Location of Urban Micro-consolidation Centers to Reduce Social Cost of Last-Mile Deliveries of Cargo: a Heuristic Approach

Mario Arrieta-Prieto[1] | **Abdelrahman Ismael[2]** | **Carlos Rivera-Gonzalez[2]** | **John Mitchell[3]**

[1]Department of Industrial and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY, 12180, USA

[2]Department of Civil and Environmental Engineering, Rensselaer Polytechnic Institute, Troy, NY, 12180, USA

[3]Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY, 12180, USA

Correspondence
Mario Arrieta-Prieto, Department of Industrial and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY, 12180, USA
Email: arriem2@rpi.edu

Urban deliveries of goods have become a rising challenge due to the increase of online shopping in addition to the traditional business-to-business (B2B) dynamics. To maintain a satisfactory level of service, carriers must conduct delivery operations under restrictive environments, creating undesirable externalities such as congestion and pollution.

Urban micro-consolidation centers (UMCs) are defined as on-street or off-street spaces where all the deliveries within a certain radius are dropped and reconsolidated to be delivered by more sustainable last-leg modes. Their main purpose is to mitigate such negative externalities.

This work formulates the last-mile delivery problem assisted by UMCs as a mixed-integer quadratically-constrained program (MIQCP) and develops a greedy heuristic solution, inspired by the decomposition algorithms for large-scale optimization. The simulation results and a case study in Manhattan (NY, USA) show how the heuristic provides reasonable results with respect to the exact solution while also shedding light on how its solution can be utilized for freight-efficient urban design and policy planning.

KEYWORDS
Bi-level mixed-integer program, decomposition algorithm, facility location, greedy heuristic, last-mile logistics, urban consolidation

1 | INTRODUCTION

The impact of urbanization has been undeniable in the development of cities for the past decades. According to [29], 54% of the world's population lives in urban areas, and this proportion will increase to 66% by 2050. These changes in urban population combined with globalization and the rise of e-commerce have had tremendous impacts on how urban logistics operate. The aforementioned phenomena have also impacted considerably on congestion, air quality, land use, among others. Almost 17% of the congestion costs in the United States can be attributed to trucks, which constitute around 7% of the traffic in urban areas [33]. Furthermore, 25% of the total transportation-related CO_2 (Carbon Dioxide) emissions are produced by freight trucks in urban settings, as well as 30-50% of the other transport pollutants, such as NO_x (Oxids of Nytrogen) and particulate matter [27].

The rise of e-commerce has exacerbated the problems of congestion and pollution as a result of the increasing demand of freight transportation services associated with delivery operations to households. The sales of retail e-commerce in the United States (US) increased from 4,476 million to 158,049 million US dollars (USD) over the twenty-year period of 1999-2019, which is a record increase of 3431% [30]. This made the percentage of retail e-commerce 11.4% of the total retail sales in the US by the fourth quarter of 2019 [31]. Forecasts anticipate that by 2025, this share will be doubled accounting for 25% of the total retail sales [4]. This new scenario has posed a challenge on traditional supply chains and congestion, due to the increased number of delivery trucks navigating in the network and trying to find available parking spaces to complete the deliveries.

The consequences of the rise of e-commerce are clearly illustrated in New York city (NYC), where the number of household deliveries tripled to 1.1 million packages daily from 2009 to 2017 [14]. This sudden change in demand has led to slower travel speeds in sectors of Manhattan where it reduced to only 4.35 kilometers/hour (km/h) [14]. Another consequence the rise of e-commerce has brought is the large number of parking fines for carriers, since drivers have to trade the compliance with legal parking regulation for a better level of service in a highly competitive market, resulting in double-parking fines due to insufficient parking space in dense urban areas. In 2019, NYC reported issuing commercial parking fines worth 123 million USD, with one quarter of them attributed to two of the largest carriers operating in the country [2]. These freight transportation problems warrant a policy intervention that can deal with the rapid changes in the supply chains due to the new dynamics of goods commercialization.

Several public agencies, nationwide, have acknowledged and addressed freight transportation as one of the key components for having more livable and sustainable cities. Freight mobility plans have started to be implemented around the country with the intent to preserve the long-term vision of the cities while giving additional insights on how to make the freight transportation within the city core more efficient and sustainable. Traditionally, the two main objectives of a freight plan are: 1) to minimize congestion, environmental pollution and conflicts with communities; and 2) to enhance the economic productivity of freight operations.

Urban Micro-Consolidation Centers (UMCs) have been given a great deal of attention in freight mobility plans because of their critical role in fostering more organized last-mile delivery operations. A UMC is defined as a type of urban consolidation center that relies on sustainable modes, such as bikes (human-powered and/or electric) or light electric vans, to perform the last-mile delivery to customers [7]. Pragmatically, a UMC is a logistical area close to the city center that ranges from sophisticated tailored-made infrastructure (construction of a staging area or repurposing of an underutilized facility) to a simple redefinition of public space (a street lane, a parking garage, or even designated curbside space). The UMC does not necessarily require a specific physical infrastructure, as its sole purpose is to provide an area for delivery trucks to unload their goods, reconsolidate if necessary, and transfer them into more sustainable modes for delivery to consumers. This micro-consolidation operation brings positive externalities including safety improvements and reductions of congestion and emissions. On the other hand, UMCs can be the result of

private carrier strategic decisions to improve efficiency of their own operations; or a public policy measure adopted by city planners, and developed in agreement with private carriers, to mitigate the negative externalities of freight transportation and delivery.

UMCs have been operating in multiple cities throughout the world with tangible satisfactory results. For example, in Germany, they have been implemented in cities such as Hamburg and Frankfurt under different names ("micro-hubs" and "micro depots"). Pre-loaded containers have been located in central spots around the city to then, transfer the goods to sustainable modes for the last-mile deliveries. In Frankfurt, the UMC was the result of a public-private partnership with a major carrier operating in the city. This UMC was estimated to decrease annual CO_2 emissions by 25.5 t (metric tons) in 300 operating days [28].

In the case of Hamburg, the same carrier has been implementing UMCs since 2012 by placing four containers throughout the city and has successfully reduced externalities equivalent to 7-10 delivery vehicles operating in the city center on a working day. A study done by Hamburg School of Business Administration estimated that these UMCs would reduce 18,000-24,000 kilometers per year due to the reduction of vehicle usage in the last-mile operations [15]. The study also estimated that using UMCs would reduce traffic jams and accidents caused by trucks while parking. The positive results encouraged the same carrier to expand the idea to other cities including Dublin, in Ireland; and Leuven, in Belgium [22].

Although UMCs have proved to be successful in reducing congestion and pollution based on the vast amount of research that has discussed their implementation and consequences, no research efforts have been conducted on how to decide their optimal location in the urban cores. On one hand, private carriers want to do the deliveries as efficiently as possible to maximize their profits. On the other hand, the public sector is aiming to maximize the welfare on society. This raises a bi-level problem of facility location, where public sector requires the tools to optimally locate these UMCs to minimize both social and private costs of deliveries within the urban core, while considering the interaction between agents reflected in the supply chain operations. From that perspective, this paper aims to contribute to the literature by being the first paper that provides a mathematical formulation of the problem at hand, as a mixed-integer quadratically-constrained program (MIQCP), and developing a heuristic methodology particularly designed to provide reasonable solutions to it, tackling its inherent prohibitive complexity which does not allow commercial solvers to reach optimal solutions in a reasonable amount of time or with limited computer memory resources.

This paper is divided as follows. Section 2 discusses the proposed problem formulation in depth. The heuristic solution method is introduced and explained as well. Section 3 shows evidence of the performance of the heuristic for a small toy example and for a sample of simulated instances. This section concludes by showing and analyzing the policy planning implications of the results, after applying the heuristic, for a case study in New York City. Finally, the conclusions and limitations are discussed in Section 4.

2 | PROBLEM FORMULATION

Urban freight is composed of the interaction between four different agents: shippers, carriers, receivers and public sector. For UMCs, the problem reduces to the interaction between carriers, which act as the supply; receivers, which act as the demand; and public sector, as a regulatory entity. The three agents might pursue different objectives. Private carriers aim for doing deliveries as efficient as possible in order to maximize their profits. Receivers are only concerned about the fulfilment of demand. Lastly, public sector aims to maximize the welfare on society by reducing the social costs [12]. The welfare can be obtained by minimizing social costs, which include private costs and the monetary quantification of the externalities produced on society, by constructing and using the UMCs. For example, impacts

on emissions and on the private carriers' operations are taken into account. This work focuses on the interaction between carriers' and policy planners' decisions regarding the creation of one or multiple UMCs as a facility location problem with interacting agents [20], since it is assumed that the level of service to receivers remains unaltered after introducing UMCs.

The problem formulation assumes that the study area is a regular-shaped rectangle, as shown in Figure 1b, that consists of $M \times N$ homogeneous rectangular blocks. This layout was inspired by the urban design and layout of blocks in Manhattan (NYC) (Figure 1a), which was used as a study case in Subsection 3.3. The mathematical formulation exploits the features of this geometry, as it is discussed in depth in the rest of this Section.

