problem-specific valid inequalities that exploit the structure of time buckets and vehicle capacity constraints. Numerical results show that the proposed algorithm outperforms Gurobi and standard Benders decomposition. The findings also provide operational insights into how congestion pricing, delivery time-window flexibility, and fleet deployment interact, showing that pricing policies can encourage off-peak deliveries and changes in fleet utilization.

**Keywords:** Middle-mile deliveries, Operations research applications, Time-dependent vehicle routing, Urban freight transport, Urban logistics

## 1 Introduction

Urban freight transportation plays a critical role in economic activity, particularly in sectors that depend on frequent and reliable goods distribution, such as manufacturing, wholesale trade, retail, and warehousing. In Canada, these sectors account for more than a quarter of national GDP (Fan & Heminthavong, 2022). At the same time, freight vehicles contribute disproportionately to urban congestion, delays, and environmental impacts (Álvarez et al., 2024). As cities grow and e-commerce expands, improving the efficiency and reliability of urban freight systems has become an important challenge for both city logistics and urban transport planning.

Among urban freight activities, middle-mile distribution—linking warehouses to retail stores and other points of sale—accounts for a substantial share of daily goods movements. These operations typically involve frequent deliveries, strict receiving schedules, and coordination across dispersed stores (Hübner et al., 2016). Because such deliveries must be synchronized with store operations while moving through congested urban networks, routing decisions are shaped by both service constraints and traffic conditions. In practice, these systems are affected not only by recurrent congestion but also by disruptions such as incidents, construction activities, and temporary road closures (Miranda & Conceição, 2016), which increase delivery costs, reduce reliability, and complicate urban freight planning (Allen et al., 2012).

To support routing decisions in these settings, the operations research literature has extensively studied the Vehicle Routing Problem (VRP) and its many variants (Toth & Vigo, 2014). Two streams are especially relevant. The Vehicle Routing Problem with Time Windows (VRPTW) captures delivery constraints related to store receiving schedules (Duan et al., 2021), while the Time-Dependent Vehicle Routing Problem (TDVRP) models travel times as functions of departure time and therefore captures recurrent congestion explicitly (Hou et al., 2021). These features are combined in the Time-Dependent Vehicle Routing Problem with Time Windows (TDVRPTW), which jointly represents service constraints and time-varying traffic conditions (Spliet et al., 2018).

However, most existing time-dependent routing models remain deterministic and rely on historical average travel times. In reality, urban networks are also affected by stochastic disruptions, including accidents, weather conditions, and road works, that alter delivery performance and service reliability (Flatberg et al., 2007). Although stochastic vehicle routing models have been developed to represent such uncertainty (Zhang et al., 2013), they often abstract from time-dependent congestion and therefore do not fully capture the interaction between predictable traffic patterns and unexpected disruptions.

In addition, most studies combining time-dependent routing with customer time windows still rely on heuristic or metaheuristic methods, while exact approaches remain relatively rare (Wei et al., 2024). Existing exact methods, including branch-and-price (Dabia et al., 2013), dynamic programming (Sun et al., 2018), branch-and-cut (Wei et al., 2024), and branch-and-price-and-cut (Rothenbächer, 2019), are largely developed for deterministic, route-based formulations with predefinable routing structures, typically in single-tour settings. As a result, they are not naturally suited to stochastic traffic-aware VRPs, where uncertainty must be integrated directly into routing and scheduling decisions and where vehicles may perform multiple tours.

