A lexicographic bi-objective approach to fleet sizing and routing for service vehicles in a real-world passenger transport system

Bouchra Z. Ben messabih[1], Walid Behiri[2], Sana Belmokhtar-Berraf[2], Tesseda Boukherroub[3], Abderrahim Sahli[2] and Iskander Zouaghi[1]

Abstract— **The Société de transport de Montréal (STM), a public transport agency in Montreal (Quebec, Canada), has committed to electrifying its entire vehicle fleet as part of its broader sustainability initiatives. This research addresses a critical challenge that arises from this commitment: optimizing the deployment and dispatching of electric service vehicles assigned to operations supervisors to ensure minimal disruption to the bus network. The problem is formulated as a mixed-integer linear program that minimizes the number of deployed vehicles and total response time using a lexicographic approach. The effectiveness of our approach is evaluated through computational experiments using the commercial CPLEX optimization solver. This research provides the STM with a valuable strategic and operational decision-support tool, aimed at optimizing the size and deployment of its new electric vehicle fleet, therefore supporting its sustainability objectives.**

I. INTRODUCTION

In recent years, sustainable urban mobility has become a crucial goal for cities worldwide. The STM, a public transportation agency in Montreal (Quebec, Canada), is leading the way with its ambitious initiative to electrify its entire fleet of vehicles by 2030. This transition is designed to tackle significant environmental challenges, such as reducing greenhouse gas emissions and minimizing air pollution [1], while also addressing the economic imperatives of cost management and operational efficiency.

This study examines the service vehicles assigned to Operations Supervisors (OS), a pivotal role in mitigating service disruptions such as detours, signage malfunctions, equipment failures, and passenger-related incidents. By electrifying these vehicles, the STM aims not only to meet environmental goals but also to enhance public transportation quality in Montreal. Reducing delays is a key measure of service satisfaction, as it significantly improves the daily commuting experience. Moreover, a modern and efficient transportation system elevates urban living, contributing to a better quality of life and increasing the appeal of large cities for both residents and businesses [2]. Achieving optimal operational efficiency within this electrified OS vehicle fleet presents a significant challenge: determining the ideal fleet size and deployment strategy. Effective resource allocation is paramount to ensure comprehensive coverage and rapid response capabilities across the STM's bus network.

The present paper proposes a decision support tool aiming to address the specific challenges of service vehicle deployment and dispatching encoutered by STM. The core focus is on two interconnected aspects: First, strategic vehicle deployment, which involves determining the optimal number of vehicles to locale in each depot to guarantee comprehensive coverage of the entire island of Montreal. Second, operational dispatching, which focuses on allocating available OS vehicles to occuring incidents over time and space. The problem is formulated as a mixed-integer linear program that minimizes, following a lexicographic approach, the number of deployed vehicles and the total response time.

The remainder of the paper is structured as follows. Section II reviews the relevant literature on vehicle deployment and dispatching. Section III provides a detailed description of the problem. Section IV presents the proposed mixed-integer linear programming model. Section V presents the results of numerical experiments. Finally, Section VI concludes the paper and outlines some perspectives for future research.

II. LITERATURE REVIEW

To address STM's challenge, the literature on emergency vehicle management offers valuable insights. Researchers have long examined problems analogous to STM's, including those encountered in ambulance, fire truck, and other emergency service operations. In this context, the literature is generally divided into several interconnected problems: location, deployment, assignment, dispatching, and relocation. The location problem focuses on identifying the best sites serving as depots, while the deployment problem determines the number of vehicles that should be assigned to each depot. The assignment problem determines which service areas or demand zones are covered by which depots, and the dispatching problem decides which vehicle should be used to respond to an incident [3]. Additionally, the relocation problem involves repositioning idle vehicles to reduce future response times. In the realm of traffic emergency systems, [4] defines deployment as the process of allocating vehicles to stations, and dispatching as the real-time decision of sending a vehicle to meet an incident demand. [5] further emphasizes that most dispatching models assign tasks to the nearest available resource, a strategy also highlighted by [6], who notes that recent work has increasingly focused on the joint