(a)   (b)

FIGURE 1 (a) Spatial arrangement of blocks in Manhattan, NY; (b) Schematic representation of the area under analysis inspired in (a)

A UMC could be theoretically located in any of these blocks. The coverage area of each UMC is reasonably defined as a buffer zone of one block in each direction, being conservative about the technical specifications of the last-mile alternative modes considered. The illustrative location and coverage area for a UMC are presented in Figure 2. The coverage area can be widened depending on the particular characteristics of the context and the operation details (see Subsection 3.3). Whenever a given block is designated as a UMC, all the trucks will stop there to deliver the packages addressed to any of the blocks in the coverage area of the UMC, and these packages will then be re-consolidated and delivered in smaller shipments by the alternative modes (freight bicycles/tricycles or handcarts). Given this operation scheme, the number of stops and the length of the tours for the carriers' operations are significantly reduced, diminishing their total cost of operation as well as the externalities in terms of air pollution due to
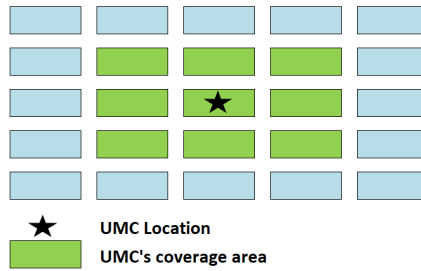
fossil fuel consumption.

FIGURE 2 Coverage area of a UMC

The boundary of the study area is a consequence of either geographic/urban design limitations (there are not urban blocks beyond that limit) or scope limitations (the study was truncated at some point to contain the size of the problem at hand). In any case, the location of a UMC in one of the blocks forming the boundary of the study area is suboptimal because its effective coverage area would not consist of eight adjacent blocks. There is another location in one of the internal blocks (defined as the blocks not on the boundary) that includes the coverage area of the location on the boundary in addition to also including other blocks in it.

To maximize the impact of each UMC in the number of stops and, ergo, in the length of the delivery tours; none of the boundary locations are selected as optimal locations (this is directly handled by the optimization problem). On the other hand, to avoid an extensive enumeration of particular cases for some of the constraints at the boundary, an additional ring of "imaginary" blocks was added to the area under study. In this ring, all the variables take the value of zero (no demand and ergo, no possible locations). These assumptions lead to the layout presented in Figure 3.
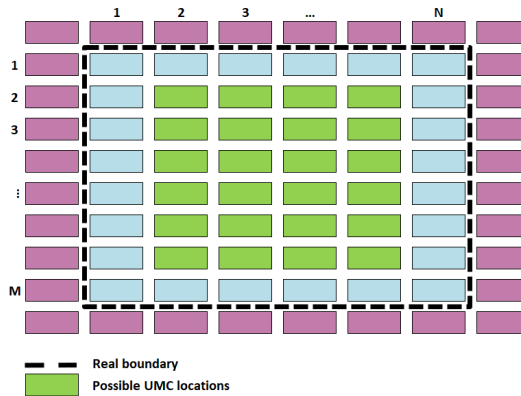

FIGURE 3 Complete representation of the study area for formulation purposes

With that layout in mind, the mathematical formulation is presented in the following Subsections. Although the characteristics of the UMCs are not described in detail due to the vast number of possibilities (that are situation-dependent and cannot be generalized), the mathematical model attempts to quantify, under a general setting, the social benefit of a reduction in the usage of trucks as the principal mode to complete the last-mile deliveries. The case

study in Section 3 highlights how these results can be interpreted towards determining the feasibility and profitability of the UMCs, by computing the maximum benefit obtained from their installation. This section also suggests how to conduct a cost-benefit analysis, once the full specification and operation definitions of the UMCs have been defined.

The problem is originally formulated with a bi-level structure considering potential differences in the objectives for carriers and policy planners. However, throughout the formulation, a justification about why this problem can be reduced to a single-level problem is described.

2.1 | Parameters and sets

Table 1 presents the parameters that the model requires as input. Some of the parameters are derived from other parameters using the expressions presented.

TABLE 1 Set and parameter definitions

Set / Parameter	Definition
M, N	Dimensions of the original study area (without the imaginary ring of blocks)
$[M]$	$:= \{0, 1, 2, \ldots, M, M+1\}$. Set of rows in the layout
$[N]$	$:= \{0, 1, 2, \ldots, N, N+1\}$. Set of columns in the layout
$\mathcal{R}$	$:= \{1, \ldots, M\} \times \{1, \ldots, N\}$. Set of real blocks in the layout
m	Number of carriers
$[m]$	$:= \{1, 2, \ldots, m\}$. Set of carriers
A	Area of the region under consideration (disregarding the imaginary ring)
ϕ	Constant used in Beardwood's approximation (see Subsection 2.4)
PL	Effective payload capacity used by a truck
$D_k \in \mathbb{R}^{(M+1) \times (N+1)}$	Demand matrix for each one of the carriers, $k \in [m]$
$d_k := \sum_{i \in [M]} \sum_{j \in [N]} (D_k)_{i,j}$	Total demand for carrier k, $k \in [m]$
$T_k := \lceil d_k / PL \rceil$	Total number of trucks used by carrier k, $k \in [m]$
$[T_k]$	$:= \{1, 2, \ldots, T_k\}$. Set of trucks used by carrier k, $k \in [m]$
λ	Cost of aggregate pollutant emissions per unit of distance
ω_k	Cost of private carrier k's operation per unit of distance; $k \in [m]$
θ	Number of UMCs to be located
$LV(i, j)$	$:= \{(i-1, j-1), (i-1, j), (i-1, j+1), (i, j-1), (i, j+1), (i+1, j-1), (i+1, j), (i+1, j+1)\}$. Set of blocks in the coverage area of block (i, j), $(i, j) \in \mathcal{R}$

2.2 | Decision variables

The variables are classified regarding the agent that has control over their values. Table 2 presents the decision variables regarding the carriers while Table 3 presents, in an analogous way, the ones related to the policy planners and the interaction between carriers and policy planners derived from the creation of UMCs.

TABLE 2 Decision variables for carriers

Variable	Definition
$(y_{sk})_{ij} \in \mathbb{R}_+$	Amount of cargo that truck s from carrier k has to deliver to block (i, j); $\forall s \in [T_k], k \in [m], (i, j) \in [M] \times [N]$
$(z_{sk})_{ij} \in \{0, 1\}$	If truck s from carrier k has to visit block (i, j), it is equal to 1; otherwise, it takes the value 0; $\forall s \in [T_k], k \in [m], (i, j) \in [M] \times [N]$
$L_{sk} \in \mathbb{R}_+$	Length of the tour for truck s corresponding to carrier k; $\forall s \in [T_k], k \in [m]$

TABLE 3 Variables related to policy planners and carrier-policy planner interactions

Variable	Definition
$x_{ij} \in \{0, 1\}$	If a UMC is located in the block (i, j), it is equal to 1; otherwise, it takes the value of 0; $\forall (i, j) \in [M] \times [N]$
$t_{ijlv} \in \{0, 1\}$	It takes the value of 1 if the demand for block (i, j) is delivered at block (l, v); otherwise, it is 0; $\forall i, l \in [M], \forall j, v \in [N]$

2.3 | Objective functions

The objectives are presented for each agent.

2.3.1 | Private carriers' objective

Each carrier seeks to minimize the operational cost of delivering the cargo, expressed as a linear function of the length of all the tours associated with each carrier [17] (assuming constant speeds in the area of study)

$$\tau_k := \min_{\mathbf{y_k}, \mathbf{z_k}, \mathbf{L_k}} \sum_{s \in [T_k]} \omega_k L_{sk}, \ k \in [m] \, ; \tag{1}$$

with $\mathbf{y_k}$, $\mathbf{z_k}$ and $\mathbf{L_k}$ being arrays of the decision variables $\{(y_{sk})_{ij}\}$, $\{(z_{sk})_{ij}\}$ and $\{L_{sk}\}$ respectively, corresponding to carrier k.

2.3.2 | Policy planner's objective

On the other hand, the policy planner attempts to mitigate the environmental impact of the carriers' operations by minimizing the total amount of emissions, which depends on the total distance traveled by all the carriers [18]

$$\min_{\{x_{ij}\}, \{t_{ijlv}\}} \sum_{k \in [m]} \lambda \tau_k. \tag{2}$$

The objective for the carriers has been presented disregarding the effect that the policy planner's decision has in their operations. However, as it is demonstrated in the Subsection 2.4, the decision made by the policy planner of locating one or several UMCs, has an impact in the way the carriers conduct their delivery operations. This effect is clearly captured by the set of constraints involving the variables $\{t_{ijlv}\}$ that are introduced in the following Subsection. It is shown how these variables are a joint decision made by both the carriers and the policy planner in a cooperative

effort.

2.4 | Constraints

The constraints are also presented for each type of agent along with the ones derived from the interaction of both agents, emphasizing the motivation behind each one of them.

