

A simulation-optimization approach for the management of the on-demand parcel delivery in sharing economy

*Original*

A simulation-optimization approach for the management of the on-demand parcel delivery in sharing economy / Perboli, Guido; Rosano, Mariangela; Wei, Qu. - In: IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS. - ISSN 1524-9050. - STAMPA. - 23:8(2022), pp. 10570-10582. [10.1109/tits.2021.3094851]

*Availability:*

This version is available at: 11583/2846123 since: 2021-08-25T08:32:13Z

*Publisher:*

IEEE

*Published*

DOI:10.1109/tits.2021.3094851

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

IEEE postprint/Author's Accepted Manuscript

©2022 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

# A simulation-optimization approach for the management of the on-demand parcel delivery in sharing economy

Guido Perboli, Mariangela Rosano and Qu Wei

**Abstract**—This paper investigates a dynamic and stochastic vehicle routing problem with time windows that considers the use of multiple delivery options and crowd drivers, reflecting the synchromodality in the urban context. We propose a multi-stage stochastic model, and we solve the problem by using a simulation-optimization strategy. It relies on a Monte Carlo simulation and a large neighborhood search (LNS) heuristic for optimization. We conduct a case study in the medium-sized city of Turin (Italy) to measure the potential impact of integrating cargo bikes and crowd drivers in parcel delivery. Experimental results show that combining crowd drivers and green carriers with the traditional van to manage the parcel delivery is beneficial in terms of economic and environmental cost-saving, while the operational efficiency decreases. Besides, the green carriers and crowd drivers are promising delivery options to deal with online customer requests in the context of stochastic and dynamic parcel delivery. The resulting set of policies are part of the outcomes of the Logistics and Mobility Plan 2019-2021 in the Piedmont region.

**Index Terms**—Crowdsourcing, Parcel delivery, Stochastic and dynamic VRPTW.

## I. INTRODUCTION

URBAN logistics aims to find efficient and effective approaches to move freights in urban areas while considering the negative impacts on congestion, the environment, and safety. In recent years, the urban population growth, the rapid boom in e-commerce, the desire for efficiency in supply chains, and the rise of the sharing economy lead to new opportunities and challenges for urban logistics. Pushed by the increasing on-demand urban delivery service, many negative impacts are generated in the city, including traffic congestion, air pollution, greenhouse gas emissions, and noise disturbance. On average, traditional vans discharge 16-50% of total vehicle emissions in the city area [1]. The use of new delivery options emerges to fulfill the on-demand requests for parcel delivery while reducing its negative impact.

Moreover, offering fast and cheap delivery has become an expectation for online customers and a challenge for logistics

companies. In this direction, companies started to adopt new business models based on more efficient and sustainable delivery options as cargo bikes, lockers, and mobile depots. In particular, crowdsourcing or the so-called “Uberization of the last mile” is an emerging application for parcel delivery that outsources the parcels to crowd drivers [2, 3]. They are a group of local and non-professional drivers who are willing to temporarily work for delivery companies and provide their assets (e.g., the vehicle) to perform the parcel delivery [4]. A certain amount of money, named compensation, is rewarded after completing the pickup or delivery tasks [5]. Crowd-drivers are required to have the driving ability and manage quickly the parcel delivery sometimes in less than an hour. The advantages of crowdsourcing are lower operation costs, higher flexibility, and lower emissions than traditional delivery options [6]. Indeed, it is a digital-driver business model with own-asset (i.e., the crowd-drivers bring their vehicles and provide for their maintenance), paperless operations reducing the overall costs and making the service attractive to the online customers. Besides, in real-world parcel delivery applications, customer demand, locations, and other attributes are usually unknown beforehand or known only probabilistically. The stochastic information on customers’ attributes becomes increasingly important, given its impact on the activities at the operational level.

In this paper, we address the dynamic and stochastic features of parcel delivery. We consider multiple delivery options and crowd drivers as sources of delivery capacity. The potential benefits generated by addressing the dynamic or stochastic contexts in parcel delivery are related to the increase of solution quality, the cost, or travel distance saving [7, 8].

The contribution of this paper is three-fold. First, we formulate a multi-stage stochastic model to capture the stochastic elements that arise in parcel delivery, namely the Dynamic and Stochastic Vehicle Routing Problem with Time Windows (DS-VRPTW). We thus, consider the uncertainty of some attributes and the possibility that some requests appear during the day, requiring adjustments in the delivery plan.

This research was funded by Politecnico di Torino under the interdepartmental Center for Automotive Research and Sustainable mobility (CARS) initiative. (*Corresponding author: Qu Wei*)

Guido Perboli is with the ICT for City Logistics and Enterprises Center, Politecnico di Torino, 10129 Turin, Italy; CARS@Polito, Politecnico di Torino, 10129 Turin, Italy; CIRRELT, Montreal, QC, Canada (e-mail: guido.perboli@polito.it).

Mariangela Rosano is with the ICT for City Logistics and Enterprises Center, Politecnico di Torino, 10129 Turin, Italy; CARS@Polito, Politecnico di Torino, 10129 Turin, Italy; (e-mail: mariangela.rosano@polito.it).

Qu Wei is with the ICT for City Logistics and Enterprises Center, Politecnico di Torino, 10129 Turin, Italy; (e-mail: qu.wei@polito.it).

Second, we introduce a simulation-optimization framework to solve the DS-VRPTW. The simulation can create realistic instances from real data, letting to simulate different policies and scenarios. We solve the problem using a Large Neighborhood Search (LNS) metaheuristic incorporating several solution improvement procedures. The reason is due to the complexity of the DS-VRPTW and the large size of the instances. Third, as mentioned before, we involve multiple and integrated transportation modes and delivery options, i.e., vans, cargo bikes, and crowd drivers. We also investigate the impact of varying customer demand. Thus, we conduct a case study in the medium-sized city of Turin (Italy) to analyze the potential influence of using multiple delivery options and crowd drivers in parcel delivery on operational cost, environmental cost, and delivery efficiency. The resulting managerial insights expressed in the form of a set of policies are part of the outcomes of the Logistics and Mobility Plan 2019-2021 led by the Regional Council of Piedmont (ICELab@Polito, the general confederation of Italian industry Confindustria, Piedmont Region, and LINKS Foundation).

The characteristics of our approach reflect the synchronomodality in the urban areas, i.e., the evolution of inter-modal and co-modal transport concepts. In this context, stakeholders of the transport chain actively interact within a cooperative network to flexibly plan transport processes and to be able to switch in real-time between transport modes tailored to available resources [37]. Our model attempts to reduce transportation by vans in favor of cargo bikes and other delivery options and synchronize the transport flow. This would aid planners in reacting to disruptions or new requests while improving the quality and the sustainability of the service through the efficient utilization of available resources [9].

This paper is organized as follows. In Section 2, we review the literature on crowdsourcing and DS-VRPTW. In Section 3, we present the description and formulation of the investigated problem. Section 4 describes the methodology we adopt to solve the problem. In Section 5, we present a case study concerning parcel delivery in the city of Turin (Italy) and we show the experimental results. Finally, we discuss the conclusion and future directions in Section 6.

## II. LITERATURE REVIEW

In this section, we review the literature on crowdsourcing and DS-VRPTW. Crowdsourcing is also known as crowd shipping [10] and crowdsourced delivery [11]. Various researchers in the literature have measured the impact of applying crowd-resources to manage deliveries.

Alnaggar et al. [12] provide a comprehensive review of crowdsourced delivery from current industry status and academic literature. They provide a taxonomy of available platforms based on their matching mechanisms, target markets, and compensation schemes. Based on the taxonomy, the authors identify four decisions that fall under *operational* (e.g., matching and routing) and *tactical levels* (e.g., scheduling and compensation) to enhance the crowdsourced delivery. Kafle et al. [13] propose a crowdsource-enabled system for urban parcel distribution. The cyclists and pedestrians are employed as

crowd drivers for the first and last legs of delivery. The investigated problem is decomposed into a winner determination problem and a simultaneous pickup and delivery problem with soft time windows and solved based on the tabu search algorithm. The authors show the reduction of truck miles traveled and related total cost, by applying the crowdsource-enabled system. Zhen et al. [14] propose six mathematical models to evaluate different operation modes of the crowdsourced delivery. The authors consider several realistic factors, such as the latest service time for each task, task cancellation rate, and range distribution of tasks. Extensive experiments validate the effectiveness of the proposed models and some managerial implications are outlined to help crowdsourced companies to make scientific decisions. Yıldız [15] proposes a “courier friendly” crowd-shipping (CS) model to carry out express package deliveries in an urban area. This model uses transshipment points to enhance operational efficiency and a company-controlled backup delivery capacity to account for the uncertainty in the crowd-provided delivery capacity. The Monte-Carlo simulation approach is used to determine “shadow costs” of capacity utilization and use them in making the assignment (matching) decisions. Le et al. [16] develop and evaluate four different pricing and compensation schemes under different demand and supply scenarios for CS systems. The platform provider’s profits are found more sensitive with the increase of willingness to pay than the rise of expected to-be-paid. The insights are helpful for CS firms to attract and retain customers and couriers in the system, by setting up optimal prices and optimal compensations based on demand and supply levels as well as the firms’ expected profits and platform users’ presuming surplus. Devari et al. [5] investigate the potential of engaging friends or acquaintances in parcel delivery. They found that this strategy reduces total emissions and delivery costs. Dayarian and Savelsbergh [17] explore a form of crowdshipping in which in-store customers supplement company drivers to fulfill online orders for a same-day delivery problem. The authors use myopic and sample-scenario planning approaches to solve respectively, the static and dynamic variants of this problem. Archetti et al. [18] consider occasional drivers as additional couriers to complete delivery. They propose a multi-start heuristic to obtain the near-optimal solution. The comprehensive computational study proves the potential advantages of applying occasional drivers for cost saving. Dahle et al. [19] investigate a pickup and delivery problem with time windows and occasional drivers that employ crowd drivers to take a detour to serve one or more customer requests. The authors model this problem through both load and flow formulation. It is then solved to optimality for up to 70 requests. The results show an average cost savings of 10-15% by engaging crowd drivers. Although good results have been reported in several papers, these researches do not capture the stochastic and dynamic features of parcel delivery. In a real-world application, there are many online requests from customers appearing dynamically. The requests information such as demand, location and time windows are revealed over time when would incur the change of planned routes or rescheduling during the execution process.