[1]École Nationale Polytechnique d'Alger, El Harrach 16200, Algerie, bouchra_zohra.ben_Messabih@g.enp.edu.dz; iskander.zouaghi@g.enp.edu.dz

[2]Univ Gustave Eiffel, ESIEE Paris, COSYS-GRETTIA, F-77454 Marne-la-Vallée, France, walid.behiri@univ-eiffel.fr; sana.berraf@univ-eiffel.fr; abderrahim.sahli@univ-eiffel.fr

[3]École de Technologie superieure (ÉTS) & CIRRELT Montréal, Canada, tasseda.boukherroub@etsmtl.ca

problem of dispatching and proactive relocation to enhance overall coverage.

These operational issues are closely linked to strategic planning. According to the classification proposed by [7], decision-making can be divided into three levels: first, there is the strategic level with long-term decisions involved such as depot location and fleet sizing. Second, the tactical level addresses medium-term issues like baseline deployment and shift scheduling. Finally, the operational level focuses on short-term decisions like dispatching and dynamic relocation. This classification is considered in [8] and [9] with strategic decisions laying the groundwork for operational efficiency by determining where vehicles should be located, while operational decisions ensure that vehicles are dispatched and, if necessary, redeployed effectively in real time.

Recent reviews, such as that by [10], indicate a trend towards integrated approaches. These models combine multiple methodologies often sequentially or iteratively to harness the strengths of each, thereby optimizing both deployment and dispatching simultaneously. This integrated perspective is particularly relevant for STM's challenge of managing an electric fleet to improve performances.

Because of the dynamic and uncertain conditions, simulation modeling has emerged as a crucial tool. Researchers employ various types of discrete event simulations [11], continuous, hybrid, Monte Carlo, and agent based [12] to capture the uncertainty inherent in emergency scenarios. Simulations help in two main ways: by assessing the impact of parameter changes on system performance and by evaluating the robustness of mathematical programming solutions [10]. Others stochastic approaches, such as queuing theory and Markov decision processes, provide additional insights into system dynamics and response times [3], [13].

Compared to similar studies in emergency services where the primary objective is minimizing rescue time due to its direct impact on saving lives, our research takes a distinct approach. Our study focuses on optimizing the fleet size of STM's electric service vehicles. This decision is crucial from an operational perspective: for instance, if multiple incidents occur simultaneously or if there is a consistent pattern in incident frequency, the STM must avoid unnecessary capital expenditure on acquiring redundant vehicles. Therefore, our goal is to reduce investments by integrating long-term strategic decisions with short-term operational ones. In addition to the decisions related to fleet sizing, deployment and dispatching, the routing of service vehicles plays an equally critical role at the operational level. Although the classic Vehicle Routing Problem (VRP) primarily focuses on minimizing routing costs [14] and has evolved to include numerous practical constraints such as vehicle capacity, time windows, and multi-trip operations [15] its adaptations in the emergency response context have been less extensively explored. For example, [16] emphasizes that while research in disaster response typically concentrates on stationing and dispatching ambulances, only limited attention has been paid to ambulance routing, particularly when multiple pick-ups are involved. In our framework, however, the routing of STM's service vehicles is crucial since a single vehicle may respond to several incidents during the same shift. Drawing inspiration from the Multi-Depot Vehicle Routing Problem with Time Windows (MDVRP-TW) [17], [18], our model leverages established routing principles to efficiently manage the complex interplay between multiple depots, dispatching, and time window constraints driven by the incident's occurrence date and the need for prompt intervention while also accommodating multi-incident pickups.