2.4.1 | Private carriers' constraints

$$\sum_{s \in [T_k]} \sum_{i \in [M]} \sum_{j \in [N]} (y_{sk})_{ij} = \sum_{i \in [M]} \sum_{j \in [N]} (D_k)_{ij}, \quad \forall k \in [m]; \tag{3}$$

$$\sum_{i \in [M]} \sum_{j \in [N]} (y_{sk})_{ij} \leq PL, \quad \forall s \in [T_k], k \in [m]; \tag{4}$$

$$(z_{sk})_{ij} \geq \frac{(y_{sk})_{ij}}{\min \{d_k, PL\}}, \quad \forall s \in [T_k], k \in [m], (i,j) \in [M] \times [N]. \tag{5}$$

Constraint 3 indicates that demand requirements for each carrier must be met, while constraint 4 expresses that the amount of cargo transported by each carrier's truck cannot exceed the payload limit. Constraint 5 enforces a logical condition: a block is visited by a specific carrier if and only if there is cargo that needs to be delivered there. The constant $\min \{d_k, PL\}$ attempts to impose a restriction as tight as possible to avoid ill-posed formulations.

The length of the tour for each truck is approximately measured using Beardwood's approximation [3], which is a function of the number of stops in the tour. The length of the tour is no less than the continuous approximation

$$L_{sk} \geq \phi \sqrt{A} \sqrt{1 + \sum_{i \in [M]} \sum_{j \in [N]} (z_{sk})_{ij}},$$

where ϕ is an appropriate constant dependent on the type of distance considered. In this particular case, the continuous approximation with $\phi = 0.765$ has been estimated to reasonably approximate the length of the tour for convex and compact areas under the Euclidean distance [13]. The "1" in the radicand indicates the stop at the end of the tour to leave the area under analysis.

This approach was followed taking into account that vehicle routing problems (VRPs) are $\mathcal{NP}$−hard [19], so having multiple VRPs embedded within a mixed-integer facility location program increases the complexity substantially. However, the continuous approximation, per se, is a complicated constraint. For that reason, by exploiting the binary nature of the variables $(z_{sk})_{ij}$, then it is true that $(z_{sk})_{ij} = (z_{sk})_{ij}^2$; and the constraint can be reformulated as

$$L_{sk}^2 \geq \phi^2 A \left[1 + \sum_{i \in [M]} \sum_{j \in [N]} (z_{sk}^2)_{ij} \right], \quad \forall s \in [T_k], k \in [m]. \tag{6}$$

Constraint 6 corresponds to a quadratic constraint, which turns out to be a convex second-order cone constraint if all the variables are handled as continuous.

2.4.2 | Policy planner's constraints

$$\sum_{i\in[M]}\sum_{j\in[N]} x_{ij} \leq \theta. \tag{7}$$

Constraint 7 expresses that the number of installed UMCs cannot exceed a pre-determined limit. Sensitivity analysis about the parameter θ is important because it allows to determine the extra benefit of installing an additional UMC in the study area, i.e., to go from θ to $\theta + 1$. This type of analysis was exploited in Subsection 3.3 to determine the feasibility of installing a determined amount of UMCs based on the additional social benefit each one of them provides.

2.4.3 | Interaction between policy planner's and carriers' decisions

$$\sum_{s\in[T_k]} (y_{sk})_{ij} \geq (D_k)_{ij} \left(1 - \sum_{(l,v)\in LV(i,j)} x_{lv}\right), \quad \forall k \in [m], (i,j) \in \mathcal{R}; \tag{8}$$

$$\sum_{s\in[T_k]} (y_{sk})_{ij} \leq (D_k)_{ij} \left(1 - x_{ij}\right) + \left[\sum_{(l,v)\in LV(i,j)} (D_k)_{lv} + (D_k)_{ij}\right] x_{ij}, \quad \forall k \in [m], (i,j) \in \mathcal{R}. \tag{9}$$

Constraints 8 and 9 attempt to establish lower and upper bounds on the amount of cargo to be delivered to a specific block in the absence/presence of a UMC that includes such block in its coverage area. If there are no UMCs around, constraint 8 states that each carrier needs to supply at least the demand of the block. On the other hand, if the block under consideration is designated as a UMC (constraint 9), the carrier might need to supply the demand of the block under consideration plus the demand of all the blocks in its coverage area.

$$t_{ijlv} \leq x_{lv}, \quad \forall (i,j) \in \mathcal{R}, (l,v) \in LV(i,j); \tag{10}$$

$$\sum_{(l,v)\in LV(i,j)} t_{ijlv} + t_{ijij} = 1, \quad \forall (i,j) \in \mathcal{R}; \tag{11}$$

$$t_{ijij} \leq 1 - \frac{1}{4} \sum_{(l,v)\in LV(i,j)} x_{lv}, \quad \forall (i,j) \in \mathcal{R}. \tag{12}$$

Constraints 10 through 12 establish the rules according to which the demand in some blocks can be delivered to adjacent blocks. For instance, constraint 10 establishes that a block's demand cannot be assigned to a different block unless this block is a UMC. Constraint 11 states that the demand of a specific block can only be assigned to be delivered to the same block or to an adjacent UMC, the latter case having priority according to constraint 12. The "1/4" factor in constraint 12 is motivated by the fact that in any optimal solution, there are at most four UMCs (out of the eight possibilities) surrounding a given block. Adjacent UMCs translate into overlapping coverage areas. The same coverage can be achieved (or even improved) by installing non-overlapping UMCs with one block of distance

between them, so adjacent UMCs are not optimal.

$$\sum_{s\in[T_k]} (y_{sk})_{ij} \leq \left\lceil \sum_{(l,v)\in LV(i,j)} (D_k)_{lv} + (D_k)_{ij} \right\rceil t_{ijij}, \quad \forall k\in[m],(i,j)\in\mathcal{R}; \tag{13}$$

$$\sum_{s\in[T_k]} (y_{sk})_{ij} \geq \left\lceil \sum_{(l,v)\in LV(i,j)} (D_k)_{lv} t_{lvij} + (D_k)_{ij} t_{ijij} \right\rceil - \left\lceil \sum_{(l,v)\in LV(i,j)} (D_k)_{lv} + (D_k)_{ij} \right\rceil (1-x_{ij}), \forall k\in[m],(i,j)\in\mathcal{R}. \tag{14}$$

Constraints 13 and 14 help refining the lower and upper bounds already established on the amount of cargo to be delivered at a specific block. For instance, constraint 13 establishes that no cargo is delivered to a block that is not self-assigned (t_{ijij} = 0). If a block is designated as a UMC, all the cargo corresponding to its own demand and the demand of all the blocks assigned to it, must be delivered at the UMC, as constraint 14 shows.

$$(y_{sk})_{ij} = (z_{sk})_{ij} = x_{ij} = t_{lvij} = 0, \ \forall s\in[T_k], k\in[m], (l,v)\in[M]\times[N], (i,j)\in[M]\times[N]\setminus\mathcal{R}. \tag{15}$$

Constraint 15 simply expresses that the imaginary ring of blocks added has no real implications in the problem, meaning that all the variables associated with the blocks in the imaginary ring take the value of zero. As it was mentioned before, the imaginary ring of blocks was considered to avoid presenting an excessively long formulation addressing particular constraint cases for blocks on the boundary of the study area. The introduction of this imaginary ring allows considering all the "real" blocks as interior blocks (not on the boundary).

2.5 | Complete formulation

It has to be noted that in order to solve a general bi-level problem, conditions to guarantee optimality for each sub-problem (each carrier's problem) would have to be expressed and included in the master problem to ensure optimal solutions for the agents in both levels. In the case of convex optimization, that can be achieved by introducing the Karush-Kuhn-Tucker (KKT) conditions of the subproblems into the master problem as additional constraints [8, 32]. However, the integer nature of some of the variables under analysis rule out any possibility of considering dual formulations, so the following features were exploited in order to achieve an optimal solution for the general bi-level problem, in a similar way to [24].

1. In a competitive market, as the freight transport operation has been described in the US [16], the price is equal to the marginal cost of the service provided, so in order to survive through the market dynamics, all carriers need to align their operational costs to the market price; meaning that the coefficients accompanying each carrier's objective, in equation 1, need to be equal ($\omega_k = \omega, \forall k$).
2. Besides, once the decisions about the location of UMCs ($\{x_{ij}\}$) and the assignment of the block demands to other blocks ($\{t_{lvij}\}$) have been made (defined as master decisions), each carrier has an independent problem of optimal allocation of their own resources. Given this fact and the previous one, the individual carrier's subproblems could be merged into a single subproblem that has as objective the summation of all the individuals objectives with the unique marginal cost, ω, multiplying that summation.
3. Under the scenario of only one subproblem for carriers, the master problem's objective (policy planner's) and the single subproblem's objective (carriers') differ only on a multiplicative constant that can be dropped out of the formulation. In that sense, the problem corresponds to the minimum (according to the policy planner) of another minimum (according to the carriers) for the same objective specification; for that reason, the problem can be

reduced to a single common-objective problem.

Let $P_k\left(\{x_{ij}\},\{t_{ijlv}\}\right)$ denote the feasible region for the subproblem of carrier k depending on the decisions of the policy planner. This feasible region is composed of constraints 3 to 6, 8 to 9, and 13 to 15 besides the nonnegativity and the binary constraints. After the considerations mentioned, the bi-level problem can be reformulated as a single problem labeled **LUMCP** (from **L**ocation of **U**rban **M**icro-consolidation **C**enter **P**roblem).