According to Ritzinger et al. [20], considering dynamic and stochastic information in VRPs performs the benefits mainly up to 20% of cost saving, carbon emissions reduction and efficiency improvement. Many papers have been published on DS-VRPTW [21, 22]. Bent and Van Hentenryck [23] investigate a dynamic VRPTW with stochastic customers, where the objective function is the maximization of the number of service customers. The authors propose a multiple scenario approach (MSA) to continuously generate the routing plans for scenarios, including known and future customer requests. The computational results show that MSA exhibits significant improvements over approaches not exploiting stochastic information.

Florio et al. [24] develop a branch-price-and-cut algorithm for VRP with stochastic demands. Instances with up to 76 nodes are solved in a reasonable time (up to five hours). They demonstrate that the solution to the stochastic problem is up to 10% less costly than the deterministic one.

Subramanyam et al. [25] apply a robust optimization method to solve a broad class of heterogeneous VRPs under demand uncertainty. To hedge against this uncertain demand, they determine a robust solution that remains feasible for all anticipated demand realizations. Both heuristic and exact approaches are used to improve robust solutions. However, the trade-off between robustness and cost is highly dependent on the choice of the uncertainty set.

Hvattum et al. [26] propose a dynamic stochastic hedging heuristic (DSHH) to solve a DS-VRP. Both locations and demands of the customer are assumed to be unknown. Thus, the Poisson distribution is applied to represent the number of customers revealing at each time interval. A multi-stage stochastic model, extending the two-stage one, is used to capture the stochastic and dynamic elements of the real-world case. When the information on customers is revealed, two recourse actions are used to rearrange the routing plans or start the new route. Comparing to a myopic dynamic heuristic (MDH) that does not consider future events, the DSHH can save more than 15% travel distances.

Sarasola et al. [27] develop an extended variable neighborhood search (VNS) algorithm to investigate a VRP with stochastic demand and dynamic requests. By applying sampling-based VNS, they improve the results by 4.39% on average.

The previous researches highlight how exploiting stochastic information creates many benefits for the routing plan in real-world applications. However, to the best of our knowledge, there is no research combining crowdsourcing with DS-VRPTW.

The main idea behind this combination is to investigate the potential benefits of employing crowd drivers in parcel delivery with stochastic and dynamic customer requests. In this paper, we model this problem as a multi-stage programming problem with recourse. The first recourse action uses crowd drivers to collect the demand of stochastic customers. Then, the second recourse action aims to relocate the customers in planned

routes. We develop a simulation-optimization-based multi-stage heuristic that gradually constructs routes by exploiting statistical information on future customer demand.

### III. PROBLEM DESCRIPTION

In this section, we present the problem setting and the mathematical model.

Different researches and projects on urban logistics have highlighted that freight networks should rely on the interoperability of several business models, stakeholders, and modes of transportation for managing the parcel delivery in the last mile [28-30]. This concept of urban synchronomodality refers to the optimization and synchronization of both the modes of transport and the parcel flows generated by online shopping, including reverse logistics. This would result in an improvement of the economic and environmental sustainability and resilience of multimodal networks. In particular, we consider a decision-maker represented by a traditional courier company (i.e., using vans) that performs a set of customers deliveries with a limited and heterogeneous fleet of vehicles within one day. In doing so efficiently, it outsources the management of some parcels to a green courier company (i.e., using cargo bikes) and crowd-sourced drivers.

We suppose that the traditional courier and the green courier start their operations from a satellite center (usually located in existing urban areas) and a mobile depot in the city center, respectively, where the parcels are consolidated. While the crowd drivers start the journey to pick up and deliver parcels in urban areas, from their original place (e.g., their home). To offer the parcel delivery services, the crowd operators require compensation that usually depends on the number of parcels delivered. A crowd driver may not be available in certain timeslots or, on the contrary, is available only in a specific timeslot.

Reflecting the real practices on the market, we suppose that a great part of customers' orders arises the day before the planning process and thus, they are known to the couriers at the beginning of the working day. On the contrary, other customers' requests could appear dynamically during the day. These requests are managed according to real-time routing decisions, making considerably difficult the quantification a priori of time and costs. The described problem is an extended variant of vehicle routing problems (VRPs) [31] that aims to generate a fleet of vehicle routes to serve customers with minimum costs.

The DS-VRPTW problem combines aspects of both the Dynamic VRP and Stochastic VRP. The number of stochastic customers, their reveal time as well as other attributes (e.g., time window, demand, and location) are only known by their probability distributions. The problem can be firstly solved as the deterministic VRPTW, generating an initial plan for these requests. When the stochastic customers reveal dynamically, some recourse actions must be taken. We allow that the

dynamic customers can be inserted into the initially planned routes or reassigned to a crowd driver who is available to complete the service. If none of the two actions are feasible, the customer requests are thus rejected. Note that if a vehicle has already moved to a new customer, then this customer must be served by that vehicle, i.e., the preemption of customers is not allowed.

### A. Model formulation

In this section, we model the problem as a new variant of the DSVRPTW. In particular, we propose a two-stage stochastic model with recourse that extends the deterministic VRPTW model. Table 1 summarizes the symbols used in the model.

The model relies on the following assumptions. Firstly, the depot  $\{0\}$  is operating during a given time horizon  $[e_0, l_0]$ , and there is a single known moment  $t \in [e_0, l_0]$  at which all unknown information is revealed. At this moment, several recourse actions can be considered so that new customers are served. Let the set of initial locations be denoted by  $N=C \cup \{0\}$ , where  $C = \{1, 2, \dots, n\}$  represents the initially known customers.

Table 1. Notation

| Symbol                     | Definition   |
|----------------------------|--|
| $C, C^+$                   | Initially known customers and stochastic customers   |
| $Q_k, Q_v$                 | Maximum capacity of vehicles and crowd drivers   |
| $N, N'$                    | Set of initial and total locations   |
| $K, V$                     | Set of heterogeneous vehicles and crowd drivers  |
| $\tilde{C}^+, \tilde{D}^+$ | Random variable vector of customer location and demand   |
| $\tilde{E}^+, \tilde{L}^+$ | Random variable vector of earliest and latest arrival time   |
| $c_{ijk}, t_{ijk}$         | Travel cost and travel time from $i$ to $j$ by vehicle $k$   |
| $e_i, l_i$                 | Earliest and latest arrival time of customer $i$ or depot 0  |
| $x_{ijk}$                  | Binary variable on whether vehicle $k$ moves from $i$ to $j$   |
| $x_{ijk}^+$                | Binary variable on whether vehicle $k$ moves from $i$ to $j$ in second-stage recourse action but not in the first-stage solution |
| $x_{ijk}^-$                | Binary variable on whether vehicle $k$ moves from $i$ to $j$ in the first-stage solution but not after recourse action           |
| $y_{iv}$                   | Binary variable on whether $i$ is served by a crowd driver $v$   |
| $p_j$                      | Binary variable on whether customer $j$ is rejected  |
| $s_{ik}$                   | The time at which the service of $i$ is exactly completed by $k$   |
| $s_{ik}^+$                 | The time at which the service of $i$ has completed by $k$ based on the second-stage recourse actions                             |
| $h, a^h$                   | Time interval and its corresponding operational decision   |
| $\xi, E_{\xi}$             | Random variable distribution and expected value of travel cost   |

Let the set of stochastic customers revealed at time  $t$  be denoted by  $C^+ = \{n+1, n+2, \dots, n+n^+\}$ . We can then denote the set of all locations as  $N'=C' \cup \{0\}$ , where  $C'=C \cup C^+$ . Every pair of locations  $i, j \in N'$  is associated with the travel time  $t_{ijk}$  and the travel cost  $c_{ijk}$  for the vehicle  $k$ . Note that the service time is included in  $t_{ijk}$ . Each customer  $i \in C'$  has a demand  $d_i$  and a time window  $[e_i, l_i]$ . The service of a customer  $i$  must be started after  $e_i$  and before  $l_i$ . Waiting at a customer  $i$  is allowed while violating the latest time window  $l_i$  would incur a penalty. A set of heterogeneous vehicles  $K = \{1, 2, \dots, k\}$ , each of maximum capacity  $Q_k$ , starts at and returns depot between time horizon  $[e_0, l_0]$  after completing all the services. Besides, a set of crowd drivers  $V=\{1, 2, \dots, v\}$ , each of capacity  $Q_v$ , starts at

their original place  $O_v$  to visit the assigned customers. Let  $K' = K \cup V$  represents all the available delivery options. The crowd drivers with limited service radius are employed to collect the demand of stochastic customers, and consolidate it at one of the customers locations. Let  $\tilde{\xi} = (\tilde{C}^+, \tilde{D}^+, \tilde{E}^+, \tilde{L}^+)$  be the random variable vector and  $\xi = (C^+, D^+, E^+, L^+)$  is one of its particular realizations.