Table I provides a comparative overview of related works, highlighting key dimensions such as the application sector, the nature of the addressed challenges, the type of demand considered (deterministic or stochastic), the approach used to handle stochasticity, the modeling framework, and the objective function(s). This comparison underscores the distinctive features of our contribution. Specifically, our work proposes a decision-support tool for operations planning that seeks to minimize response times when feasible, while simultaneously optimizing fleet size to balance cost-efficiency with demand fulfillment.

Despite the extensive focus on emergency vehicle deployment and dispatching, to the best of our knowledge, no previous work has specifically addressed the optimization of electric fleet sizing, deployment and routing under urban operational conditions for public transit operations, while simultaneously considering both strategic and operational challenges. This study aims to fill this gap. Our contribution lies in formulating a real-world problem and developing a novel mathematical model that incorporates detailed factors such as incident timing and location, travel times, distances, and the specific constraints of electric vehicles.

III. Problem Description

This study focuses on the island of Montreal, which is geographically divided into 15 distinct sectors. Across these sectors, the STM operates bus network, supported by 8 depots. These depots are not only used for bus storage, but also play a crucial role in maintenance operations, electric charging, and other essential services. To ensure smooth operation of this network, the STM relies on a dedicated team of OS, each assigned to one of the eight depots and work in rotating 6-hour shifts.

When an incident occurs at any point within the network, the Central Operations Center (COC) is immediately activated to coordinate a rapid response. The COC connects the department responsible for transportation planning, management and monitoring with the one in charge of infrastructure and equipment management, maintenance, and repair. Its main task is to rapidly dispatch the closest OS to the incident location. Each incident is characterized by its specific geographical coordinates and an estimated handling duration.

The proposed approach is not designed to function as a real-time operational system for the STM to manage incident responses, given the inherently unpredictable nature of such events in both time and space. Rather, its primary objective is to optimize the sizing and spatial allocation of the service

TABLE I: Comparison of different papers addressing similar problems

Papers	Sector	Problem Treated			Problem Features	Approach	Modeling	Objective(s)
		D	DS	RT				
[3]	Emergency	✓	✓	-	Stochastic demand and traffic state	Scenario Generation, SAA (MC)	MIP BD	Total cost
[4]	Traffic	✓	✓	-	Stochastic demand, time and location	Scenario Generation SAA (MC)	MIP VNS	Total cost
[5]	Medical	-	✓	✓	Deterministic	-	MILP	Total cost
[6]	Fire trucks	-	✓	-	Stochastic demand and driving time	MDP	OSI, OSIA	Fraction of late arrivals
[7]	Medical	✓	✓	-	Stochastic demand	Scenario Generation	MIP	Fraction of calls with late response times
[9]	Medical	-	✓	-	Stochastic demand	Scenario Generation, SAA (MC)	MIP HDT, SBG	Total cost
[12]	Emergency	-	✓	-	Stochastic demand and response time	MAS	MIP	Travel time
[16]	Medical	-	✓	✓	Deterministic	-	MIP LNS	Latest service completion time
[19]	Army	✓	✓	-	Deterministic	-	Bi-obj MILP GA	Total cost Travel time
[20]	Traffic	-	✓	-	Stochastic traffic state	Queuing theory	MIP	Travel time
Our work	**Service vehicle**	✓	**-**	✓	**Deterministic**	**-**	**Bi-obj MILP**	**Fleet size Response time**

*Abbreviations: **D**: Deployment, **DS**: Dispatching, **RT**: Routing, **SAA**: Sample average approximation,**MC**: Monte Carlo, **MAS**: Multi-Agent System, **VNS**: Variable Neighborhood Search,**MDP**: Markov Decision Process, **BD**: Benders Decomposition, **OSI**: One-Step Improvement, **OSIA**: One-Step Improvement Approximation Heuristic, **GA**: Genetic Algorithm, **HDT**: Temporal Decomposition Heuristic, **SBG**: Lagrangian Relaxation-based Heuristic, **LNS**: Large Neighborhood Search.*

vehicle fleet across depots, aiming to achieve cost-efficient investment decisions without compromising service quality or responsiveness. Since the core mission of these vehicles is to address operational incidents, it is crucial to incorporate this operational dimension into the modeling framework. To that end, we use incident data generated by a forecasting model trained on several years of historical observations. The model also supports post hoc analysis, enabling the evaluation of ideal fleet sizes and deployment strategies for effectively managing past incidents, thereby offering valuable insights for future planning. This work constitutes an initial step toward the development of a more operational decision-support system capable of integrating real-time incident data into the decision-making process.