[LUMCP]:

$$\min \sum_{k\in[m]} \sum_{s\in[T_k]} L_{sk} \tag{16a}$$

$$s.t. \sum_{i\in[M]} \sum_{j\in[N]} x_{ij} \leq \theta, \tag{16b}$$

$$t_{ijlv} \leq x_{lv}, \qquad\qquad \forall(i,j)\in\mathcal{R}, (l,v)\in LV(i,j); \tag{16c}$$

$$\sum_{(l,v)\in LV(i,j)} t_{ijlv} + t_{ijij} = 1, \qquad\qquad \forall(i,j)\in\mathcal{R}; \tag{16d}$$

$$t_{ijij} \leq 1 - \frac{1}{4}\sum_{(l,v)\in LV(i,j)} x_{lv}, \qquad\qquad \forall(i,j)\in\mathcal{R}; \tag{16e}$$

$$x_{ij} = t_{lvij} = 0, \qquad \forall(l,v)\in[M]\times[N], (i,j)\in[M]\times[N]\setminus\mathcal{R}; \tag{16f}$$

$$x_{ij} \in \{0,1\}, \qquad\qquad \forall(i,j)\in\mathcal{R}; \tag{16g}$$

$$t_{lvij} \in \{0,1\}, \qquad\qquad \forall(l,v),(i,j)\in\mathcal{R}; \tag{16h}$$

$$\mathbf{y_k, z_k, L_k} \in P_k\left(\{x_{ij}\},\{t_{ijlv}\}\right), \qquad\qquad \forall k\in[m]. \tag{16i}$$

It has to be noted that each feasible region, P_k, is determined by two types of constraints: the ones related to each carrier's operations regardless of the UMC location decisions, and the ones dependent on the UMC location decisions. This split is exploited in developing a heuristic algorithm to more accurately solve the entire problem.

2.6 | Heuristic

Given the combinatorial structure of this mixed integer quadratically constrained problem (MIQCP) and the number of variables required to model the operation dynamics, the problem has an outstanding level of complexity even for modest-size layouts. The first attempts to solve this problem directly using `CPLEX` showed that the algorithms available to solve generic MIQCP instances could benefit from tailored strategies designed to boost their performance based on the nature and characteristics of the problem at hand. In this section, the motivation behind the heuristic designed to improve the performance of `CPLEX` for this problem is described as well as the details of the heuristic developed.

According to [6], a complicating variable is defined as one that once is given a value, the remaining problem can be easily solved in independent blocks of less size and complexity. For instance, it has been mentioned before that once the variables associated with the location of UMCs ($\{x_{ij}\}$), and the assignment of block demands to other blocks ($\{t_{lvij}\}$) are assigned a value, the remaining problem can be solved independently for each carrier since there are no other ties that make the carrier formulations dependent on each other.

When the problem involving complicating variables is linear, it can be split into a master problem and a surrogate problem that iteratively converge to the optimal solution of the original problem [6]. The main idea behind this

decomposition is to treat the complicating variables as parameters with a given value in the general problem (now surrogate problem) and try to determine these given values in a master problem. The master problem will update the complicating variables' values in the surrogate problem once a master solution has been obtained. The surrogate problem will be solved for the rest of the variables using the values of the complicating variables as given (fixed) and will also provide feedback to the master problem in the form of optimality cuts. These cuts are determined using dual variables of the surrogate problem. The convergence of this loop is guaranteed provided that the original problem is linear. Also, a first (arbitrary) value needs to be given to the complicating variables for the first iteration of the loop starting at the surrogate problem.

Unfortunately, even though **LUMCP** has an appealing structure for the complicating variable framework, it is neither linear nor continuous, ruling out the possibility of exploiting the algorithm aforementioned because of the absence of dual equivalent problems. A heuristic solution method was then proposed in an analogous way. The schematic structure of the heuristic is presented in Algorithm 1.

Algorithm 1: Heuristic to solve a **LUMCP** instance

Result: Feasible solution for **LUMCP**

Input: Parameters in Table 1 ;

$\theta_C := 0$;

$x_{ij}^{fixed} := 0, \forall (i,j) \in [M] \times [N]$;

while $\theta_C \leq \theta$ **do**

    1. Solve **LUMCP**$\left(\left\{ x_{ij} = x_{ij}^{fixed} \right\}, \theta_C \right)$, the surrogate problem with the UMC location decisions fixed and the maximum number of UMCs given by θ_C;

    2. From the optimal solution of **LUMCP**$\left(\left\{ x_{ij} = x_{ij}^{fixed} \right\}, \theta_C \right)$, compute the amount of stops around a given block that are not related to UMCs

$$v_{ij} := \sum_{(l,v) \in LV(i,j)} \left[\sum_{k \in [m]} \sum_{s \in [T_k]} (z_{sk}^*)_{lv} \cdot t_{lvlv}^* \cdot \left(1 - x_{lv}^{fixed} \right) \right];$$

    3. Locate the next UMC based on the values for $(v_{ij})_{i,j}$: $x_{gh}^{fixed} \leftarrow 1$ if

$$(g,h) \in \arg \max_{(i,j) \in \mathcal{R}} v_{ij} \quad \text{s.t.} \quad \sum_{(l,v) \in LV(i,j)} x_{lv}^{fixed} + x_{ij}^{fixed} = 0;$$

    4. $\theta_C \leftarrow \theta_C + 1$;

end

The heuristic only deals with determining the location of the UMCs and not with the block assignment variables because their interactions and impact on the objective are not as simple as the ones of the locations, especially when there are multiple UMCs installed. This implies that the surrogate problem, **LUMCP**$\left(\left\{ x_{ij} = x_{ij}^{fixed} \right\}, \theta_C \right)$, is not decomposable into k subproblems, one for each carrier. However, the problem gets significantly simpler if the location decisions are made separately.

This is a greedy heuristic in the sense that it builds up on previous results to locate additional UMCs. Although there is no guarantee of optimality for this heuristic, Section 3 provides evidence of its advantages over the brute-

force approach of simply letting the solver handle the entire problem and attempt to find a solution. The heuristic steps can be described as follows:

Step 0. An initial value of 0 for the location variables is provided. The counter for the number of UMCs located is set to 0 as well.

Step 1. Given as input the value for the location variables (either from **Step 0** or after **Step 4** if there are previous iterations), the surrogate problem for the carriers, **LUMCP**$\left(\left\{x_{ij} = x_{ij}^{fixed}\right\}, \theta_C\right)$, is solved using CPLEX or any other MIQCP solver.

Step 2. With the solution provided from **Step 1** for the optimal number and the location of the stops for each carrier, a matrix of cumulated stops is computed. The matrix gathers the total number of stops within the coverage area of a given block for all carriers. Only the stops related to blocks that have not yet been designated as UMCs or that are not already being serviced by UMCs are considered.

Step 3. The new UMC is located in the block with the maximum aggregate count (see **Step 2**) that does not have a UMC already at it or at any of the blocks in its coverage area. There might be multiple optima for this step. In any case, the algorithm chooses the first one it finds.

Step 4. The counter for the number of UMCs located is increased. Then, it is checked if the total number of UMCs is less than or equal to the maximum number of UMCs to be considered. If that is the case, another iteration of the heuristic is run (**Steps 1** through **4**). If not, the problem has reached its solution and the procedure ends.

The outcome of the heuristic routine is a feasible solution for the original MIQCP problem that defines an upper bound for the objective value. In Section 3, it is shown that, besides being feasible, this solution is usually close to the optimal one in terms of the objective value.

3 | RESULTS

This section presents a comparison of the performance for both the solution of **LUMCP** using a generic solver (which is referred to as **LUMCP**) and the solution via the implementation of the heuristic described in Algorithm 1.

First, both models were compared in a small setting (denominated as "toy example") for which **LUMCP** could be solved to optimality with CPLEX. This scenario allows to compare how the heuristic performs with respect to the original problem in terms of the proximity to the optimal value and computational efficiency.

In a second experiment, multiple instances were randomly generated from a stochastic mechanism to evaluate the performance of both solutions by means of statistical analyses given the size of the simulated sample. The stochastic structure generates more complex instances than the toy example, so optimality was not achieved within a reasonable amount of time. In order to still compare the performances of the competing alternatives, a time limit was established and the quality of the best solutions reached up to this limit was compared.

Finally, a case study inspired by the freight delivery patterns to households in Manhattan, NY was solved using the heuristic approach (given its good performance) and the interpretation and implications of the results are discussed from a policy planning perspective to show how promising this formulation, along with the heuristic algorithm, is in providing guidance for implementation of UMCs.

3.1 | Toy example

Assuming a layout of 30 blocks with two different carriers operating and a payload of 6 units per truck, the spatial distribution for the demand shown in Figure 4 was used. The general parameters for this example are presented in

Table 4.