The two-stage stochastic programming problem can be defined as follows:

$$\min E_{\tilde{\xi}} \left[ \sum_{i \in N} \sum_{j \in N} \sum_{k \in K} c_{ijk} x_{ijk} + Q(x, p, s, \tilde{\xi}) \right] \quad (1)$$

$$Q(x, p, s, \tilde{\xi}) = \sum_{i \in N'} \sum_{j \in N'} \sum_{k \in K'} c_{ijk} (x_{ijk}^+ - x_{ijk}^-) + H_i \sum_{i \in C'} \sum_{v \in V} y_{iv} + P \sum_{j \in C'} p_j \quad (2)$$

$$\text{s.t.} \quad \sum_{i \in N} \sum_{k \in K} x_{ijk} = 1 \quad \forall j \in C \quad (3)$$

$$\sum_{j \in N} x_{0jk} = 1 \quad \forall k \in K \quad (4)$$

$$\sum_{i \in N} x_{ihk} - \sum_{j \in N} x_{hjk} = 0 \quad \forall h \in C, \forall k \in K \quad (5)$$

$$\sum_{i \in N} x_{i0k} = 1 \quad \forall k \in K \quad (6)$$

$$\sum_{i \in N} \sum_{k \in K'} (x_{ijk} + x_{ijk}^+ - x_{ijk}^-) = 1 \quad \forall j \in C' \quad (7)$$

$$\sum_{j \in N'} (x_{0jk} + x_{0jk}^+ - x_{0jk}^-) = 1 \quad \forall k \in K' \quad (8)$$

$$\sum_{i \in N'} (x_{ihk} + x_{ihk}^+ - x_{ihk}^-) - \sum_{j \in N'} (x_{hjk} + x_{hjk}^+ - x_{hjk}^-) = 0 \quad \forall h \in C', \forall k \in K' \quad (9)$$

$$\sum_{i \in N'} (x_{i0k} + x_{i0k}^+ - x_{i0k}^-) = 1 \quad \forall k \in K' \quad (10)$$

$$\sum_{i \in C} d_i \sum_{j \in N} x_{ijk} \leq Q_k \quad \forall k \in K \quad (11)$$

$$\sum_{i \in C} d_i \sum_{j \in N} (x_{ijk} + x_{ijk}^+ - x_{ijk}^-) \leq Q_k \quad \forall k \in K' \quad (12)$$

$$x_{ijk}^- \leq x_{ijk} \quad \forall i, j \in N, \forall k \in K \quad (13)$$

$$x_{ijk}^- = 0 \quad \forall i, j \in C^t, \forall k \in K \quad (14)$$

$$t x_{ijk}^- \leq s_{ik} \quad \forall i, j \in N, \forall k \in K \quad (15)$$

$$s_{ik} + t_{ij} - M_{ij}(1 - x_{ijk}) \leq s_{jk} \quad \forall j \in C, \forall i \in N, \forall k \in K \quad (16)$$

$$e_i \leq s_{ik} \leq l_i \quad \forall i \in N, \forall k \in K \quad (17)$$

$$s_{ik} + t_{i0} - M_{i0}(1 - x_{i0k}) \leq l_0 \quad \forall i \in C, \forall k \in K \quad (18)$$

$$s_{ik}^+ + t_{ij} - M_{ij}(1 - x_{ijk} - x_{ijk}^+ + x_{ijk}^-) \leq s_{jk}^+ \quad \forall j \in C', \forall i \in N', \forall k \in K' \quad (19)$$

$$e_i \leq s_{ik}^+ \leq l_i \quad \forall i \in N', \forall k \in K' \quad (20)$$

$$s_{ik}^+ + t_{i0} - M_{i0}(1 - x_{i0k} - x_{i0k}^+ + x_{i0k}^-) \leq l_0 \quad \forall i \in C', \forall k \in K' \quad (21)$$

$$s_{0k} = e_0 \quad \forall k \in K \quad (22)$$

$$s_{0k}^+ = e_0 + (t - e_0)x_{00k} \quad \forall k \in K' \quad (23)$$

$$x_{ikk} = 0 \quad \forall i \in C, \forall k \in K \quad (24)$$

$$x_{ikk}^+ = 0 \quad \forall i \in C', \forall k \in K' \quad (25)$$

$$x_{ijk} \in \{0,1\} \quad \forall i, j \in N, \forall k \in K \quad (26)$$

$$x_{ijk}^+, x_{ijk}^- \in \{0,1\} \quad \forall i, j \in N', \forall k \in K' \quad (27)$$

$$y_{iv} \in \{0,1\} \quad \forall v \in V, \forall i \in C' \quad (28)$$

$$p_j \in \{0,1\} \quad \forall j \in J \quad (29)$$

$$s_{ik} \in [e_0, l_0] \quad \forall i \in C, \forall k \in K \quad (30)$$

$$s_{ik}^+ \in [e_0, l_0] \quad \forall i \in C', \forall k \in K' \quad (31)$$

The objective function (1) minimizes the expected value  $E_{\xi}$  of total travel costs for the distribution of  $\xi$ . The random quantity  $Q(x, p, s, \xi)$  is the expected cost at the second stage, which includes the cost for adjusting the routing, using crowd drivers and a penalty  $P$  to be paid for the customer rejection. The value of  $P$  would be set very high so that the minimization of the objective function would result in minimizing the number of rejected customers as well. Note that a high value of  $P$  would make the rejection of a customer request unaffordable. Thus, in some cases, a suitable value of  $P$  should be considered.

Constraints (3)-(6) are similar to (7)-(10) and ensure that all customers are served once, and any vehicle must start and end at one depot, respectively. Constraints (5) and (9) guarantee flow conservation. Constraints (11) and (12) ensure that the vehicle capacities are not violated in both first and second stage solutions. Constraints (13)-(15) ensure that a tour between the customer  $i$  and  $j$  in the first-stage solution cannot be skipped in second stage solution recourse if the service in departure location is finished before time  $t$  where  $C^t$  denotes the customers revealed at time  $t$ . Constraints (16)-(18) ensure that the time windows of both customers and depot are not violated, where  $M_{ij}$  is a sufficiently large constant, e.g.,  $M_{ij} = l_i + t_{ij} - e_j$ . Constraints (19)-(21) are applied to track the time at which service is finished in the second-stage solution. Constraints (22) and (23) ensure that vehicle  $k$  cannot leave the depot before time  $t$  in the second stage decision if it does not leave the depot in the first stage. Finally, constraints (26) to (31) express the domain of decision variables.

Note that the crowd drivers can be considered as sources of special delivery capacity that move within the city. When they are employed to pick up the demand from customers, constraints such as capacity, time windows, start and return at the original place, are akin to the constraints of traditional drivers. Note that the crowd drivers have limited service distance. The stochastic requests are assigned to crowd drivers by checking the service distance and feasibility in terms of capacity and time windows constraints. We assume that crowd drivers should consolidate the parcels to a traditional van at one of the nearest customer location or a mobile depot. As the schedules of both vans and crowd drivers are tight, waiting for consolidation would incur the violation of time window constraints. The synchronization between crowd drivers and traditional vans is thus crucial to maintain the parcel delivery functioning, i.e., the crowd drivers and vans should arrive at a

selected customer location or mobile depot at the same time to complete the consolidation.

In this paper, we assume that the operational context of parcel delivery can be divided into a predefined number of time intervals  $h$ . The problem is modified with information revealed during each interval. A multi-stage model can thus be extended from a two stage-stochastic model into an  $h$ -stage model for any given  $h$  by adding additional variables and constraints for each stage [26].

#### IV. METHODOLOGY DESCRIPTION

The exact approach for multi-stage stochastic VRPs currently fails to solve the problem with a large number of customers. The evaluation of recourse cost function (2) can become extremely difficult, depending on the distribution of random variables. Developing a practical heuristic is thus one of the promising approaches to solve this complex and large-size problem. Instead of assuming particular distributions of stochastic variables, we develop a simulation-optimization based multi-stage heuristic based on sample information. Sample scenarios are first generated using a Monte Carlo simulation and then applied to guide a heuristic approach that constructs a routing plan for each time interval in turn. We suppose that the time horizon can be divided into  $h$  intervals  $H_1, \dots, H_h$ , which is related to stages in the multi-stage model. At the beginning of each interval, the algorithm aims to generate a routing plan that minimizes the expected travel cost of serving both known and stochastic customers. More precisely, at each time interval  $h \in [H_1, H_h]$ , an action  $a^h$  must be decided. Each action  $a^h$  contains two parts: first, for each customer  $i \in C'$  revealed at a time interval  $h$ , the action  $a^h$  must accept or reject the customer based on the given constraints. Second, the action  $a^h$  must provide the operational decisions for traditional vehicles or crowd drivers at time interval  $h$  (i.e., service a customer, travel to the next customer, or wait at the current position). Before the online execution, the first action  $a^1$  at time interval  $H_1$  is computed based on a set of known offline customers. A solution is a sequence of actions  $a^{1..h}$  that covers the whole operational horizon. On the one hand, we consider the rescheduling and adoption of the crowd drivers as two recourse actions to deal with the dynamic feature of the problem. On the other hand, using stochastic information during the planning is to capture the stochastic elements. The key idea is to solve each sample scenario as a deterministic VRPTW and then select the distinguished plan from the solutions [23]. A post-optimization procedure is finally applied to compute additional key performance indicators (KPIs).