In order to clearly define the problem context, we introduce the following assumptions:

- All incidents or interventions are treated with equal priority.
- Each incident requires the intervention of exactly one vehicle.
- The service time for each incident is determined from empirical historical data.
- It is assumed that the time when an incident occurs is known based on prior predictions.
- The fleet consists of homogeneous vehicles, meaning that all vehicles share identical characteristics.
- The speed of the vehicles is assumed to be known and constant, which facilitates travel time estimations.
- The travel distances between nodes are computed using the Haversine formula to accurately reflect real-world geodesic distances.

IV. Mathematical model

As the company has already defined the priority between its two objectives within the framework of this study, with fleet size minimization taking precedence over response time, our formulation adopts a lexicographic bi-objective Mixed-Integer Linear Programming (MILP) approach that balances these conflicting goals. This model simultaneously supports strategic planning and operational execution by determining the number of vehicles to deploy, selecting the originating depot for each vehicle, and efficiently assigning them across the service area, while also optimizing the routing of each vehicle during their service shift.

Formally, the problem is defined using the following sets and parameters. Let $\mathcal{K}$ be the set of OS vehicles, each with a constant speed v and a battery capacity B, expressed as a maximum travelable distance. Let $\mathcal{D}$ be the set of depots, each with a capacity Q, and $\mathcal{N}$ the set of incidents to be handled. The set of all nodes is denoted by $\mathcal{V} = \mathcal{D} \cup \mathcal{N}$. For every pair of nodes $i, j \in \mathcal{V}$, let t_{ij} denote the travel time between them. Each incident $i \in \mathcal{N}$ is associated with a service time s_i, representing the duration of the intervention carried out by an OS vehicle. The remaining notations used in this problem are as follows:

Parameters:

$[e_i, l_i]$	Time window for each incident i.
S	Vehicle shift duration.
M	A large positive number.

Decision Variables:

x_{ij}^k	1 if vehicle k travels from i to j, 0 otherwise.
y_k	1 if vehicle k is used, 0 otherwise.
o_{dk}	1 if vehicle k is assigned to depot d, 0 otherwise.
a_{ik}	Arrival time of vehicle k at node i.
w_{ik}	Waiting time of vehicle k at node i.

The lexicographic bi-objective MILP model is defined as follows:

$$\min Z_1 = \sum_{k \in \mathcal{K}} y_k \tag{1}$$

$$\min Z_2 = \sum_{i \in \mathcal{N}} \sum_{k \in \mathcal{K}} a_{ik} \tag{2}$$

Subject to:

$$\sum_{i \in \mathcal{V}, i \neq j} \sum_{k \in \mathcal{K}} x_{ij}^k = 1, \quad \forall j \in \mathcal{N} \tag{3}$$

$$\sum_{j \in \mathcal{V}, i \neq j} \sum_{k \in \mathcal{K}} x_{ij}^k = 1, \quad \forall i \in \mathcal{N} \tag{4}$$

$$\sum_{i \in \mathcal{V}, i \neq j} x_{ij}^k = \sum_{i \in \mathcal{V}, i \neq j} x_{ji}^k, \quad \forall j \in \mathcal{V}, \forall k \in \mathcal{K} \tag{5}$$

$$e_i \cdot \sum_{j \in \mathcal{V}, i \neq j} x_{ji}^k \leq a_{ik} \leq l_i \cdot \sum_{j \in \mathcal{V}, i \neq j} x_{ji}^k, \quad \forall i \in \mathcal{N}, \forall k \in \mathcal{K} \tag{6}$$