To address these limitations, this study develops a stochastic time-dependent vehicle routing problem with time windows (STDVRPTW) for middle-mile urban freight that integrates recurrent congestion, disruption-driven travel-time uncertainty, and store receiving windows within a unified two-stage stochastic optimization framework. The proposed model links routing and scheduling decisions to key urban freight policy levers, particularly congestion pricing and delivery-window flexibility, thereby enabling evaluation of both operational responses and policy effects. To solve the resulting large-scale stochastic program, the paper develops an improved Benders decomposition algorithm as an exact solution approach that enhances computational tractability and supports traffic-aware scenario analysis. The study therefore makes three contributions: it develops an integrated stochastic traffic-aware routing formulation, proposes an exact decomposition-based solution approach, and generates managerial and policy insights into how congestion pricing and delivery-window flexibility jointly influence route design, fleet deployment, and peak-period freight exposure.

mdf - start

## 2 Methodology

***Study Setting***

This study addresses an urban middle-mile delivery system where a distribution center serves retail stores under demand and receiving time-window constraints. As illustrated in Figure 1, the routing environment integrates two traffic components: recurrent time-dependent congestion, which varies systematically throughout the day, and stochastic disruptions (e.g., incidents or work zones) that introduce additional uncertainty. Our framework accounts for both to ensure operationally robust

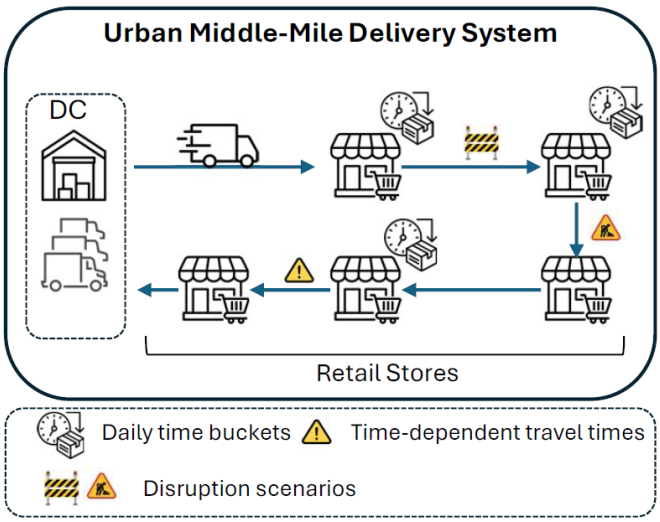



Figure 1: Study setting for the urban middle-mile delivery system, including customer time windows, time-dependent traffic patterns, and disruption scenarios

urban freight planning.

***Modeling Approach***

The problem is first formulated as a deterministic time-dependent vehicle routing problem with time windows (TDVRPTW).

The core mathematical model in compact form is as follows

$$\min_{x,y} \quad C^{\text{travel}}(x) + C^{\text{traffic}}(x) + C^{\text{setup}}(x) + C^{\text{TW-violation}}(y) \tag{1}$$

$$\text{s.t.} \quad Ax = b \tag{2}$$

$$Tx + Wy \leq r \tag{3}$$

$$x \in X,\ y \geq 0. \tag{4}$$

The objective function minimizes total cost consisting of travel cost, congestion-related traffic penalty cost, vehicle setup cost, and time-window-violation cost. The routing-design vector $x$ aggregates the binary decisions for arc traversal ($x^{\text{arc}}_{ijmt}$), node-bucket synchronization ($x^{\text{node}}_{imt}$), and vehicle activation ($x^{\text{veh}}_{m}$). The constraint set $Ax = b$ represents the following routing-flow connectivity constraints, as well as capacity-budget synchronization constraints, to ensure customer coverage, depot balance, and fleet requirements.

*Routing and Flow Connectivity:*

$$\sum_{t,m,i} x^{\text{arc}}_{ijmt} = \sum_{t,m,j} x^{\text{arc}}_{ijmt} = 1, \qquad \forall j \in V \setminus \{0\} \tag{5}$$

$$\sum_{t,j} x^{\text{arc}}_{0jmt} = \sum_{t,i} x^{\text{arc}}_{i0mt} = x^{\text{veh}}_{m}, \qquad \forall m \in M \tag{6}$$

*Capacity and Bucket Synchronization:*

$$\sum_{j} x^{\text{arc}}_{ijmt} = x^{\text{node}}_{imt}, \quad \sum_{t,i} D_i x^{\text{node}}_{imt} \leq C_m, \qquad \forall i \in V^c, m, t \tag{7}$$

The timing vector $y$ defines continuous arrival times $y_{im}$, and the constraint set $Tx + Wy \leq r$ in (3) represents temporal feasibility and time window constraints, ensuring the synchronization of arrivals with time-dependent durations and active buckets while enforcing store receiving windows.