Figure 1 depicts the framework, and the remaining part of this section describes the details of this algorithm.

##### A. Operational context generation

The operational context is defined by using different sources of information, e.g., city network, vehicle fleet and travel time, customer attributes, and the company's objective and

constraints. The city network is generated considering: i) the city map provided by the local government; ii) geographical coordinates and empirical distributions of customers and depots by courier companies. The data related to the vehicle fleet are provided by courier companies and include the specific attributes of vehicles such as capacity, speed, fuel consumption. Travel times are measured through the sensors spread in the city. We also collect customer attributes, i.e., locations, demand, and time windows. In dynamic applications, stochastic knowledge is known as the probability distributions (e.g., demand and service time). The objective and constraints are defined based on the specific optimization problem.

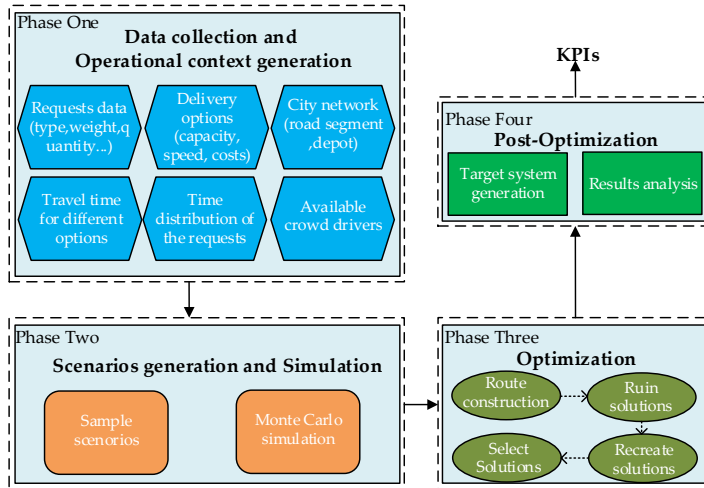


Figure 1. Simulation-Optimization Framework

Some data may be stochastic since the uncertainty of some components in the operational context is involved. These data can be described by random variables including service or travel time, customer demand, etc. and they are typically obtained from the historical data gathered by delivery companies. In other words, these data describe the structure of the problem which would be solved in the next few phases.

We remark that the original data for generating the operational context is provided in the work [32].

### B. Scenarios generation and simulation

To estimate the expected (or recourse) costs we use the Monte Carlo simulation or *multi-scenarios* approach. An approximation algorithm for deterministic VRPTW is thus applied by evaluating a solution on a set of scenarios. Metaheuristic such as LNS and artificial bee colony algorithm [33] can be efficiently engaged with sampling approaches to implement a simulation-based multi-stage stochastic optimization approach,

The scenarios are generated based on the well-defined operational context as well as the knowledge about the probability distributions of the stochastic variables. Each scenario corresponds to a specific realization of all the random variables in an operational context. The distributions of demand, reveal time, time windows and locations are used to create possible future customer requests. Monte Carlo sampling

is used to generate a set of instances. Each instance contains both the known customers and the stochastic ones drawn from the given distributions. These sample instances show that likely events are associated with a high probability. For any given time interval  $h$ , solutions of these sample instances are generated using simple local search algorithms such as *insertion* heuristic and *regret* heuristic. The implementation starts with the routes which are already executed. In particular, each start depot matches the vehicles' current position, while each vehicle capacity is calculated by reducing the weight of goods collected up to the current time. The end depot remains the same, and time windows are appropriately modified. The feasibility of assigning the customers to crowd drivers are also checked based on service distance and other constraints. The frequently visited customers among the sample instances are identified, and a decision is thus made to serve these customers during  $h$  in the final plan. We implement the Google Earth application programming interfaces (APIs) through a georeference module to generate more accurate travel time matrices.

### C. Optimization

The optimization algorithm is an extended meta-heuristic that combines the ruin and recreate principle and a group of general heuristics proposed by [34] and [35]. It aims to generate and improve the routing plans for multiple scenarios ruin and recreate operations. We first suppose that an initial solution  $s$  has been constructed through a simple insertion heuristic. Then, a quantity  $q$  of customers is removed from the solution. We conduct the ruin and recreate operations on the current solution  $s_{new}$ , to diversify the search space and improve the solution. The algorithm ends when it reaches a certain number of iterations (i.e., 5000 iterations). Since the parameter  $q$  determines the neighborhood size, we select an appropriate value  $q$  that balances the computational efforts and solution quality. In our case, we set  $q$  equal to 10% of the maximum customers for each instance. The performance and robustness of LNS are dependent on the selected ruin and recreate operations. Note that in each ruin and recreate operation process, only one heuristic is selected, according to the well-known roulette wheel selections. We describe more details about the whole process in the following subsections.

#### 1) Construct an initial solution

For the given scenario, the algorithm starts with an initial solution generated by the basic greedy heuristic. This heuristic aims to repeatedly insert a request at the cheapest possible position. It means that the request is always inserted into a position with a minimum insertion cost in each iteration. More formally, let  $U$  be the set of unserved customers and  $\Delta f_{ik}$  be the change of objective value generated by inserting request  $i \in U$  at the *cheapest* position in the vehicle  $k$ , if request  $i$  fails to insert in the vehicle  $k$ , the value  $\Delta f_{ik}$  is set to infinite. We thus calculate all the potential insertion and insert request  $i$  in vehicle  $k$  at its minimum cost position as follows:

$$(i, k) := \arg \min_{i \in U, k \in K} (\Delta f_{ik}) \quad (32)$$

Note that only one route is changed in each iteration and the process does not end until all requests are inserted, or no

feasible requests exist. As a simple construction heuristic, the basic greedy heuristic has the potential problem of postponing the placement of expensive customers (i.e., with larger  $\Delta f_{ik}$ ) to the last iterations. It makes it difficult to serve expensive customers. Indeed, many routes might have no space at the last iterations, leading to creating new routes or reject customers. To overcome this potential issue, we consider an alternative approach presented in subsection 3).

## 2) *Ruin operations*

After constructing the initial solution for each scenario, we apply the following four ruin strategies to destroy the initial solution: random removal, worst removal, related removal, and cluster removal. These heuristics take a given solution  $s$  as input and then output a partial solution together with  $q$  removed requests.

The *random removal* is the simplest heuristic that selects  $q$  requests randomly and removes them from the current solution. This process aims to diversify the search space.

The *worst removal* selects some requests that have high costs in their current position. Given a solution  $s$  and a request  $i$ , we define  $f(s, i)$  as the objective value that request  $i$  has been removed from solution  $s$ . The change of objective value  $\Delta f_{-i}$  is defined as  $\Delta f_{-i} = f(s) - f(s, i)$ . The *worst removal* repeatedly selects a new request  $i$  with the highest cost of  $\Delta f_{-i}$  until  $q$  requests are removed. The purpose of the *worst removal* heuristic aims to remove the requests at the worst positions and insert them at other positions to obtain better objective value in the recreate process. However, it is important to maintain the randomization of this heuristic, in case the same customers with expensive costs are removed repeatedly. This is achieved by using a parameter  $p \geq 1$  that controls the selection process. A less expensive customer associated to a high value of  $p$ , may be selected. This probability decreases with the value  $\Delta f_{-i}$ . It means that if the value of  $p$  is small, then the most expensive customer is chosen.

The *related removal* is used to remove the requests that are similar to each other in some sense. The motivation of this heuristic is that we may not obtain any improvement when reinserting the removed requests in the case they are very different from each other. The similarity of request  $i$  and request  $j$  is defined as *relatedness*  $R(i, j)$ . The main idea is to measure the similarity by calculating the difference value in terms of capacity, service-starting time, and distance between requests  $i$  and  $j$ , as indicated in the equation (33).

$$R(i, j) = \varphi c_{ij} + \sigma(|s_i - s_j|) + \tau(|q_i - q_j|) \quad (33)$$

Note that all terms in this equation are normalized in the range  $[0, 1]$ . The related removal procedure removes a random customer, and in the next iterations, it selects customers that are similar to the already removed customers. The parameter  $p$  is again used to control the selection process as we do in the *worst removal*. For further details on the heuristic, the interested reader can refer to [36, 37].

Finally, the *cluster removal*, as a variant of the *related removal*, is applied to remove clusters of related requests from a few routes. The purpose of this heuristic is to remove the clusters of requests entirely from different routes in case the single removed request would be inserted back into the route.

## 3) *Repair operations*

After applying the ruin operation, a group of repair operation is used to generate new solutions for each scenario. The operation is conducted in parallel since there are different scenarios. The *basic greedy* heuristic is applied again to recreate the new solutions. However, this simple heuristic may insert some requests back in their previous position. The *regret* heuristics are then used to prevent this problem by using a kind of look-ahead information. Let  $\Delta f_i^q$  be the change of objective value generated by inserting customer  $i$  at its best position in the  $q$ th *cheapest* route for customer  $i$ . The value of  $\Delta f_i^2$  is thus the change of objective value by inserting customer  $i$  into the route where the customer could be inserted *second-cheapest*. In each iteration, the customer  $i$  is selected based on  $i := \arg \max_{i \in N} (\Delta f_i^2 - \Delta f_i^1)$ . This operation aims to maximize the difference in the cost of inserting customer  $i$  at its best route and second-best route. This process is repeated until no more customers can be inserted. Instead of using a simple acceptance criterion that only accepts the solution with better objective value, the *simulated annealing* strategy is used to select the solutions based on a varying probability. The detail of this strategy can be found in the work by [34].