$$a_{ik} \leq t_{di} + w_{dk} + M \cdot (1 - x_{di}^k), \quad \forall i \in \mathcal{N}, \quad \forall d \in \mathcal{D}, \forall k \in \mathcal{K} \tag{7}$$

$$a_{ik} \geq t_{di} + w_{dk} - M \cdot (1 - x_{di}^k), \quad \forall i \in \mathcal{N}, \quad \forall d \in \mathcal{D}, \forall k \in \mathcal{K} \tag{8}$$

$$a_{jk} \leq a_{ik} + s_i + w_{ik} + t_{ij} + M \cdot (1 - x_{ij}^k), \quad \forall i \in \mathcal{N}, \quad \forall j \in \mathcal{V}, i \neq j, \forall k \in \mathcal{K} \tag{9}$$

$$a_{jk} \geq a_{ik} + s_i + w_{ik} + t_{ij} - M \cdot (1 - x_{ij}^k), \quad \forall i \in \mathcal{N}, \quad \forall j \in \mathcal{V}, \quad i \neq j, \quad \forall k \in \mathcal{K} \tag{10}$$

$$a_{ik} \leq S, \quad \forall i \in \mathcal{V}, \forall k \in \mathcal{K} \tag{11}$$

$$v \cdot \sum_{i \in \mathcal{V}, i \neq j} x_{ij}^k \cdot t_{ij} \leq B, \quad \forall k \in \mathcal{K} \tag{12}$$

$$\sum_{d \in \mathcal{D}} o_{dk} = 1, \quad \forall k \in \mathcal{K} \tag{13}$$

$$o_{dk} \geq \sum_{j \in \mathcal{N}} x_{dj}^k, \quad \forall d \in \mathcal{D}, \forall k \in \mathcal{K} \tag{14}$$

$$\sum_{i \in \mathcal{V}} x_{id}^k \leq \sum_{j \in \mathcal{V}} x_{dj}^k, \quad \forall d \in \mathcal{D}, \forall k \in \mathcal{K} \tag{15}$$

$$\sum_{i \in \mathcal{V}} x_{id}^k \leq 1, \quad \forall d \in \mathcal{D}, \forall k \in \mathcal{K} \tag{16}$$

$$\sum_{j \in \mathcal{V}} x_{dj}^k \leq 1, \quad \forall d \in \mathcal{D}, \forall k \in \mathcal{K} \tag{17}$$

$$M \cdot y_k \geq \sum_{i \in \mathcal{V}} \sum_{j \in \mathcal{V}} x_{ij}^k, \quad \forall k \in \mathcal{K} \tag{18}$$

$$\sum_{k \in \mathcal{K}} o_{dk} \leq Q, \quad \forall d \in \mathcal{D} \tag{19}$$

$$x_{ij}^k, y_k, o_{ik}, \in \{0, 1\}, \quad \forall i, j \in \mathcal{V}, \forall k \in \mathcal{K} \tag{20}$$

$$a_{ik}, w_{ik}, \in \mathbb{R}, \quad \forall i \in \mathcal{V}, \forall k \in \mathcal{K} \tag{21}$$

In this model, the two objective functions (1–2) are prioritized using a lexicographic method, representing the minimization of fleet size and total response time, respectively. Given that incident occurrence times are constant, minimizing the total response time is equivalent to minimizing the total arrival time of vehicles. Constraints (3-4) ensure that each incident is served by exactly one vehicle, thus ensuring full coverage of all incidents. Constraints (5) guarantee flow conservation, meaning that every vehicle that arrives at an incident node will leave it. Time-related constraints, including time windows, waiting time, and service time, are defined in constraints (6-10). Constraints (11) ensure that no vehicle is assigned to an incident occurring after the end of its designated shift. Constraints (12) impose battery capacity limitations on each vehicle. Constraints (13-17) require that each vehicle is assigned to a specific depot and that each route both starts and ends at the same depot. Constraints (18) define the variable y_k, indicating whether vehicle k is used. Finally, constraints (19) impose the capacity limits of the depots.