*Temporal Feasibility and Windows:*

$$y_{im} + U_i + d_{ijt} - M_{ijmt}(1 - x^{\text{arc}}_{ijmt}) \leq y_{jm}, \qquad \forall i, j \in V^c(i), m, t \tag{8}$$

$$\sum_{t} S_t x^{\text{node}}_{imt} \leq y_{im} \leq \sum_{t} F_t x^{\text{node}}_{imt}, \qquad \forall i, m \tag{9}$$

$$E_i \leq \sum_{m} y_{im} \leq L_i, \qquad \forall i \in V \tag{10}$$

For notational brevity and adhering to the symposium's word limit, the rest of this paper is presented in the compact form defined in (1)-(4), while the complete extensive formulation is available upon request.

To account for disruption-related uncertainty impacting the across arcs (i.e., parameter $d_{ijt}$ in the extensive model), the deterministic TDVRPTW is extended to a two-stage stochastic formulation. In the stochastic model, the route design is fixed before uncertain travel times are realized, while the timing variables become scenario-dependent recourse variables. The two-stage model is

$$\min_{x} \quad C^{\text{travel}}(x) + C^{\text{traffic}}(x) + C^{\text{setup}}(x) + E_{\omega\in\Omega}\big[Q(x,\omega)\big] \tag{11}$$

$$\text{s.t.} \quad Ax = b$$

$$x \in X,$$

where $\Omega$ is the set of travel-time scenarios. In the stochastic extension, the deterministic timing variables $y$ are replaced by scenario-dependent timing variables $y^\omega$, so their contribution is captured through the expected recourse term. For a fixed first-stage solution $x$ and scenario $\omega$, the second-stage recourse problem is

$$Q(x,\omega) = \min_{y^\omega} \quad C^{\text{TW-violation}}(y^\omega) \tag{12}$$

$$\text{s.t.} \quad W_\omega y^\omega \leq r_\omega - T_\omega x,$$

$$y^\omega \geq 0.$$

Operationally, the first stage determines the planned routing and service structure, whereas the second stage evaluates its reliability once recurrent congestion and disruptions are realized. This makes the framework suitable for traffic-management applications in urban freight, where routes must be planned in advance but schedule performance depends on realized network conditions.

***Solution approach***

The resulting STDVRPTW is a large-scale mixed-integer stochastic optimization problem whose extensive form becomes difficult to solve directly as the numbers of customers, time periods, and scenarios increase. To retain the two-stage structure while improving tractability for traffic-management applications, we solve the model using an exact improved Benders decomposition framework. The master problem contains the scenario-independent first-stage routing decisions, while the scenario subproblems evaluate the second-stage recourse cost under realized travel times. Because classical Benders may converge slowly when the master problem admits weak fractional routing solutions (Toth & Vigo, 2014), we strengthen it by deriving proper valid inequalities based on bucket time-budget inequalities, lifted-cover (Wolsey, 1998) inequalities, and rounded-capacity inequalities (Toth & Vigo, 2014), generated dynamically during the Benders process. In compact form, the improved Benders master problem is

$$\begin{aligned} \min_{x,\theta} \quad & C^{\text{travel}}(x) + C^{\text{traffic}}(x) + C^{\text{setup}}(x) + \sum_{\omega \in \Omega} p_\omega \theta^\omega \qquad (13) \\ \text{s.t.} \quad & Ax = b \\ & x \in X \\ & \text{Benders optimality cuts} \\ & \text{Benders feasibility cuts} \\ & \text{Bucket time-budget inequalities} \\ & \text{Cover and lifted-cover inequalities} \\ & \text{Rounded-capacity inequalities.} \end{aligned}$$