(a) Demand for Carrier 1   (b) Demand for Carrier 2
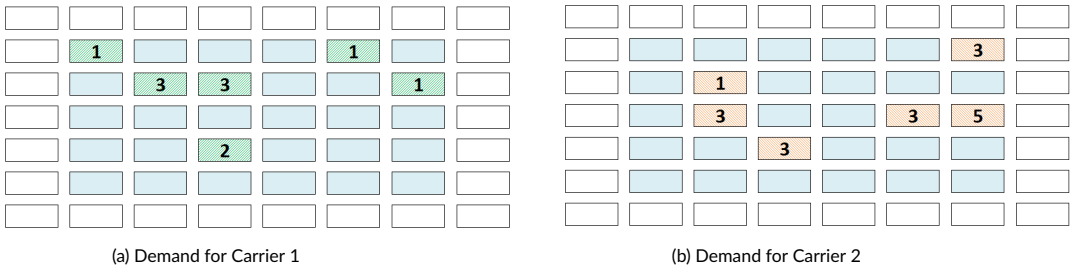
FIGURE 4   Spatial distribution of the demand (in units) for each carrier under consideration

TABLE 4   General parameters for the toy example

Parameter	Value
Layout dimensions	$M = 5$, $N = 6$
Number of carriers	$m = 2$
Beardwood's constant	$\phi = 0.765$
Payload per truck	$PL = 6$ units
Area	$A = M \times N$ units^2
Demand	D_k (see Figure 4)
Number of UMCs	$\theta \in \{0, 1, \ldots, 5\}$

The results for both the **LUMCP** and the heuristic are presented in Table 5. Given the simplicity of this illustrative example, the **LUMCP** was solved to optimality. The relative comparison of the gap between the **LUMCP** and the heuristic indicators with respect to the heuristic is also shown as a percentage.

TABLE 5   Presentation of the results for the toy example

	LUMCP		Heuristic		Relative gap w. r. t. Heuristic	
θ	Objective	Time (s)	objective	Time (s)	Objective (%)	Time (%)
0	3.437	1.578	3.437	0.781	0.00	102.05
1	3.198	3.828	3.198	1.266	0.00	202.37
2	2.991	6.485	2.991	1.531	0.00	323.58
3	2.991	5.797	2.991	1.797	0.00	222.59
4	2.991	5.469	2.991	1.875	0.00	191.68
5	2.991	5.000	2.991	1.938	0.00	157.99

In Table 5, it can be seen that, as the value of θ increases, the objective experiences a reduction until it stabilizes

(at θ = 2) for both the model and the heuristic. The computation time increases significantly for the **LUMCP**, as a result of the problem being $\mathcal{NP}$-hard. The heuristic, on the other hand, manages to achieve lower computation times for all the scenarios considered, reaching the same optimal value of the **LUMCP** two times faster in the worst case (**LUMCP** taking 102.05% more than the heuristic), and four times faster in the best case (**LUMCP** taking 323.58% more than the heuristic). A graphical comparison can be seen in Figure 5a.

The gap between the objective without UMCs (θ = 0) and with UMCs ($\theta \geq 1$) is crucial in the decision-making process, especially, from the policy planner perspective; because it allows to conclude how sustainable in terms of cost (when cost valuations are available) the installation of a given UMC is, by contrasting it with the marginal benefit it provides to the system. Figure 5b compares the value of the objective with the marginal reduction obtained after installing each additional UMC. A more detailed analysis of the economic implications of the results will be presented in the Subsection 3.3.
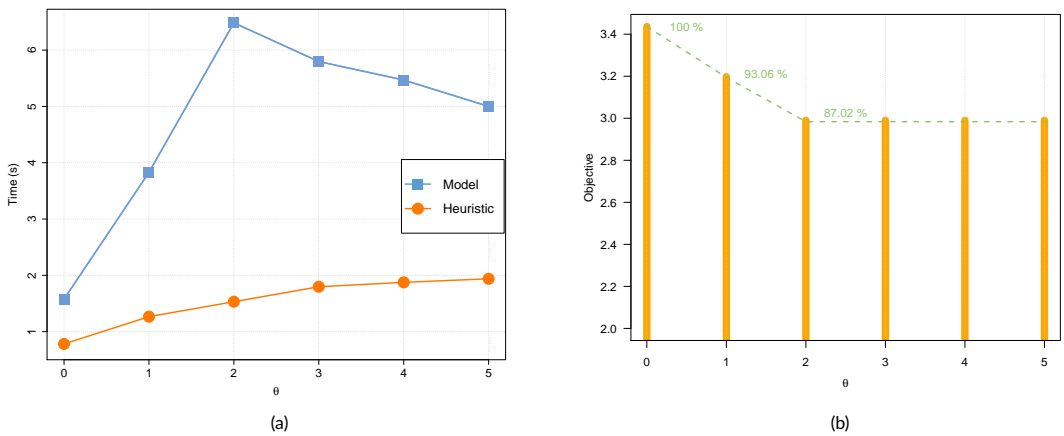

FIGURE 5   (a) Comparison of computation times for the heuristic and the model, (b) Representation of the objective and the marginal benefit of each UMC expressed as the reduction in the objective for the baseline scenario (θ = 0)

3.2 | Simulated instances

An expanded example of 100 blocks with three different carriers and payload of 6 units was created to test the performance of the heuristic with respect to the **LUMCP** by means of statistical tools. For the purpose of comparison of both solution methods, 50 instances were simulated by randomly assigning positive demands to approximately 10% of the blocks, i.e., 10 blocks, within the proposed layout; and assuming null demand values for the rest of the blocks. The block selection was done by simulating a Bernoulli outcome for each block with probability of success equal to 0.1.

The demand for the selected blocks was simulated as the realization of a uniformly-distributed random variable over the interval $[0, PL/2]$. Table 6 shows the parameters used in these simulated instances which were run on a computer with Intel Core i7-6700 CPU at 3.40 GHz and 16.0 GB of RAM. A time limit of six hours per instance was set (corresponding to one hour per each value of θ).

TABLE 6   General parameters for the simulation experiment

Parameter	Value
Layout dimensions	$M = 10$, $N = 10$
Number of carriers	$m = 3$
Beardwood's constant	$\phi = 0.765$
Area	$A = 0.022 \times M \times N$ km^2
Payload per truck	$PL = 6$ units
Number of UMCs	$\theta \in \{0, 1, \ldots, 5\}$
Demand	$(D_k)_{i,j} = V_{i,j} \cdot W_{i,j}, \{V_{i,j}\} \overset{\text{i.i.d.}}{\sim} U[0, PL/2] \perp \{W_{i,j}\} \overset{\text{i.i.d.}}{\sim} Ber(\pi = 0.1); (i, j) \in \mathcal{R}$
Number of instances	50
Time limit	6 h/instance

The summary of the results for the **LUMCP** and the heuristic is shown in Tables 7 and 8. In Table 7, the summary statistics of the run times (in hours) for the 50 instances are reported. The average total run time per instance (defined as the time required to obtain a solution for all the values of θ considered) was estimated to be 5.88 hours for the **LUMCP**, compared to 1.04 hours for the heuristic. Most of the instances for the **LUMCP** were not solved to optimality, as they were terminated by the imposed time limit. This proves the computational complexity of solving the **LUMCP** directly, as previously mentioned. All instances solved by the heuristic were solved without reaching the time limit, as the maximum time taken by the heuristic in one instance was 3.76 hours.

It should also be noted that the heuristic was faster (in average and median terms) in obtaining a solution for a given θ in all instances, except at $\theta = 5$. This is due to the greedy nature of the heuristic, where the running time is cumulative and adds in the running time of previous iterations. This means that the average run time of the heuristic at $\theta = 5$ of 1.04 hours is the same as the average of the total run time for a given instance, since the maximum value of θ is five. This number is higher than the average run time of the **LUMCP** as, for each θ in the **LUMCP** context, the run starts from the beginning with no record of the previous UMC locations.

TABLE 7   Summary statistics of running time for the sample of instances

θ	LUMCP time (h)				Heuristic time (h)			
	Minimum	Maximum	Median	Average	Minimum	Maximum	Median	Average
0	2×10^{-4}	1.00	1.00	0.94	1×10^{-4}	1.00	0.34	0.27
1	0.002	1.00	1.00	0.98	2×10^{-4}	1.72	0.67	0.50
2	0.019	1.00	1.00	0.98	3×10^{-4}	2.41	0.75	0.69
3	0.107	1.00	1.00	0.98	5×10^{-4}	3.08	0.77	0.83
4	1.00	1.00	1.00	1.00	6×10^{-4}	3.42	0.81	0.94
5	1.00	1.00	1.00	1.00	7×10^{-4}	3.76	0.82	1.04

Despite this fact, the heuristic still proves its superiority in terms of running time because the **LUMCP** would have to be run for all possible values of θ to select the most convenient number of UMCs in the study area. This leads to much less running time per instance in case of using the heuristic. In this experiment, this running time would be

17.6% of the one of the **LUMCP** per instance, in average.

Table 8 contrasts the relative objective gap percentage (substracting the **LUMCP**'s objective from the heuristic's objective) with respect to the **LUMCP**. The largest gap where the heuristic outperformed the **LUMCP** was -3.02%, while it was 6.96% for the opposite case. The overall average gap is 0.77%, and there is an increase in the value of the gap as values of θ increase. At $\theta = 0$, the heuristic's objective is always either better or at least as good as that of the **LUMCP**, however, that advantage reduces when the number of UMCs increases. This is clearly exhibited in Figure 6, that shows the percentage of instances for which the heuristic outperformed or performed equally well as the **LUMCP**.