These steps described above are repeated until the given termination criterion is met, i.e., reach the maximum running time or have no improvement for continuous iterations. As phase 3 continues, a set of plans is maintained at each interval. To decide which customers should be fixed in the routes for the current interval, a natural operation is used to exploit the common features among the maintained routing plans, i.e., find the customer that is most frequently visited in the current time interval in the multiple sample scenarios. Once these customers are identified, the action  $a^h$  is then fixed to visit these customers during the current time interval  $h$  in the final plan. At the end of the last time interval, a post-optimization and results analysis are conducted in phase 4, and we compute the related KPIs.

## V. CASE STUDY

In this section, we analyze the potential impact of crowdsourcing and multiple delivery options in terms of economic, environmental, and operational sustainability for on-demand parcel delivery. In doing so, we conduct a real case study related to urban logistics in the city of Turin (Italy). We first present the description of the case study and the related operational context. Then, we discuss the computational results to provide useful insights for decision-makers.

### A. *Description of the operational context*

Last mile presents serious challenges to the operations of the supply chain network. Thus, alternative distribution systems

architectures have been proposed to tackle these challenges and enhance the efficiency of last-mile distribution. A promising solution that we consider in this study is the adoption of a two-tier system[38]. In the first level, trucks perform deliveries from distribution centers located in a strategic node of the city to Urban Consolidation Centers named satellite. They are generally transshipment points situated in the proximity of a city center. At the second level, orders are consolidated in city freighters, i.e., small vehicles that can move easily along any street in the city center area managed even partially with crowdsourcing contracts [39].

In this case study, we consider four benchmarks integrating van, bike, and crowdsourcing:

- Benchmark 1 (B1): only traditional vans (fossil-fueled) are used to manage the parcel delivery.
- Benchmark 2 (B2): green carriers, such as drivers using cargo bikes or bicycles with the messenger bag, are used as an environment-friendly delivery option, providing economic and operational benefits for parcel delivery. In the B2, we consider both van and bike as transportation modes.
- Benchmark 3 (B3): we consider crowdsourcing as a flexible delivery option. The crowd driver plays the role of additional capacity to meet the on-demand requests. In practice, the delivery tasks are assigned based on the distance between crowd drivers and customers.
- Benchmark 4 (B4): all the above delivery options are integrated.

The urban distribution setting and data used in the paper are inspired from the analysis of a real case study of the city Turin conducted by the CARS@Polito [40], and the ICELab@Polito [41], with the collaboration of the Torino Living Lab project [42] and the Amazon Innovation Award. The managerial insights coming from this work will be part of the new Logistics and Mobility Plan to be activated in 2022 in the Piedmont region, for which one of the authors is responsible.

In particular, to generate specific operational contexts, we have fused the parameters and data coming from the following sources:

- URBeLOG project for the distribution of customers and real (and anonymized) information about their location [43];
- Municipality of Turin for what concerns the satellite location, city map and data on the road network from sensors in the city;
- the study by [32], regarding vehicle characteristics, costs and revenue structure of parcel delivery companies.

In particular, the city network presented in Figure 2 is generated using a  $2.805 \times 2.447$  km<sup>2</sup> area in Turin that includes both the center of the city and a semi-central area. We consider that area because, according to [44], it is the area in which the different modes can coexist sustainably (from economic, environmental, social, and operational perspectives) and is also the most populated area of the city, covering more than the 80% of the total population. We use a distribution center located on the outskirts of the city to serve the traditional carriers while a mobile depot in the city center is a satellite facility for the green carrier and crowd drivers. Road segments of this network are represented by two sequential connected points. The roads' information is extracted from the shapefiles provided by the local public authority in Turin. The average speed on each road segment is monitored by the speed sensors around the city area. Each point on this network is associated with a unique ID number and real GPS coordinates.

Table 2 presents the values of the capacity, speed and service time for each delivery option. As these parameters are provided by an international parcel delivery company operating in Turin, their values are assumed to be fixed and thus, we consider them as input in our case study. The crowd drivers move within the city. We thus randomly generate the location of crowd drivers on this graph. We conduct some preliminary experiments to obtain a suitable number of available crowd drivers. The aim is to test and analyze the impact of the different numbers of available crowd drivers for B4. The detailed results are represented in subsection V.B.1. We classify the parcels based on their weights as mailers (0-3kg), small parcels (3-6kg), and large parcels (over 6kg), while the percentage of total parcels for each type is defined as 57%, 13%, and 30% respectively, according to [28].

In the operational context, we separate the eight working hours as four time buckets with the same length, according to the current standard for timeslots in time-sensitive urban delivery (e.g., Amazon Prime Now, Uber Freight). Each bucket is split by 1 minute time unit. For each potential customer, the demand is generated based on its parcel type while the time window is specified for the time bucket by the simulator.

In our simulation, we consider instances with 550, 350, and 150 potential customers, respectively. For each context, there are 70% offline requests which are known before scheduling. Meanwhile, 30% of potential customers are assigned as prime members with a priority which restricts their time window to the first two time buckets. The expected behavior of each potential customer for the investigated problem is described.

Table 2. Input parameters

| Delivery Options | Maximum parcel size(kg) | Capacity  | Coverage (km) | Speed (km/h) | Service time(min) |               |               |
|------------------|-------------------------|-----------|---------------|--------------|-------------------|---------------|---------------|
|                  |                         |           |               |              | Mailers           | small parcels | large parcels |
| Van              | 70                      | 700kg     | NA            | 40           | 4                 | 4             | 5             |
| Cargo Bike       | 15                      | 70kg      | NA            | 20           | 2                 | 2             | —             |
| Crowd Driver     | 6                       | 4 Parcels | 2             | 15           | 2                 | —             | —             |

For each potential customer location  $i$  and each time unit  $t$  of the time horizon, a probability is associated with an online request (i.e., picking up a parcel) that may reveal at a time  $t$  for location  $i$ .

Once all the locations are defined, the mutual distances between these customers and depots in the network are then generated. Instead of using Euclidean distance, the distance matrix among the points is calculated using Dijkstra's shortest path method. The travel time matrixes are generated based on these distance matrixes and original speed among the road segments in the input data.

We consider three different degrees of dynamism (DOD) to address dynamic online requests, i.e., 15%, 30%, and 45%. For example, DOD-15% means that there are 15% of dynamic requests in total customers. For each operational context, we thus obtain three different sub-contexts, which are used to measure the impact of DOD. For each sub-context, we generate 10 independent scenarios with 10 sets of dynamic online requests, obtained sampling their probability distributions. Each online request is associated with its location, demand, time windows, and the time when they appear in the time horizon. This yields a total of 360 instances. They are available on the Github repository at the following link: <https://github.com/gcmswm/Benchmarks>.

The objective function minimizes first the total travel cost (expected) of parcel delivery and second the number of rejected requests.

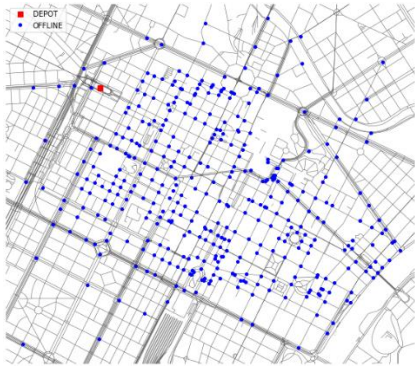


Figure 2. City network in the case study. Note that the red square represents the mobile, while the blue circles are the offline customers.

### B. Numerical analysis

In this section, we analyze the impact of adopting crowd drivers and multimodality on the sustainability of the parcel delivery.

We conduct experiments based on some randomly generated test problems. For each benchmark and operational context, we perform 10 independent tests. Therefore, we solve the 360 instances independently, by the optimization procedure. To measure the experimental results, we calculate different KPIs that reflect the mix of economic, environmental, and operational facets of the service:

- *Economic sustainability.* According to the current real practices, the delivery cost is associated with the number

of parcels served by different delivery options since different options have different contract costs for each parcel. We define the KPI as cost per delivery for each option:

- cost per delivery (van), the unit cost of each parcel delivered by traditional van;
- cost per delivery (bike), the unit cost of each parcel served by bike.

For the sake of simplicity of exposition and brevity, we refer to costs structure analyzed in [44]. In particular, we consider the operating costs related to the vans and cargo bikes. The latter are usually adopted by crowd drivers.

These costs are computed per kilometer traveled in the last mile segment of the supply chain, and include both variable costs (e.g., gasoline) and the total cost of ownership of the vehicle [44]. In particular, we consider both costs directly related to the vehicles (e.g., purchase cost, taxes, insurance, fuelling, and maintenance costs) and personnel costs (e.g., drivers/bikers salaries and related taxes). The typical contract scheme in the parcel delivery industry imposes then the conversion from a cost per kilometer to a cost per stop.

In the crowdsourcing, we also consider the compensation per delivery that reflects the unit paid by the company to crowd driver for each delivered parcel. The exact compensation is dependent on many factors such as distance, weight, and the local market. However, for the medium distance, inter-city market, we follow the investigation from [6] and use an average value (i.e., 1.8€ per delivery) as compensation for crowd drivers.

- *Environmental sustainability.* We consider the emissions and costs of the overall last-mile chain, according to the latest regulation, the ISO/TS 14067:2013 “Greenhouse gases - Carbon footprint of product - Requirements and guidelines for quantification and communication”. In particular, we consider three types of emissions: direct emissions from the fuel combustion process, indirect emissions, emitted by the fuel production process, and the long-haul shipment of the fuel, CO<sub>2</sub> equivalent to including pollutants, such as nitrogen oxide. Thus, we calculate the CO<sub>2</sub> emission saved in B2, B3, and B4, based on the lower travel distance by applying green delivery options (i.e., bike and crowdsourcing).
- *Operational sustainability.* The operational efficiency of the delivery system is normally evaluated in terms of fulfilled deliveries. It becomes challenging when the demand increases. We thus use the total number of parcels served per hour ( $nS/h$ ) to measure operational sustainability.