where $\theta^\omega$ denotes the master approximation of second-stage cost under scenario $\omega$ with probability $p_\omega$. For a given incumbent first-stage solution, the scenario subproblem is the recourse model in (12). If the recourse problem is feasible and bounded, Benders optimality cuts refine the expected second-stage cost approximation; if infeasibility occurs, feasibility cuts exclude first-stage solutions that cannot support a valid schedule under the corresponding scenario.

## 3 Results and discussion

Numerical experiments were conducted on middle-mile food and grocery distribution instances. Travel times were modeled for four daily periods—AM Peak, Midday, PM Peak, and Late Day—and uncertainty was represented through scenarios reflecting recurrent congestion variability, incidents, and construction-related disruptions. All experiments were implemented in Python and solved using Gurobi.

***Model behavior and validation***

Figure 2 presents the customer time-window profile used in a small illustrative network with 1 depot and 8 customers. We compare three settings, presented in Table 1: (i) deterministic TDVRPTW without traffic penalties, (ii) deterministic TDVRPTW with traffic penalties, and (iii) stochastic TDVRPTW with traffic penalties. Table 1 and Figure 3 report the resulting routes.

Table 1: Comparison of routing settings under different traffic modeling assumptions

| Model | Setting | Obj. ($) | TWV Cost ($) | Traffic Cost ($) | TWV (hh:mm) | Num. of Vehicles | Avg. Route Length (hr) |
|---|---|---|---|---|---|---|---|
| TDVRPTW | Deterministic Travel Time | 344.47 | 11.50 | N/A | 0:23 | 1 | 6:13 |
| TDVRPTW | Deterministic Travel Time + Traffic Penalty | 539.23 | 87.00 | 0 | 2:54 | 2 | 3:13 |
| STDVRPTW | Stochastic Travel Time + Traffic Penalty | 624.75 | 20.84 | 0 | 0:42 | 3 | 2:28 |

The results show a clear progression across the three settings. Without traffic penalties, the deterministic model consolidates deliveries into a single route, yielding the lowest cost but high

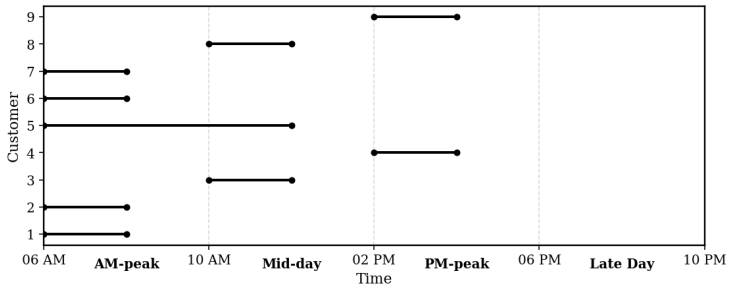
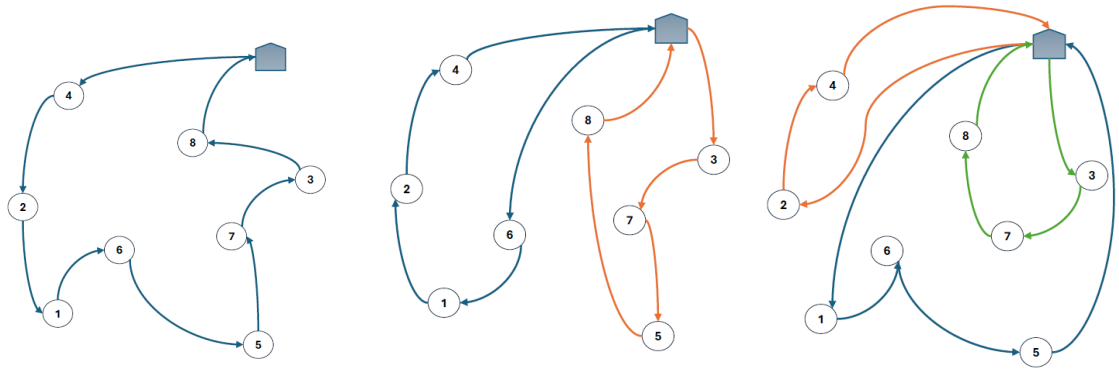
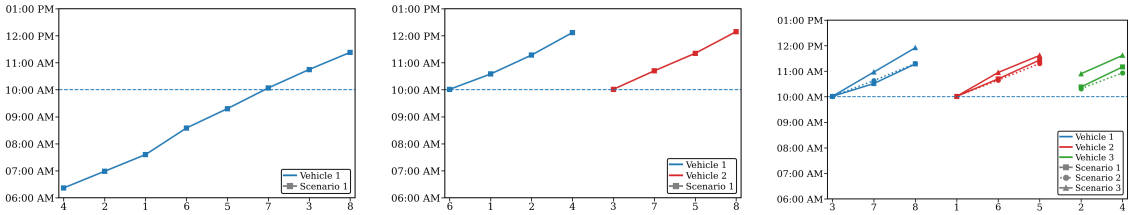