TABLE 8 Summary statistics of objective gap with respect to LUMCP

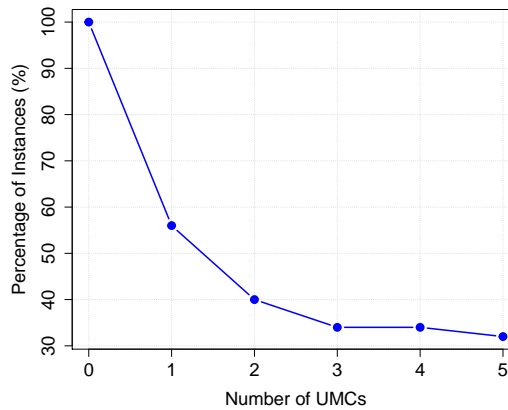| θ	Objective gap (%) w. r. t. LUMCP				
	Minimum	Maximum	Median	Average	p-value
0	-0.25	0.00	0.00	-0.02	0.031[†] (*)
1	-1.08	2.11	0.00	0.30	0.755
2	-1.93	3.99	0.40	0.64	0.079
3	-2.86	4.66	0.78	0.81	0.013 (*)
4	-2.38	5.42	1.21	1.21	0.008 (**)
5	-3.02	6.96	1.64	1.68	0.002 (**)
All	-3.02	6.96	0.11	0.77	1×10^{-5} (***)

†: the p-value is unreliable due to multiple ties

FIGURE 6 Percentage of instances where the heuristic achieved an objective less than or equal to the **LUMCP** model

The decline in the performance of the heuristic with the increase of UMCs could be attributed to the cumulative nature of the greedy heuristic in which, for adding another UMC, the location of the previous ones is taken into account, i.e., the heuristic builds upon the solution found at the previous value of θ.

Table 8 also shows the p-values for the one-sample sign test. This nonparametric tool allows to test whether or not the median of a given population is equal to a given value, based on the evidence provided by the data. The sign test was preferred over a classic t-test because the distribution of the data is far from being Gaussian (see Figures 7 and 8), and the sample is of moderate size, casting doubt on the appropriateness of the Central Limit Theorem for this case.

The purpose of this test was to generalize, based on the data at hand, if the results for the population of instances (characterized by the stochastic simulation algorithm already described) favored the heuristic or the original model in terms of their accuracy to solve the mathematical problem. This question translates to the following hypothesis system

$$\begin{cases} H_0 : \text{The median of the relative gaps is equal to zero,} \\ \text{v.s.} \\ H_1 : \text{The median of the relative gaps is different from zero.} \end{cases}$$

Given a significance level of 5%, rejecting the null hypothesis means that the data support the idea that one of the two competing models (the heuristic of the **LUMCP**) performs better in terms of median performance. The results of the sign test show that for $\theta \in \{1, 2\}$, the null hypothesis could not be rejected, while it is certainly rejected for the rest. Not being able to reject the null hypothesis in these two cases provides evidence in favor of the heuristic because it implies that the heuristic can achieve the same level of median performance than the original model using less computational resources, as it was shown previously.

Although the p-value for the test when $\theta = 0$ is unreliable, all the corresponding values in the sample are either negative or zero showing that the heuristic performs as well as the model in the worst case, and demonstrating the superiority of the heuristic in this scenario too. On the other hand for the cases when $\theta \geq 3$, it can be seen in Figures 6-8 that the evidence favors the **LUMCP** model. In general, the performance level of the heuristic decreases as long as θ increases, however, its value exceeds the **LUMCP**'s objective value by at most 7%.

The histogram of the relative objective gap with respect to the **LUMCP** for all values of θ (from zero to five) is shown in Figure 7. Furthermore, to understand in detail the performance of the models and the role of the variability, an estimation of the density for the relative gaps as a function of θ is presented in Figure 8. This figure summarizes the findings from the simulated instances in the sense that it shows how (1) the performance level of the heuristic (in median terms) decreases as long as θ increases, (2) the variability and the range of the values for the objective gap also widens when θ increases, and (3) how the heuristic in general provided results within a 7% tolerance with respect to **LUMCP**.

This experiment of multiple simulated instances showed that the heuristic is faster than the **LUMCP** in providing a solution, even with instances larger than the toy example. For this case, it also showed that the heuristic could obtain a solution that is better, at least as good or with a reasonably small percentage of inaccuracy, especially for small values of θ. The reasons mentioned above highlight the advantages the heuristic has in handling large instances in terms of both time efficiency and solution quality, although the heuristic's limitations should be considered as the number of UMCs in the study area increase.
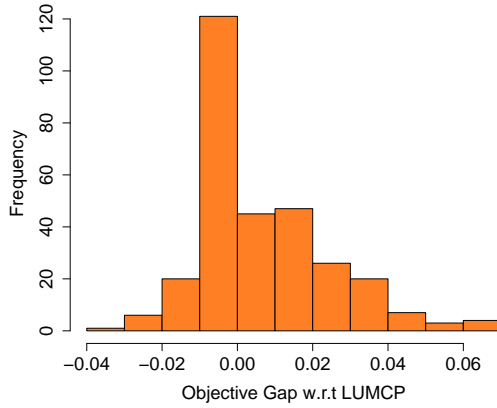
FIGURE 7 Percentage of instances where the heuristic achieved an objective less than or equal to the **LUMCP** model
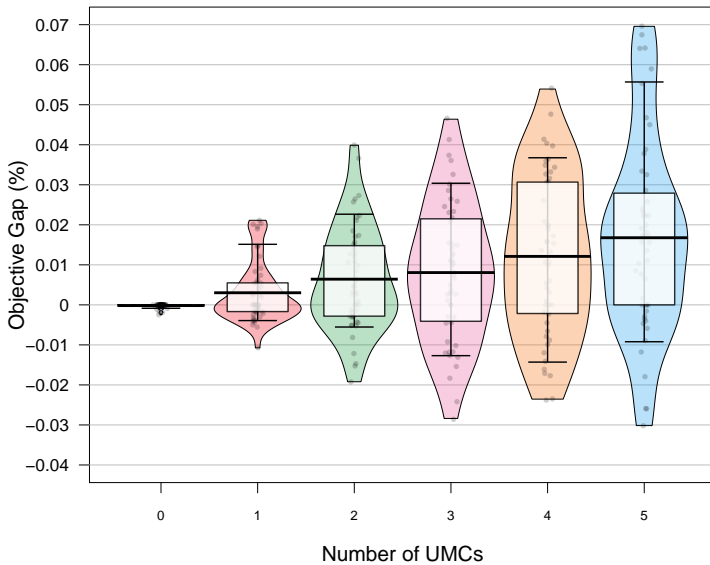

FIGURE 8 Percentage of instances where the heuristic achieved an objective less than or equal to the **LUMCP** model

3.3 | Case study

3.3.1 | Data preparation

The authors decided to do a case study in Midtown Manhattan to quantify the impact that (one or several) UMCs would have in the last-mile logistical operations in New York City. The quantification of impacts is measured in terms of social costs: private costs and externalities. The geographic area in which the problem is inspired locates between

Fifth and East Avenues, and between 76th and 87th streets (a 10×8 grid of blocks), as shown in Figure 1a. This location was selected because census tract data of e-commerce residential packages per day are available [5].

In addition to the e-commerce data, the city of New York has a robust geo-database, PLUTO, which combines tax lot data and features from the Department of Finance's Digital Tax Map (DTM). The database contains over 50 attributes regarding land use characteristics and intensity of usage, i.e., residential units by lot or density of commercial establishments [21]. Combining these two datasets provides the required information on the daily residential e-commerce demand of 30 out of the 80 blocks of the proposed area. From the 30 blocks where information was available, a ratio of number of packages per day per each residential unit was computed. The ratio showed that every residential unit in Manhattan produces an average of 0.92 packages per day. The demand for the remaining 50 blocks was imputed proportionally using this ratio. It should be noted that each block in this 10×8 grid has an average dimension of 80 meters by 274 meters, which leads to an extremely small coverage area for each UMC when defined as one block away in each direction of the UMC. Since an electric cargo bike is estimated to be able to travel two kilometers [25], then, without loss of generality; both the area and, by default, the demand can be blown up by a factor of nine without affecting the validity of the model. This expansion creates a coverage radius, for each UMC, of 1.2 km lengthwise and 0.36 km widthwise (neglecting the road widths), which gives a conservative estimate of the distance traveled by bikes (to consider regular bikes as well) for the last-mile operation.

The following step was to transform the total number of packages per day to a weight measure to quantify the total number of trucks needed to do the deliveries. Leveraging the results of an online survey developed in France (conducted to 108 e-commerce shippers) in 2016, the demand for packages by block was transformed to weight of packages per block in kilograms by using the mean of the probability distribution reported [26].