V. Experimental Results

To assess the computational performance of the proposed model, we implemented it in Java. It was solved using the CPLEX 22.1 optimization solver. Numerical experiments were conducted on a machine equipped with an Intel Core i5-6300U (2 cores, 4 threads, 2.4 GHz base frequency) and 8 GB of RAM. A time limit of 600 seconds was imposed for each optimization run. The following subsections describe the data generation procedure and provide a detailed performance analysis of the MILP model.

A. Data description

The instance generation process focused on a representative region within the Island of Montreal. We constructed seven families of instances, each consisting of 25 individual instances. Each family is defined by a specific combination of the number of incidents $|\mathcal{N}|$, the number of depots $|\mathcal{D}|$, and the number of available vehicles $|\mathcal{K}|$. To systematically evaluate the model performance, we incrementally increased these parameters, primarily using a step size of 1. For each instance:

TABLE II: Performance of the MILP

Class $(\|\mathcal{N}\|, \|\mathcal{K}\|, \|\mathcal{D}\|)$	**Nb Var**	**Nb Cons**	**Solved Opt**	**Gap %**		**CPLEX time (s)**	
				Average	**Standard deviation**	**Average**	**Standard deviation**
(6, 2, 4)	332	601	25	-	-	0.14	0.07
			25	-	-	0.37	0.16
(7, 3, 5)	620	1100	25	-	-	0.22	0.08
			25	-	-	1.17	0.56
(8, 3, 5)	735	1322	25	-	-	0,55	0,36
			25	-	-	5,28	6,05
(9, 4, 6)	1032	1873	25	-	-	1.35	1.02
			25	-	-	13.94	16.20
(10, 4, 7)	1603	2870	25	-	-	51.44	73.79
			14	15.03%	7.60%	336.06	209.23
(11, 4, 7)	1820	3292	24	33.33%	0.00%	135.40	355.42
			7	23.29%	8.71%	164.75	198.40
(12, 4, 7)	2051	3742	21	33.33%	0.00%	275,93	252,14
			1	28,71%	8,42%	476,62	134,59

- Geographic coordinates of incidents are generated randomly within the city of Montreal. Pairwise distances are computed using the Haversine formula and converted into travel times, assuming a fixed and known vehicle speed.
- Service times s_i at each incident are sampled from the historical distribution of intervention durations.
- Time window parameters $[e_i, l_i]$ are generated randomly, inspired by patterns observed in the operational database.

Given the bi-objective nature of our problem, we employed the lexicographic approach. This method prioritizes the objectives sequentially. First, we minimize the number of vehicles required (1). Once the minimum fleet size was determined, we solved the model a second time, focusing on minimizing the total response time (2). In this second optimization run, we added a constraint to fix the number of vehicles to the optimal value obtained in the first step. This value served as an upper bound constraint on the fleet size, ensuring that the second objective was optimized without increasing the previously determined minimal fleet size.

B. *Performance of the MILP model*

To evaluate the computational efficacy of the MILP model, we conducted a series of experiments across varying instance families. Table II summarizes the key performance metrics obtained from these experiments. For each instance family, we report the number of variables and constraints. Then, for each objective function, two rows are displayed: the first corresponds to the results obtained with objective (1), and the second with objective (2). We indicate the number of instances (out of 25) that were solved optimally (i.e., with a gap of 0), along with the average and standard deviation of the optimality gap for the remaining instances. We also report the average execution time in seconds, along with its standard deviation, for each objective function. The 10-minute time limit corresponds to the total cumulative solving time allocated to both objectives within the lexicographic approach. The results indicate that the first objective minimizing fleet size is generally solved more rapidly than the second objective, which aims to minimize the arrival time. The model demonstrates its ability to solve instances with up to 9 incidents within the 10-minute time limit. With 10 incidents, some instances are no longer solved to optimality within the 10-minute time limit. Indeed, while the first objective function is solved for all 25 instances, there are 11 instances for which the second objective yields an average optimality gap of approximately 15%, indicating that the exact approach struggles with larger instances. Moreover, the performance varies notably across instances of similar size, indicating that scenario characteristics (e.g., spatial distribution of incidents) directly impact problem complexity and solver efficiency. These highlight the limitations of the current exact method in handling complex or adverse scenarios and point to the need for more scalable, adaptive approaches, such as heuristic or decomposition-based techniques.