Figure 2: Customer time-window profile used in the demonstrative instance



(a) (b) (c)

Figure 3: Vehicle routing results under (a) TDVRPTW (no traffic penalty), (b) TDVRPTW (with traffic penalty), and (c) STDVRPTW




(a) (b) (c)

Figure 4: Vehicle arrival times at customers under (a) TDVRPTW (no traffic penalty), (b) TDVRPTW (with traffic penalty), and (c) STDVRPTW

congestion exposure. Adding traffic penalties shifts service to later periods and increases fleet use, reducing peak exposure but increasing time-window violations. Under the stochastic model, an additional vehicle is deployed and arrivals are spread more evenly across the day, improving robustness to incidents and construction-related delays. Although total cost increases, extreme violations are reduced and schedule stability improves. Overall, the comparison shows how the proposed framework responds to congestion pricing, delivery flexibility, and traffic uncertainty, while highlighting the trade-off between cost efficiency, congestion avoidance, and operational robustness.

***Computational Analysis***

Figure 5 highlights the computational superiority of the proposed Improved Benders over baseline benchmarks such as Gurobi and Standard Benders, primarily due to its strengthened master problem achieved through the systematic addition of Benders and problem-specific cuts, which tighten bounds, improve search guidance, and accelerate convergence.

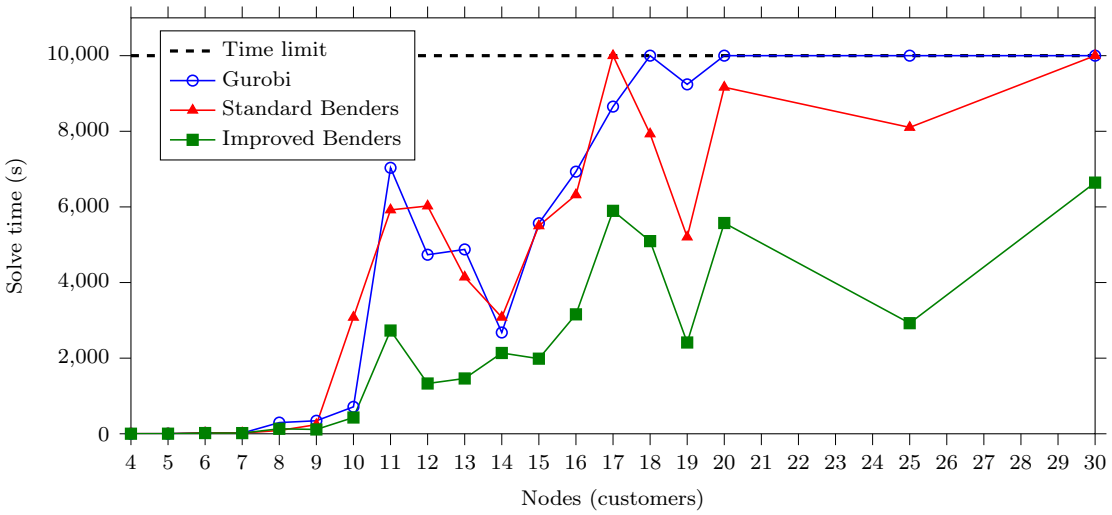



Figure 5: Computational Performance (Solve Time) of Gurobi, Standard Benders and Improved Benders.

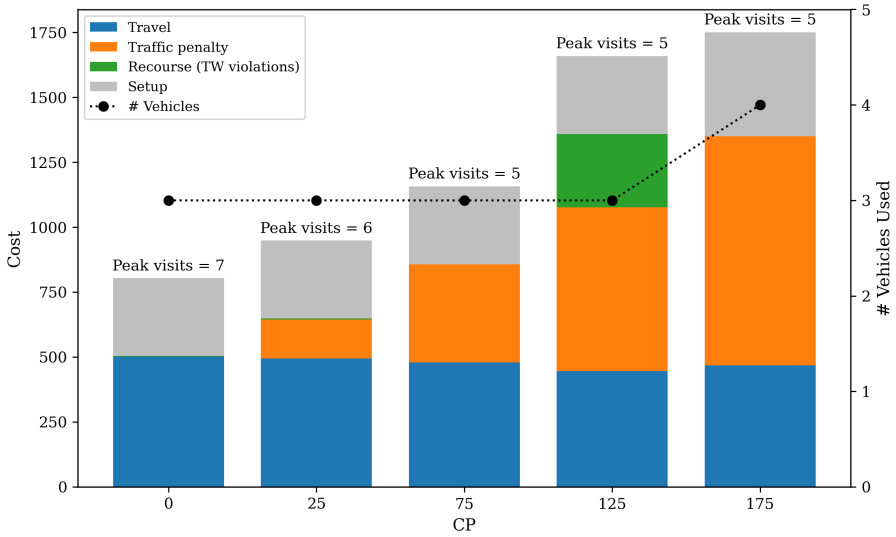



Figure 6: Peak-exposure ratio under different combinations of traffic penalty cost ($C_P$).

***Cost Decomposition under Congestion Pricing***

Figure 6 shows how the cost structure changes as congestion charges increase when the customer time-window violation penalty is fixed at $C_V = 100$. As the traffic penalty parameter $C_P$ rises, deliveries shift away from peak periods and the number of peak-period visits declines. At moderate penalty levels, this adjustment is achieved mainly through schedule changes, whereas beyond a critical threshold the model deploys an additional vehicle to maintain service reliability. The figure reports the corresponding decomposition into travel cost, traffic penalties, time-window violation cost, and vehicle setup cost.

The results suggest that congestion pricing can reduce peak-hour freight activity, but if set too high it may also encourage fleet expansion, increasing setup cost. Congestion charges should therefore be calibrated carefully and, where possible, combined with measures that increase delivery-time flexibility.