Analogously to the actual market dynamics in which three companies control over 95% of the market share on delivering urban packages in the US [9], the case study considers only three companies acting as carriers. This market share power has tremendous implications in their logistics operations, since it allows carriers to consolidate cargo, optimize routing and have more decision power in their supply chain operations. As a consequence of that, it was assumed that the three carriers employ trucks with a reference payload of 20 t out of an installed capacity of 34 t, as a compromise between their ability to consolidate given by the market power and the high level of service (characterized mainly by fast deliveries) demanded by online shoppers [23].

The objective of the case study was to conduct a social cost-benefit analysis of whether using UMCs or not as a freight initiative to improve urban logistics. The social cost is the combination of private costs and any other external costs faced by society derived as a byproduct of the production and consumption dynamics [12]. Regarding the private costs, the carriers' operational costs that are considered are: driver salaries and benefits, fuel, lease and truck payments, repairs and maintenance and other costs including vehicle insurance, permits, tolls and tires which in total adds to 1.14 USD/km in the northeast part of the US [1]. In relation to the external costs that deliveries bring to society, pollutants by Vehicle-Miles Travelled (VMT) are considered. In fact, [18] presents the findings of three cases studies of off-hour deliveries (Bogota, Colombia; New York City, USA; and Sao Paulo, Brazil), along with a comparison of emissions when deliveries are done at regular hours versus off-hours. That research quantified emission factors per kilometer for Carbon Monoxide (CO), Carbon Dioxide (CO_2) and Oxides of Nitrogen (NO_x). These emission factors are used in the case study to quantify the total amount of pollutants emitted by all truck tours. Additionally, monetary valuation of these pollutants has been done by the Federal Highway Administration and the Environmental Protection Agency. A metric ton of CO, CO_2 and NO_x is valuated in 5,400 USD, 42 USD and 20,000 USD, respectively [10, 11].

Due to limitations in memory, a limit of three hours per each UMC was set for the heuristic to provide a solution, totalling 48 hours of computation capacity for the heuristic to locate the 15 UMCs. All the surrogate problems involved in the heuristic algorithm reached the time limit established, meaning that the entire 48 hours were utilized to reach

the solution presented in the following Subsection. The parameters used for the case study are summarized in Table 9.

TABLE 9   Specification of parameters for the case study

Parameter	Value
Layout dimensions	$M = 10$, $N = 8$
Number of carriers	$m = 3$
Beardwood's constant	$\phi = 0.765$
Area	$A = 1.775 \times M \times N$ km^2
Payload per truck	$PL = 20$ t
Number of UMCs	$\theta \in \{0, 1, \ldots, 15\}$
Operational private cost	$\omega = 1.13$ USD/km
Emission factor of CO	3.63×10^{-6} t/km
Emission factor of CO_2	1.59×10^{-4} t/km
Emission factor of NO_x	5.84×10^{-7} t/km
Valuation of CO	$\lambda_{CO} = 5,400$ USD/t
Valuation of CO_2	$\lambda_{CO_2} = 42$ USD/t
Valuation of NO_X	$\lambda_{CO} = 20,000$ USD/t

3.3.2 | Findings

Results from the heuristic highlight not only the reduction of VMTs in the network, but also the marginal benefit each UMC brings. For the purpose of the heuristic, it was determined that up to 15 UMCs would be tested in the selected grid because (1) a stabilizing point could be identified, after which, having an additional UMC would not impact the objective function significantly; (2) marginal benefit analysis could be done to compute the explicit reduction on VMT by adding an extra UMC. Figure 9a illustrates the decreasing pattern of the VMTs as a function of the number of UMCs. It also shows that the stabilizing point corresponds in this case to 10 UMCs, with a total reduction of circa 52% with respect to the base case (no UMCs). A total amount of 10 UMCs has the potential of reducing around 150 VMTs per day. It must be noted that if more computational time and memory were available, the solution obtained from the heuristic might have exhibited better values for the objective. In any case, the results presented must be interpreted as a conservative outcome in the sense that the impact of increasing the number of UMCs might have a more dramatic decay than the one shown in Figure 9a.

Given the results in terms of daily VMTs, an extrapolation to a yearly basis projection of reductions of the social cost was conducted. Figure 9b exhibits the reduction in social cost as UMCs increase in the study area. A significant potential reduction in the social cost is observed (more than 100,000 USD per year) when 10 UMCs are allocated. However, around 50% of that reduction is achieved only with the installation of the first UMC, as it can be seen in Figure 9c.
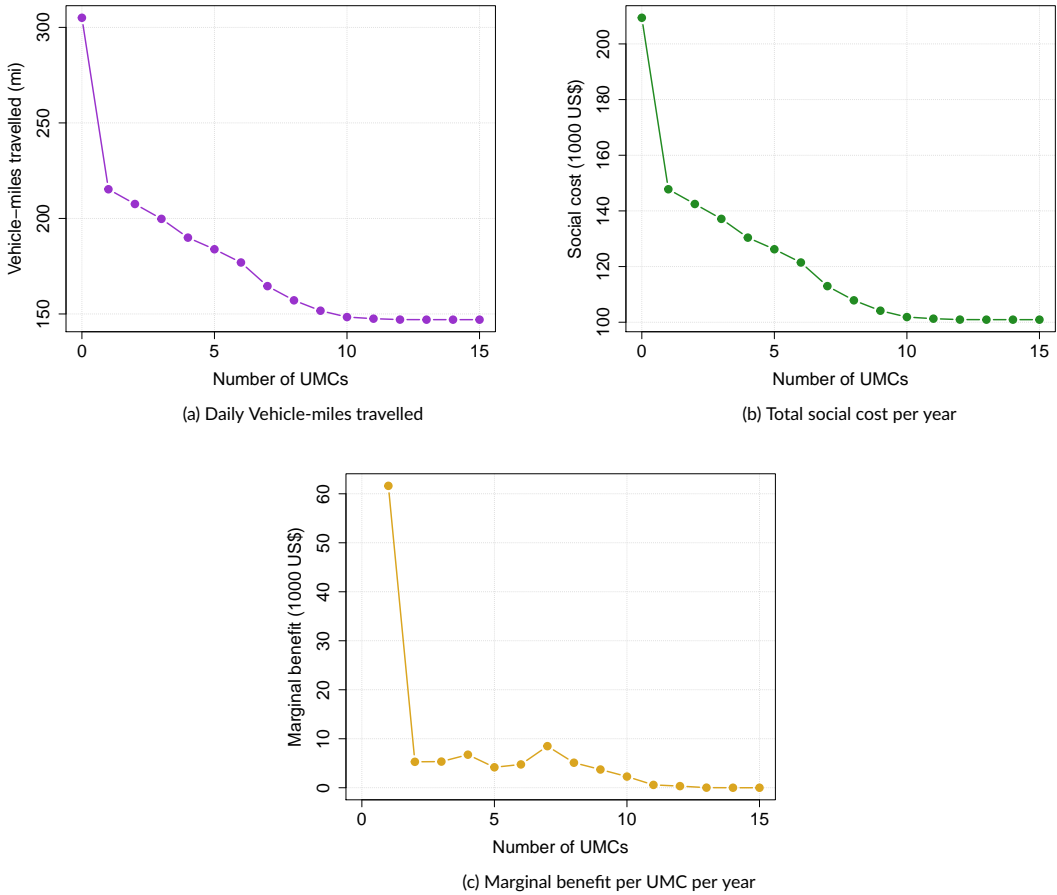
(a) Daily Vehicle-miles travelled

(b) Total social cost per year

(c) Marginal benefit per UMC per year

FIGURE 9 Visualization of the results for the case study

3.3.3 | Implications for policy planning

Given the fact that the characteristics and the corresponding expenses associated with the installation of UMCs are very dependent on the context (the technologies available, the urban design plans, the willingness of the private carriers to collaborate, etcetera), it is impossible to determine an optimal number of UMCs to be installed based only on the potential benefits in terms of social cost that they might generate.

Nevertheless, the marginal benefit analysis does suggest an upper bound on the investment budget for UMC installation in order for the initiative to be financially sustainable. For instance, the first UMC installed in this case study brings a monetary benefit to the system of (at least) 57,000 USD per year, which translates in a budget limit for its installation of that same amount. If the cost analysis reveals that this UMC will cost more than 57,000 USD a year, then this initiative would not be financially sustainable. When considered together, the first two UMCs amount a total benefit of (at least) 62,000 USD per year, which establishes a budget of 31,000 USD for the annual installation and maintenance of each. These two examples show that the highest benefits are achieved by the first UMCs and at

some point, the installation of an additional UMC brings negligible benefits that might not compensate its cost. For this policy to be self-sustainable, only a few UMCs will be actually set in place, unless this measure is implemented jointly with taxation initiatives or public incentives.

4 | CONCLUSIONS

This paper constitutes a novel contribution to the literature because it is the first effort addressed towards (1) formulating the location of Urban micro-consolidation Centers as a mathematical program, and (2) providing a time-efficient heuristic method that can reach a good feasible solution by effectively tackling the particular features of the problem under consideration.