#### 1) Preliminary analysis for crowdsourcing

This section conducts a preliminary experiment to measure the suitable number of crowd drivers in our case study. The average distance traveled by vans and the total average economic cost are calculated for comparison. In this

experiment, the number of crowd drivers is set equal to 5, 10, 15, 20, and 25, while their locations are randomly generated spreading within the urban area. The experiment is analyzed by using benchmark B4 with 350 potential customers.

In general, there are some limitations to the number of crowd drivers in real-world applications due to different reasons: the customers' requests, the availability and motivation of crowd drivers, and suitable compensation strategies.

To find a proper number of crowd drivers in our case study, we analyze the results shown in Figure 3. In particular, the values in Figure 3a highlight that when the number of available crowd drivers is low (i.e., 5 drivers), and the degree of dynamism is low (i.e., 15%), these drivers are used to manage the online requests given their flexibility. When the degree of dynamism increases while the number of available crowd drivers is still low and thus, insufficient to cope with the high online requests, vans remain the promising choice.

Increasing the number of crowd drivers, they are adopted to manage a greater part of the requests than the previous scenario, also when the degree of dynamism increases. This leads to a reduction of the distances traveled by vans (e.g., when DOD is 45%, increasing the number of crowd drivers from 5 to 25, the average van distance decreases by 20%). Moreover, when the number of crowd workers is equal to 20, we obtain a reduction of the average van distance for all the DODs. In particular, for DOD equals to 15%, we reach the lowest value of the distance traveled using vans. This is reflected in the lowest value of the economic costs faced by the traditional courier company (Figure 3b). While the economic cost for DOD-30% and DOD-45% has no significant variation for the different number of crowd drivers.

We thus decide to use 20 as the baseline number for crowd drivers for our case study. This choice is based on two reasons in the real-world application. First, when the number of crowd drivers is too large, some of the crowd drivers cannot receive enough delivery tasks during the execution, which will decrease their motivation for participating in the parcel delivery. Second, if the number of crowd drivers is too small, there can a lack of available crowd drivers in some local areas so that the expected service quality may not be guaranteed.

However, the optimal capacity planning of crowd drivers is a complex problem due to many factors including the compensation of delivery, the motivation of crowd drivers, and the available customer requests. The interested readers are referred for more details to [45].

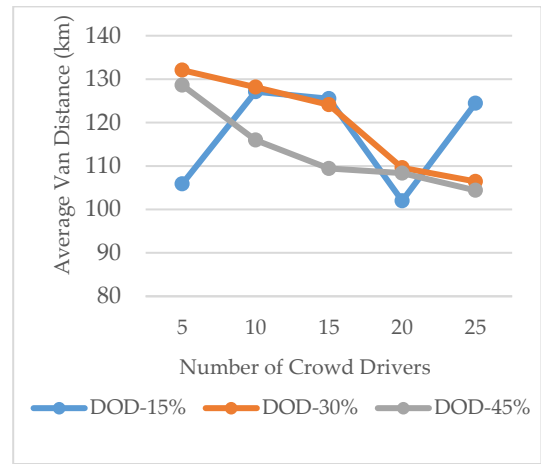


Figure 3a. Comparison of van distance

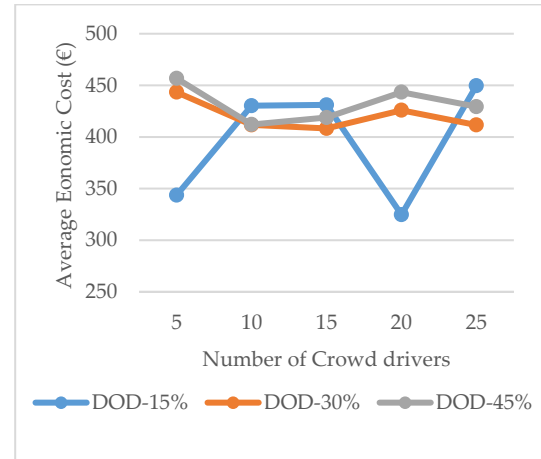


Figure 3b. Comparison of economic cost

## 2) Performance comparison for benchmarks

To demonstrate the performance of using different options for parcel delivery, we calculate the percentage of each KPI compared with benchmark B1. Figure 4 presents the performance of the traditional vans in B2, B3, and B4. The figures are calculated as a percentage variation of each KPI to the value of the same KPI in B1. The operating and environmental cost savings obtained by using cargo bikes and crowdsourcing are denoted by  $\Delta OC$  and  $\Delta EC$  respectively. While  $\Delta Efficiency$  refers to the reduction of efficiency due to the reduced number of services (see [25] for a detailed description of the computation of the KPIs).

Figure 4 illustrates the improvement of economic and environmental sustainability when applying different delivery options (vans, cargo bikes, and crowdsourcing) for parcel delivery. As shown in Figure 4, the adoption of cargo bikes (B2), crowdsourcing (B3), and their combination (B4) lead to the reduction of economic and environmental costs. In particular, these reductions are obtained by reducing the number of vans and their total travel distance, which leads to 25%, 21%, and 44% reduction of average economic cost for B2, B3, and B4, respectively. Meanwhile, the average decreases of CO<sub>2</sub> emissions for B2, B3, and B4 are 116.68kg, 115.01kg, and

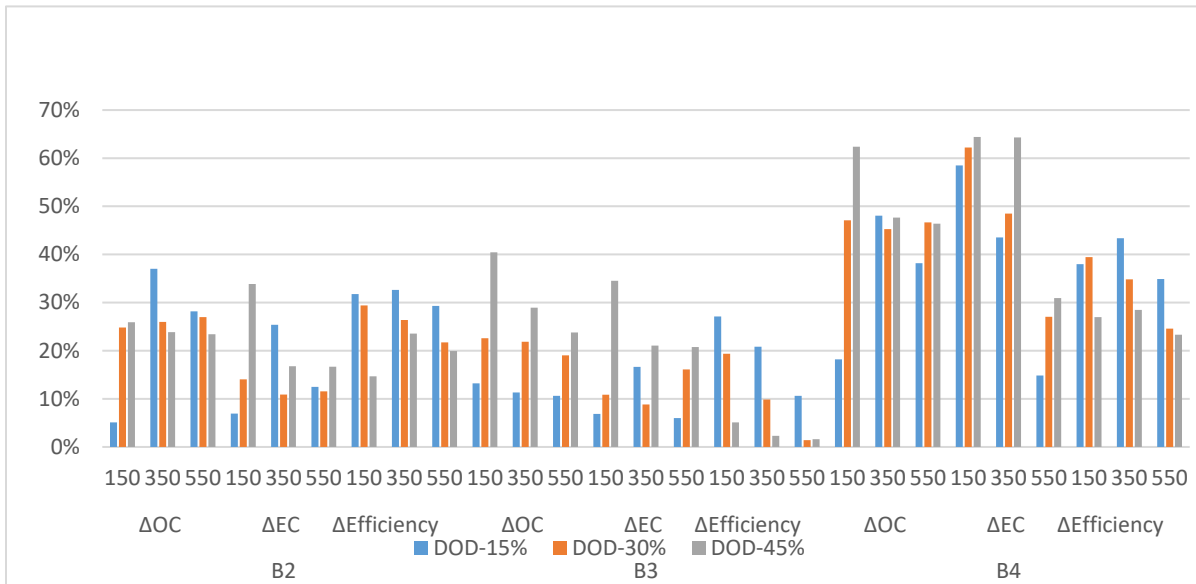


Figure 4. Performance comparison of different benchmarks

297.05kg that contribute to 17%, 16%, and 46% total average reductions of the environmental cost, respectively. In particular, when the degree of dynamism is equal to 45%, the  $\Delta EC$  values are larger than the other two counterparts, i.e., DOD-15% and DOD-30%. According to the results, the potential benefits of applying cargo bikes and crowdsourcing as green carriers are demonstrated for investigated three benchmarks. The most significant finding is that combining both cargo bikes and crowdsourcing into traditional van delivery reaches the highest reduction of economic and environmental costs. In addition, the loss of efficiency for B2, B3, and B4 is 25%, 11%, and 33% on average, respectively. B3 reaches the minimum loss of efficiency for parcel delivery, while the economic and environmental cost saving is promising, i.e., 21% and 16% on average. Though the B4 reaches the maximum average loss of efficiency at 33%, the maximum economic and environmental cost savings (44% and 46%) are reached. In practice, when crowd drivers and cargo bikes are involved with traditional vans, there should be a balance between the increase in profits and service quality. The integration of different delivery options should be managed wisely in terms of balancing the workload, working conditions, efficiency, and service quality.

In this paper, we consider two strategies to deal with the high demand for dynamic requests. The first one is to assign the online requests to the spreading crowd drivers according to the available recourse actions proposed by the optimization solver.

The second is to accommodate the online requests to the existing vehicles with spare capacity.