The STM problem exhibits natural decomposability due to its temporal and spatial structure, offering a promising strategy to address the scalability challenges of the exact MILP model. Historical data indicates that a typical day for the STM involves approximately 200 incidents requiring service vehicle intervention. The day is divided into four shifts, with fewer incidents occurring during off-peak periods. The region is also partitioned into 15 sectors, some of which generate fewer incidents. By focusing on a single shift and applying a sector-based decomposition, the problem can be reduced to smaller subproblems, each involving approximately 10 incidents. These subproblems can be solved independently using the MILP formulation, with a post-processing step to integrate solutions into a cohesive global solution. Preliminary tests on instances representing one-third of a typical operational day demonstrate that these subproblems are solved within seconds, consistently requiring fewer vehicles than the current STM operational fleet.

These results suggest that the decomposition-based approach, currently under development, could significantly enhance the efficiency and scalability of the proposed method.

VI. CONCLUSIONS & PERSPECTIVES

This paper introduced a bi-objective optimization framework for the strategic deployment and operational dispatching of electric service vehicles within the bus network of the société de transport de Montréal (STM). The problem was formulated as a mixed-integer linear programming (MILP) model that simultaneously minimizes the fleet size and the cumulative response time while respecting time windows, depot capacity and electric vehicle operation constraints.

The lexicographic approach was applied to prioritize the two objectives. Through extensive computational experiments on synthetically generated yet realistic STM-inspired instances, the model demonstrated its ability to handle small sized instances efficiently. However, scalability limitations emerged beyond ten incidents, due to the inherent complexity of the MILP formulation.

To address real-size data, a preliminary decomposition heuristic was proposed. It partitions the service area into sectors and iteratively solves localized subproblems. Early results indicate promising potential for application to larger instances, with no major deterioration in solution quality observed so far.

As a research perspective, we plan to further develop and evaluate the decomposition-based heuristic to confirm its potential. In addition, our literature review has highlighted genetic algorithms as a promising metaheuristic for this class of problems, which could be an interesting direction to explore in future work. Furthermore, hybrid metaheuristic frameworks and advanced decomposition techniques, such as Lagrangian Relaxation or Benders Decomposition, will be investigated to improve computational scalability and solution quality on large-scale instances. To enhance the model's realism, future extensions will also aim to incorporate stochastic demand, scheduling of vehicle charging operations at depot facilities, and dynamic dispatch policies under uncertainty.

ACKNOWLEDGMENTS

We would like to thank our collaborators from the Société de transport de Montréal (STM) and IVADO for their valuable support and insights throughout this work. We also gratefully acknowledge Mitacs – SVP for their support, which enabled our participation in this conference.

REFERENCES

[1] M. Hossain, L. Kumar, M. Islam, and J. Selvaraj, "A comprehensive review on the integration of electric vehicles for sustainable development," *Journal of Advanced Transportation*, vol. 2022, no. 1, p. 3868388, 2022.

[2] M. Mylonakou, A. Chassiakos, S. Karatzas, and G. Liappi, "System dynamics analysis of the relationship between urban transportation and overall citizen satisfaction: A case study of patras city, greece," *Systems*, vol. 11, no. 3, 2023.