The formulation described represents, in an accurate way, the interactions between carriers and policy planners when joint initiatives are undertaken. Although innovative, this formulation is also the result of gathering input from various experts and previous research efforts in the area, whose findings materialize in a sound and realistic representation of the last-mile delivery operations and the impact on sustainability and efficiency of installing UMCs. In spite of being based on an approximation of the length for the truck tours, its outcome provides the location of the UMCs that minimizes (indirectly) the number of stops, since that is the only quantity impacting the tour lenghts. The formulation is also general enough to contemplate different specifications for the UMCs depending on the context to be analyzed. The economic analysis presented focuses on the benefits of reducing the use of fossil fuel delivery vehicles, so it can also be adapted to conduct the cost-benefit analysis under the given specifications.

The greedy heuristic proposed to solve the **LUMCP** is a second major outcome of this research effort. Its performance was tested under multiple conditions showing in general the heuristic's competitiveness with respect to the exact solution method. The results showed that, in median and average terms, the heuristic was faster in reaching a solution in all instances; however, the relative running times increase with the number of UMCs to be considered, as a result of its greedy nature. The heuristic also proves its capability of providing a solution of fair quality, revealing under which conditions the solution obtained tends to exceed the quality of the **LUMCP**'s solution. In any case, even in the scenarios where the heuristic is outperformed, the error margins lie below 7% (with an average difference of only 0.77%).

The results of the case study clearly show the potential that UMCs have as an initiative to increase the sustainability of urban logistics. It should be evident, at this point, how policy planners and private carriers should be interested in pursuing a public-private partnership to implement this type of initiative, since this collaborative effort would result in a urban freight policy with the potential of reducing the private costs for carriers and improving the air quality for society. By contrasting the results of the marginal benefit analysis with the number of UMCs used, it can be determined, on a case-by-case basis; the economic viability and environmental impacts that the initiative brings in order to be self-sustainable. Furthermore, e-commerce is still at its initial stage, hence, a significant growth of the demand will be experienced. Besides, the scope of this analysis can be extended to account for, at least partially, business-to-business deliveries in addition to the deliveries to residential units, to magnify the benefits obtained.

The authors acknowledge the limitations of both the problem formulation and the heuristic solution method, which offer an opportunity for future research efforts. The limitations in the problem formulation include (1) the geometric assumptions of the study area and the coverage area of a UMC, (2) the treatment given to the embedded vehicle routing problems by means of an approximate solution, and (3) the technical definition and characterization of a UMC to fulfill the tasks considered and its impact on the problem formulation. As for the heuristic, the greedy and cumulative nature of the algorithm that builds up on previous solutions is a limitation, causing an impact in its

performance with the increase of UMCs.

References

[1] American Transportation Research Institute - ATRI, *An Analysis of the Operational Costs of Trucking: 2018 Update*, Tech. report, 2018.

[2] L. Baker, 2020. New york city hit ups with $23m in parking fines in 2019. Available at `https://www.freightwaves.com/news/ups-hit-with-22m-in-nyc-parking-fines`. Last accessed 21 March 2020.

[3] J. Beardwood, J.H. Halton, and J.M. Hammersley, *The shortest path through many points*, Math. Proc. Cambridge Philosophical Soc. **55** (1959), 299–327.

[4] K.M. Borland, 2019. Ecommerce will continue to shape retail for years. Available at `https://www.globest.com/2019/12/23/ecommerce-will-continue-to-shape-retail-for-years/?slreturn=20200222223715`. Last accessed 21 March 2020.

[5] Q. Chen and A. Conway, *Commercial vehicle parking availability and behavior for residential delivery in new york city*, Transportation Research Board 95th Annual Meeting, 2016.

[6] A.J. Conejo, E. Castillo, R. Minguez, and R. Garcia-Bertrand, *Decomposition techniques in mathematical programming: engineering and science applications*, Springer Science & Business Media, 2006.

[7] A. Conway, P.E. Fatisson, P. Eickemeyer, J. Cheng, and D. Peters, *Urban micro-consolidation and last mile goods delivery by freight-tricycle in manhattan: Opportunities and challenges*, Conference proceedings, Transportation Research Board 91st Annual Meeting, 2012.

[8] S. Dempe and A.B. Zemkoho, *The bilevel programming problem: reformulations, constraint qualifications and optimality conditions*, Math. Program. **138** (2013), 447–473.

[9] S. Dolan, 2019. Here is a list of last mile delivery companies and startups fighting to ship items same day for less. Available at `https://www.businessinsider.com/last-mile-delivery-companies-startups`. Last accessed 28 Oct 2019.

[10] Environmental Protection Agency, *Social Cost of Carbon*, Tech. report, 2016.

[11] Federal Highway Administration, *Cost Effectiveness Tables Summary*, Tech. report, 2017.

[12] B.C. Field and M.K. Field, *Environmental economics: an introduction*, Sustainable Human Development Review **105** (1997).

[13] M.A. Figliozzi, *Planning approximations to the average length of vehicle routing problems with varying customer demands and routing constraints*, Transp. Res. Record **2089** (2008), 1–8.

[14] M. Haag and W. Hu, *1.5 million packages a day: The internet brings chaos to n.y. streets*, New York Times (Oct 2019), Available at `https://www.nytimes.com/2019/10/27/nyregion/nyc-amazon-delivery.html`. Last accessed 21 March 2020.

[15] Hamburg News, 2018. Intelligent last mile solutions for hamburg. Available at `https://www.hamburg-news.hamburg/en/port-logistics/intelligent-last-mile-solutions/`. Last accessed 21 March 2020.

[16] J. Holguín-Veras, *Necessary conditions for off-hour deliveries and the effectiveness of urban freight road pricing and alternative financial policies in competitive markets*, Transp. Res. Part A: Pol. Practice **42** (2008), 392–413.

[17] J. Holguín-Veras and M. Brom, *Trucking costs in a metropolitan area: a comparison of alternative estimation approaches*, Transportation Research Board Annual Meeting, 2008.

[18] J. Holguín-Veras, T. Encarnación, C.A. González-Calderón, J. Winebrake, C. Wang, S. Kyle, N. Herazo-Padilla, L. Kalahasthi, W. Adarme, V. Cantillo, et al., *Direct impacts of off-hour deliveries on urban freight emissions*, Transp. Res. Part D: Transport Environment **61** (2018), 84–103.

[19] S.N. Kumar and R. Panneerselvam, *A survey on the vehicle routing problem and its variants*, Sci. Res. Publishing (2012).

[20] L. Mallozzi, E. D'Amato, and P.M. Pardalos, *Spatial Interaction Models*, Springer, 2017.

[21] NYC Planning, 2019. Pluto and map pluto. Last accessed 1 July 2019.

[22] T. Parr, 2017. Rippl no. 31: Ups – reducing vehicle movements with city centre container hubs. Available at `http://www.rippl.bike/en/rippl-31-ups-reducing-vehicle-movements-with-city-centre-container-hubs/`. Last accessed 21 March 2020.

[23] G. Radu and T. Miller, 2019. Why are overweight containers and misdeclared weights a problem?. Available at `https://www.ukpandi.com/knowledge-publications/article/overweight-container-guide-28/`. Last accessed 2 February 2020.

[24] G.K. Saharidis and M.G. Ierapetritou, *Resolution method for mixed integer bi-level linear problems based on decomposition technique*, J. Global Optim. **44** (2009), 29–51.

[25] M. Sheth, P. Butrina, A. Goodchild, and E. McCormack, *Measuring delivery route cost trade-offs between electric-assist cargo bicycles and delivery trucks in dense urban areas*, Eur. Transport Res. Review **11** (2019), 11.

[26] Statistica, 2019. Average weight of parcels sent for export by french e-merchants in 2016. Available at `https://www.statista.com/statistics/771219/e-trade-weight-way-parcel-sent-at-l-export-france/`. Last accessed 09 September 2019.

[27] The Alliance for Logistics Innovation through Collaboration in Europe and The European Road Transport Advisory Council, *Urban Freight - Research and Innovation Roadmap*, Tech. report, 01 2015, Available at `http://www.etp-logistics.eu/wp-content/uploads/2015/08/W56mayo-kopie.pdf`.

[28] F. Twisse, 2018. Micro depots in meisengasse, frankfurt am main. Available at `https://www.eltis.org/discover/case-studies/micro-depots-meisengasse-frankfurt-am-main`. Last accessed 21 March 2020.

[29] United Nations, Department of Economic and Social Affairs, Population Division, *World Urbanization Prospects: The 2014 Revision*, Tech. report, 2015, Available at `https://esa.un.org/unpd/wup/`.

[30] U.S. Census Bureau, 2020. Estimated quarterly u.s. retail sales. Available at `https://www.census.gov/retail/index.html`. Last accessed 21 March 2020.

[31] U.S. Census Bureau, 2020. Quarterly retail e-commerce sales. Available at `https://www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf`. Last accessed 21 March 2020.

[32] J.J. Ye, *Constraint qualifications and kkt conditions for bilevel programming problems*, Math. Oper. Res. **31** (2006), 811–824.

[33] A. Zaleski, 2017. Cities seek deliverance from the e-commerce boom. Available at `https://www.citylab.com/transportation/2017/04/cities-seek-deliverance-from-the-e-commerce-boom/523671/`. Last accessed 21 March 2020.