***Time-window flexibility***

Figure 7 shows the effect of time-window flexibility, where $\theta$ represents the level of time-window relaxation or flexibility. Greater flexibility improves both routing efficiency and congestion management, but the form of flexibility matters. Allowing earlier deliveries, or relaxing both bounds

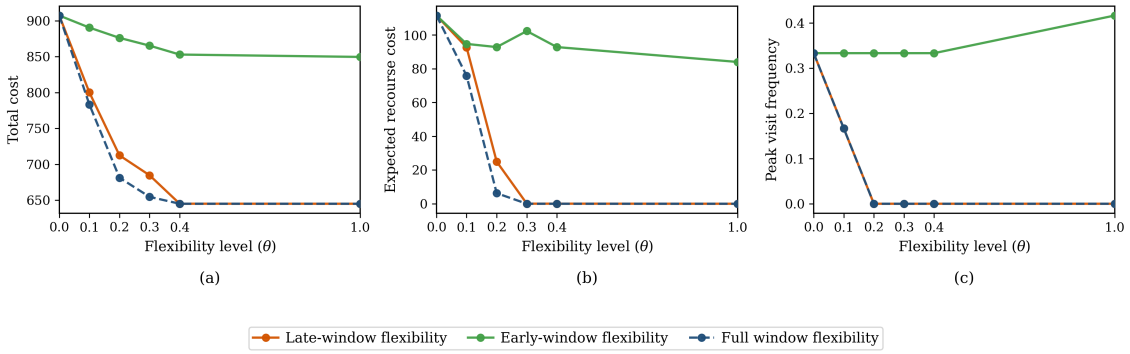



Figure 7: Impact of time-window flexibility on traffic-aware performance metrics: (a) Total cost, (b) Recourse cost, (c) Peak exposure.

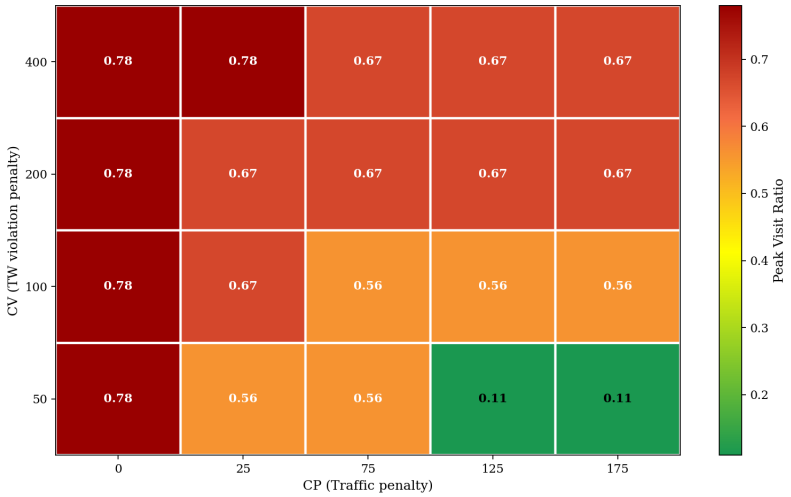



Figure 8: Peak-exposure ratio under different combinations of congestion pricing ($C_P$) and recourse penalty ($C_V$).

of the service window, yields the largest reductions in total cost, expected recourse cost, and peak exposure, whereas extending delivery deadlines alone provides only limited benefits. From a city logistics perspective, this suggests that policy should focus less on tolerating late arrivals and more on enabling earlier or more adaptable delivery access, for example through extended receiving hours or off-peak delivery programs.

***Interaction Effects of Peak Pricing and Service Flexibility***

Figure 8 shows that congestion pricing reduces peak-period freight activity only when operators have sufficient scheduling flexibility. When $C_P = 0$, peak exposure remains high across all values of $C_V$, indicating that time-window penalties alone do not shift deliveries away from congested periods. As congestion pricing increases, peak exposure declines, but the effect is much stronger when recourse penalties are low and operators can adjust delivery schedules more easily. Under strict time-window enforcement, pricing alone produces only limited changes in routing behaviour. This suggests that congestion pricing is most effective when combined with flexibility-enabling measures such as off-peak receiving programs or wider appointment windows.