To analyze the impact of crowd drivers for on-demand parcel delivery, we compare the rejected customer requests for each benchmark. Figure 5 presents a boxplot of the results for the four benchmarks. This figure represents the case with 550 potential customers, as there are no rejected requests in

instances with 150 and 350 customers. Three different DODs are considered since the different number of dynamic requests may have different influences on the number of rejected requests. As shown in Figure 6, the number of rejected requests in B1 is the largest independently by the DOD. When the DOD increases, the number of rejected requests increases significantly. This result shows that only using traditional vans as the delivery option would cause more rejected requests when the dynamic requests are higher. When the cargo bikes are integrated, as shown in B2, the number of rejected requests decreases while the result has the same trend with B1 since the number of rejected requests is also increased with DOD.

However, the trend in B3 and B4 is different. The number of rejected requests in both B3 and B4 remains stable for different DODs. Note that the results of B4 are better than B3 for each DOD since they have a lower minimum, average and maximum number of rejected requests according to the boxplot. In addition, there are only a few numbers of rejected requests in both B3 and B4. This result indicates that introducing crowd drivers as a delivery option can significantly reduce the number of rejected requests for our investigated instances. Therefore, considering these results, we conclude that green carriers and crowd drivers are promising delivery options to deal with online customer requests in the context of stochastic and dynamic parcel delivery.

### 3) Influence of different customer demand

To illustrate the influence of varying customer demand on sustainability performance (i.e., operational cost, environmental cost and efficiency), we conduct a group of experiments on B4 by changing customer demand. We decide to conduct a sensitivity analysis on this parameter as the uncertainty on the composition of the demand will affect in the near future the congestion and the development of urban areas [28, 46].

In doing so, we create three new groups of customer demand varying the composition of the demand as follows:

- reduction of 20% resulting in a market downturn;
- increase of 20% and 40% to suppose a market expansion.

The results are represented in Table 3, with respect to the current situation of demand.

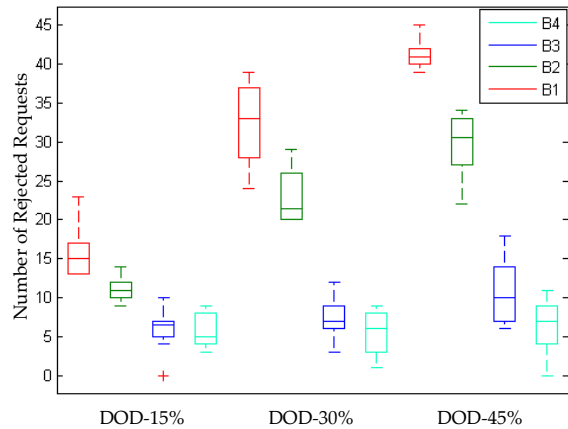


Figure 5. Comparison of rejected requests for different benchmarks

The values of  $\Delta OC$ ,  $\Delta EC$  and  $\Delta Efficiency$  represent the percentage variations of operational cost, environmental cost and efficiency, respectively, between the normal customer demand and three other different groups of demand. As shown in Table 3, when the customer demand decreases to 80%, the operational cost, environmental cost and delivery efficiency of the delivery system decrease for all investigated instances. On the other hand, when the customer demand increases to 120%, the operational cost increases ranging from 7.8%-14.1% for the investigated instances. The environmental cost increases by 9.7%, 17.3% and 14.8% for three different instance sizes. The efficiency of parcel delivery is witnessed to a few increases ranging from 0.2%-6.3%. Moreover, the demand expansion of 40%, has the most significant change among other instances. For example, the operational cost increases by 17.1%, 18.2%, and 14.6%, respectively, and the environmental cost increases by 35.6%, 34.7%, and 28.2%, respectively.

The delivery efficiency has a significant increase in all instances. The results show that customer demand has a significant impact on operational cost as well as environmental cost. When the customer demand decreases/increases, operational and environmental cost decreases/increases. The potential reason behind this phenomenon is that the total required vehicles are changed in terms of both vans and bikes as well as crowdsourcing. However, it is not a linear function between customer demand and costs since many other factors must be considered. Thus, it is difficult for companies to decide how many vehicles should be prepared for varying customer demand. To solve this issue, one of the potential solutions, we believe, is to introduce more flexible delivery options like cargo bikes from a third party, or crowd drivers from the social community. Both two options are suitable for the on-demand market and would not cause much more fixed costs for delivery companies.

Table 3. Impact of different customer demand

| Demand | instance size | $\Delta OC$ | $\Delta EC$ | $\Delta Efficiency$ |
|--------|---------------|-------------|-------------|---------------------|
| 80%    | 150           | 13.9%       | 9.4%        | 10.0%               |
|        | 350           | 5.3%        | 13.8%       | 29.1%               |
|        | 550           | 3.9%        | 12.5%       | 36.7%               |
| 120%   | 150           | -7.8%       | -9.7%       | -6.3%               |
|        | 350           | -14.1%      | -17.3%      | -0.2%               |
|        | 550           | -8.2%       | -14.8%      | -2.6%               |
| 140%   | 150           | -17.1%      | -35.6%      | -18.7%              |
|        | 350           | -18.2%      | -34.7%      | -11.7%              |
|        | 550           | -14.6%      | -28.2%      | -8.1%               |

## VI. CONCLUSIONS AND FUTURE DIRECTIONS

In this paper, we addressed a DSVRPTW problem with crowdsourcing for on-demand parcel delivery. This research topic is inspired by the practice in the parcel delivery market including the advent of crowdsourcing in our community. Indeed, we integrate multiple delivery options into this problem together with crowd drivers.

To cope with this research topic, we proposed a multi-stage stochastic model, and the problem is solved by following a simulation-optimization framework.

We conducted a case study in the medium-sized city of Turin (Italy) to measure the potential impact of using cargo bikes, crowdsourcing in parcel delivery. The numerical experiments show that combining crowd drivers and green carriers into traditional van delivery is beneficial thanks to economic and environmental cost-saving, while the delivery efficiency decreases. In particular, the total travel distance and CO<sub>2</sub> emissions are reduced in our investigated instances. Besides, green carriers and crowdsourcing are promising and flexible solutions when dealing with many online requests. We varied the customer demand investigating its potential impacts on the system. The results show that operational costs and environmental costs are sensitive to variations of customer demand.

Future directions will consider the optimal workforce capacity planning in crowdsourcing applications as well as their compensation strategies, i.e., hourly compensation, per-delivery compensation, and driver-determined compensation [12]. Besides, it is promising to address the stochastic and dynamic pickup and delivery problem with time windows, especially for tighter and overlapping time windows.

## ACKNOWLEDGMENT

While working on this paper, the first author was the Director of the ICT for City Logistics and Enterprises Lab and the head of the Urban Mobility and Logistics Systems (UMLS) initiative of the interdepartmental Center for Automotive Research and Sustainable mobility (CARS) at Politecnico di Torino, Italy.