[3] W. Wang, S. Wang, L. Zhen, and X. Qu, "EMS location-allocation problem under uncertainties," *Transportation Research Part E: Logistics and Transportation Review*, vol. 168, 12 2022.

[4] L. Zhen, J. Wu, F. Chen, and S. Wang, "Traffic emergency vehicle deployment and dispatch under uncertainty," *Transportation Research Part E: Logistics and Transportation Review*, vol. 183, 3 2024.

[5] Z. Zeng, W. Yi, S. Wang, and X. Qu, "Emergency Vehicle Routing in Urban Road Networks with Multistakeholder Cooperation," *Journal of Transportation Engineering, Part A: Systems*, vol. 147, 10 2021.

[6] D. Usanov, P. M. de Ven, and R. D. der Mei, "Dispatching fire trucks under stochastic driving times," *Computers and Operations Research*, vol. 114, 2 2020.

[7] D. Bertsimas and Y. Ng, "Robust and stochastic formulations for ambulance deployment and dispatch," *European Journal of Operational Research*, vol. 279, pp. 557–571, 12 2019.

[8] G. Zhang, R. Ma, Y. Kong, C. Lian, H. Guo, and S. Zhai, "A multi-period capacitated facility location problem with maximum travel time and backup service for locating and sizing EMS stations," *Computational Urban Science*, vol. 4, 12 2024.

[9] R. Boujemaa, A. Jebali, S. Hammami, and A. Ruiz, "Multi-period stochastic programming models for two-tiered emergency medical service system," *Computers and Operations Research*, vol. 123, 11 2020.

[10] Y. Frichi and F. Jawab, "A review and classification of ambulance deployment and redeployment models in emergencies 'A review and classification of ambulance deployment and redeployment models in emergencies'," Tech. Rep. 1, 2024.

[11] S. A. Rodriguez, R. A. De la Fuente, and M. M. Aguayo, "A simulation-optimization approach for the facility location and vehicle assignment problem for firefighters using a loosely coupled spatio-temporal arrival process," *Computers and Industrial Engineering*, vol. 157, 7 2021.

[12] S. Ibri, M. Nourelfath, and H. Drias, "A multi-agent approach for integrated emergency vehicle dispatching and covering problem," *Engineering Applications of Artificial Intelligence*, vol. 25, pp. 554–565, 4 2012.

[13] V. Bélanger, A. Ruiz, and P. S. and, "Déploiement et redéploiement des véhicules ambulanciers dans la gestion d'un service préhospitalier d'urgence," *INFOR: Information Systems and Operational Research*, vol. 50, no. 1, pp. 1–30, 2012.

[14] P. Toth and D. Vigo, *The vehicle routing problem*. SIAM, 2002.

[15] J. Tassone and S. Choudhury, "A Comprehensive Survey on the Ambulance Routing and Location Problems," 1 2020.

[16] L. Talarico, F. Meisel, and K. Sörensen, "Ambulance routing for disaster response with patient groups," *Computers and Operations Research*, vol. 56, pp. 120–133, 2015.

[17] A. Paul, R. Shankar Kumar, C. Rout, and A. Goswami, "Designing a multi-depot multi-period vehicle routing problem with time window: hybridization of tabu search and variable neighbourhood search algorithm," 2021.

[18] M. Rabbani, M. Akbarpour, M. Hosseini, and H. Farrokhi-Asl, "A multi-depot vehicle routing problem with time windows and load balancing: A real world application," *International Journal of Supply and Operations Management*, vol. 8, pp. 347–369, 9 2021.

[19] H. He, X. Zhang, J. Yang, X. Zhou, and H. Yan, "Optimization and Research on Army Vehicle Deployment in Emergency Situations," in *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, pp. 5314–5318, Institute of Electrical and Electronics Engineers Inc., 2024.

[20] G. Chai, J. Cao, W. Huang, and J. Guo, "Optimized traffic emergency resource scheduling using time varying rescue route travel time," *Neurocomputing*, vol. 275, pp. 1567–1575, 1 2018.