## 4 Conclusions

This paper developed a two-stage stochastic, time-dependent routing framework for middle-mile urban freight and applied it to evaluate congestion-management policies under realistic traffic uncertainty in Halifax, Canada. The results show that congestion pricing can reduce peak-period freight activity, but its effectiveness depends strongly on the degree of delivery-window flexibility

available to operators. When flexibility is limited, pricing alone produces modest changes; when flexibility is increased, pricing leads to substantially larger shifts toward off-peak operations. The analysis also shows that explicitly representing disruption scenarios yields routing plans that are more robust to incidents and construction-related delays. Overall, the study demonstrates how an operations-research model can support urban freight policy design by linking routing decisions, service reliability, and congestion exposure within a single analytical framework.

## References


Allen, J., Browne, M., & Cherrett, T. (2012). Survey techniques in urban freight transport studies. *Transport Reviews*, *32*(3), 287–311. doi: 10.1080/01441647.2012.665949

Álvarez, P., Serrano-Hernandez, A., Lerga, I., & Faulin, J. (2024). Optimizing freight delivery routes: The time-distance dilemma. *Transportation Research Part A: Policy and Practice*, *190*, 104283.

Dabia, S., Ropke, S., Van Woensel, T., & De Kok, T. (2013). Branch and price for the time-dependent vehicle routing problem with time windows. *Transportation science*, *47*(3), 380–396.

Duan, J., He, Z., & Yen, G. G. (2021). Robust multiobjective optimization for vehicle routing problem with time windows. *IEEE Transactions on Cybernetics*, *52*(8), 8300–8314.

Fan, D., & Heminthavong, K. (2022, May). *Road transportation: Heavyweight of the canadian economy* (Tech. Rep. No. Publication No. 2022-04-E). Library of Parliament, Canada. Retrieved from `https://lop.parl.ca/sites/PublicWebsite/default/en_CA/ResearchPublications/202204E.pdf`

Flatberg, T., Hasle, G., Kloster, O., Nilssen, E. J., & Riise, A. (2007). Dynamic and stochastic vehicle routing in practice. *Dynamic fleet management: concepts, systems, algorithms & case studies*, 41–63.

Hou, D., Fan, H., Ren, X., Tian, P., & Lv, Y. (2021). Time-dependent multi-depot heterogeneous vehicle routing problem considering temporal–spatial distance. *Sustainability*, *13*(9), 4674.

Hübner, A., Holzapfel, C., & Meier, J. (2016). Distribution systems in omni-channel retailing. *Business Research*, *9*(2), 255–296. doi: 10.1007/s40685-016-0034-7

Miranda, D. M., & Conceição, S. V. (2016). The vehicle routing problem with hard time windows and stochastic travel and service time. *Expert Systems with Applications*, *64*, 104–116.

Rothenbächer, A.-K. (2019). Branch-and-price-and-cut for the periodic vehicle routing problem with flexible schedule structures. *Transportation science*, *53*(3), 850–866.

Spliet, R., Dabia, S., & Van Woensel, T. (2018). The time window assignment vehicle routing problem with time-dependent travel times. *Transportation Science*, *52*(2), 261–276.

Sun, P., Veelenturf, L. P., Dabia, S., & Van Woensel, T. (2018). The time-dependent capacitated profitable tour problem with time windows and precedence constraints. *European Journal of Operational Research*, *264*(3), 1058–1073.

Toth, P., & Vigo, D. (2014). *Vehicle routing: Problems, methods, and applications*. SIAM.

Wei, J., Poon, M., & Zhang, Z. (2024). A branch-and-cut algorithm for the time-dependent vehicle routing problem with time windows and combinatorial auctions. *Computers & Operations Research*, *172*, 106807.

Wolsey, L. A. (1998). *Integer programming*. John Wiley & Sons.

Zhang, J., Lam, W. H., & Chen, B. Y. (2013). A stochastic vehicle routing problem with travel time uncertainty: trade-off between cost and customer service. *Networks and Spatial Economics*, *13*(4), 471–496.