## REFERENCES

- [1] R. G. Thompson, "Vehicle orientated initiatives for improving the environmental performance of urban freight systems," in *Green logistics and transportation*: Springer, 2015, pp. 119-129.
- [2] F. Basik, B. Gedik, H. Ferhatosmanoglu, and K.-L. Wu, "Fair task allocation in crowdsourced delivery," *IEEE Transactions on Services Computing*, 2018.
- [3] Y. Wang, D. Zhang, Q. Liu, F. Shen, and L. H. Lee, "Towards enhancing the last-mile delivery: An effective crowd-tasking model with scalable solutions," *Transportation Research Part E: Logistics and Transportation Review*, vol. 93, pp. 279-293, 2016.
- [4] A. M. Arslan, N. Agatz, L. Kroon, and R. Zuidwijk, "Crowdsourced delivery—A dynamic pickup and delivery problem with ad hoc drivers," *Transportation Science*, vol. 53, no. 1, pp. 222-235, 2018.
- [5] A. Devari, A. G. Nikolaev, and Q. He, "Crowdsourcing the last mile delivery of online orders by exploiting the social networks of retail store customers," *Transportation Research Part E: Logistics and Transportation Review*, vol. 105, pp. 105-122, 2017, doi: 10.1016/j.tre.2017.06.011.
- [6] A. Punel and A. Stathopoulos, "Modeling the acceptability of crowdsourced goods deliveries: Role of context and experience effects," *Transportation Research Part E: Logistics and Transportation Review*, vol. 105, pp. 18-38, 2017, doi: 10.1016/j.tre.2017.06.007.
- [7] T. C. Du, E. Y. Li, and D. Chou, "Dynamic vehicle routing for online B2C delivery," *Omega*, vol. 33, no. 1, pp. 33-45, 2005, doi: 10.1016/j.omega.2004.03.005.
- [8] V. Pillac, M. Gendreau, C. Guéret, and A. L. Medaglia, "A review of dynamic vehicle routing problems," *European Journal of Operational Research*, vol. 225, no. 1, pp. 1-11, 2013, doi: 10.1016/j.ejor.2012.08.015.
- [9] R. Giusti, D. Manerba, G. Bruno, and R. Tadei, "Synchronodal logistics: An overview of critical success factors, enabling technologies, and open research issues," *Transportation Research Part E: Logistics and Transportation Review*, vol. 129, pp. 92-110, 2019.
- [10] I. Dayarian and M. Savelsbergh, "Crowdshipping and same-day delivery: Employing in-store customers to deliver online orders," *Optimization Online*, pp. 07-6142, 2017.
- [11] H. Paloheimo, M. Lettenmeier, and H. Waris, "Transport reduction by crowdsourced deliveries – a library case in Finland," *Journal of Cleaner Production*, vol. 132, pp. 240-251, 2016, doi: 10.1016/j.jclepro.2015.04.103.
- [12] A. Alnaggar, F. Gzara, and J. H. Bookbinder, "Crowdsourced delivery: A review of platforms and academic literature," *Omega*, p. 102139, 2019/10/08/ 2019, doi: <https://doi.org/10.1016/j.omega.2019.102139>.
- [13] N. Kafle, B. Zou, and J. Lin, "Design and modeling of a crowdsourced-enabled system for urban parcel relay and delivery," (in English), *Transport Res B-Meth*, vol. 99, pp. 62-82, May 2017, doi: 10.1016/j.trb.2016.12.022.
- [14] L. Zhen, Y. Wu, S. Wang, and W. Yi, "Crowdsourcing mode evaluation for parcel delivery service platforms," *International Journal of Production Economics*, vol. 235, p. 108067, 2021.
- [15] B. Yıldız, "Express package routing problem with occasional couriers," *Transportation Research Part C: Emerging Technologies*, vol. 123, p. 102994, 2021.
- [16] T. V. Le, S. V. Ukkusuri, J. Xue, and T. Van Woensel, "Designing pricing and compensation schemes by integrating matching and routing models for crowd-shipping systems," *Transportation Research Part E: Logistics and Transportation Review*, vol. 149, p. 102209, 2021.
- [17] I. Dayarian and M. Savelsbergh, "Crowdshipping and same-day delivery: Employing in-store customers to deliver online orders," *Production and Operations Management*, vol. 29, no. 9, pp. 2153-2174, 2020.
- [18] C. Archetti, M. Savelsbergh, and M. G. Speranza, "The Vehicle Routing Problem with Occasional Drivers," *European Journal of Operational Research*, vol. 254, no. 2, pp. 472-480, 2016, doi: 10.1016/j.ejor.2016.03.049.
- [19] L. Dahle, H. Andersson, M. Christiansen, and M. G. Speranza, "The Pickup and Delivery Problem with Time Windows and Occasional Drivers," *Computers & Operations Research*, 2019/04/28/ 2019, doi: <https://doi.org/10.1016/j.cor.2019.04.023>.
- [20] U. Ritzinger, J. Puchinger, and R. F. Hartl, "A survey on dynamic and stochastic vehicle routing problems," *International Journal of Production Research*, vol. 54, no. 1, pp. 215-231, 2016.
- [21] G. Kim, Y. S. Ong, T. Cheong, and P. S. Tan, "Solving the dynamic vehicle routing problem under traffic congestion," *Ieee T Intell Transp*, vol. 17, no. 8, pp. 2367-2380, 2016.
- [22] Z. Cao, H. Guo, J. Zhang, D. Niyato, and U. Fastenrath, "Finding the shortest path in stochastic vehicle routing: A cardinality minimization approach," *Ieee T Intell Transp*, vol. 17, no. 6, pp. 1688-1702, 2015.
- [23] R. W. Bent and P. Van Hentenryck, "Scenario-Based Planning for Partially Dynamic Vehicle Routing with Stochastic Customers," *Operations Research*, vol. 52, no. 6, pp. 977-987, 2004, doi: 10.1287/opre.1040.0124.
- [24] A. M. Florio, R. F. Hartl, and S. Minner, "New exact algorithm for the vehicle routing problem with stochastic demands," *Transportation Science*, vol. 54, no. 4, pp. 1073-1090, 2020.
- [25] A. Subramanyam, P. P. Repoussis, and C. E. Gounaris, "Robust optimization of a broad class of heterogeneous vehicle routing problems under demand uncertainty," *INFORMS Journal on Computing*, 2020.
- [26] L. M. Hvattum, A. Løkketangen, and G. Laporte, "Solving a Dynamic and Stochastic Vehicle Routing Problem with a Sample Scenario Hedging Heuristic," *Transportation Science*, vol. 40, no. 4, pp. 421-438, 2006, doi: 10.1287/trsc.1060.0166.
- [27] B. Sarasola, K. F. Doerner, V. Schmid, and E. Alba, "Variable neighborhood search for the stochastic and dynamic vehicle routing problem," *Annals of Operations Research*, vol. 236, no. 2, pp. 425-461, 2015, doi: 10.1007/s10479-015-1949-7.
- [28] G. Perboli and M. Rosano, "Parcel delivery in urban areas: Opportunities and threats for the mix of traditional and green business models," *Transportation Research Part C: Emerging Technologies*, vol. 99, pp. 19-36, 2019/02/01/ 2019, doi: <https://doi.org/10.1016/j.trc.2019.01.006>.
- [29] T. G. Crainic, G. Perboli, and M. Rosano, "Simulation of intermodal freight transportation systems: a taxonomy," *European Journal of Operational Research*, vol. 270, no. 2, pp. 401-418, 2018.
- [30] J. F. de Sousa and J. Mendes-Moreira, "Urban Logistics Integrated in a Multimodal Mobility System," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 2015: IEEE, pp. 89-94.
- [31] G. B. Dantzig and J. H. Ramser, "The truck dispatching problem," *Management science*, vol. 6, no. 1, pp. 80-91, 1959.
- [32] G. Perboli, M. Rosano, M. Saint-Guillain, and P. Rizzo, "Simulation-optimisation framework for City Logistics: an application on multimodal last-mile delivery," *IET Intelligent Transport Systems*, vol. 12, no. 4, pp. 262-269, 2018, doi: 10.1049/iet-its.2017.0357.

- [33] Q. Wei, Z. Guo, H. C. Lau, and Z. He, "An artificial bee colony-based hybrid approach for waste collection problem with midway disposal pattern," *Applied Soft Computing*, 2019.
- [34] D. Pisinger and S. Ropke, "A general heuristic for vehicle routing problems," *Computers & Operations Research*, vol. 34, no. 8, pp. 2403-2435, 2007, doi: 10.1016/j.cor.2005.09.012.
- [35] P. Shaw, "A new local search algorithm providing high quality solutions to vehicle routing problems," *APES Group, Dept of Computer Science, University of Strathclyde, Glasgow, Scotland, UK*, 1997.
- [36] P. Shaw, "Using constraint programming and local search methods to solve vehicle routing problems," in *International conference on principles and practice of constraint programming*, 1998: Springer, pp. 417-431.
- [37] S. Ropke and D. Pisinger, "An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows," *Transportation science*, vol. 40, no. 4, pp. 455-472, 2006.
- [38] G. Perboli, R. Tadei, and D. Vigo, "The two-echelon capacitated vehicle routing problem: Models and math-based heuristics," *Transportation Science*, vol. 45, no. 3, pp. 364-380, 2011.
- [39] G. Perboli, L. Brotcorne, M. E. Bruni, and M. Rosano, "A new model for Last-Mile Delivery and Satellite Depots management: The impact of the on-demand economy," *Transportation Research Part E: Logistics and Transportation Review*, vol. 145, p. 102184, 2021.
- [40] CARS@Polito, "Center for automotive research and sustainable mobility web site," URL <http://www.cars.polito.it/>. Last access: 10/05/2021.
- [41] ICELAB@Polito, "ICT for City Logistics and Enterprises Lab web site," URL <http://icelab.polito.it/>. Last access: 10/05/2021.
- [42] Municipality of Turin, "Torino living lab," URL <http://torinolivinglab.it/en/>. Last access: 10/05/2021.
- [43] A. De Marco et al., *Business Modeling of a City Logistics ICT Platform*. 2017, pp. 783-789.
- [44] L. Brotcorne, G. Perboli, M. Rosano, and Q. Wei, "A Managerial Analysis of Urban Parcel Delivery: A Lean Business Approach," *Sustainability*, vol. 11, no. 12, p. 3439, 2019.
- [45] H. Dai and P. Liu, "Workforce planning for O2O delivery systems with crowdsourced drivers," *Annals of Operations Research*, 2019, doi: 10.1007/s10479-019-03135-z.
- [46] R. Giusti, D. Manerba, G. Perboli, R. Tadei, and S. Yuan, "A New Open-source System for Strategic Freight Logistics Planning: the SYNCHRO-NET Optimization Tools," *Transportation Research Procedia*, vol. 30, pp. 245-254, 2018/01/01/ 2018, doi: <https://doi.org/10.1016/j.trpro.2018.09.027>.



**Guido Perboli** is Professor of Operations Management in Politecnico di Torino and Associate Member of the CIRRELT - Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport, Montréal, Canada. In 2016 he founded the ICT for City Logistics and Enterprises (ICE) center of Politecnico di

Torino, a research center focused on two of the main activities supporting the Urban growth: logistics and enterprises. He is presently director of the ICE Lab. Author of more than 100 conference and journal papers, in 2019 funded Arisk, one of the main Fintech companies in Italy working on

the prediction of companies' business interruption with AI, Spinoff of Politecnico di Torino.



**MARIANGELA ROSANO** was born in Turin, Italy, in 1987.

She received the Ph.D. degree in Computer and Control Engineer from Politecnico di Torino, Turin, Italy in 2019, and the M.S. degree in Industrial engineering and management from the same university in 2015.

She is presently research assistant at the ICT for City Logistics and Enterprises (ICE) center of Politecnico di Torino and professor of Logistic Trade in the University of Turin.

From 2009 to 2014, she was consultant on the management and monitoring of health and safety at work for several small-medium enterprises in Turin.

In 2017 she joined as visiting researcher at the Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport (CIRRELT), Montréal, Canada.

Her research interests include City Logistics, business development, and business models and operations management in the last-mile logistics for Smart Cities applications.



**Qu Wei** received the master's degree from the Sichuan University of China in 2018.

He is currently pursuing the Ph.D. degree with Politecnico di Torino, Italy. His current research interests include optimization method, intelligent transportation, and computational intelligence.