

Model and exact solution for a two-echelon inventory routing problem

Katyanne Farias

University of Grenoble Alpes, CNRS, Grenoble INP, G-SCOP, F-38000 Grenoble, France
katyanne.farias-de-araujo@grenoble-inp.fr

Khaled Hadj-Hamou

University of Lyon, INSA Lyon, DISP, 69621 Villeurbanne, France
khaled.hadj-hamou@insa-lyon.fr

Claude Yugma

Ecole des Mines de Saint-Etienne, CNRS LIMOS Gardanne, France
claude.yugma@emse.fr

The classic version of the Inventory Routing Problem considers a system with one supplier that manages the stock level of a set of customers. The supplier defines when and how much products to supply and how to combine customers in routes while minimizing storage and transportation costs. We present a new version of this problem that considers a two-echelon system with indirect deliveries and routing decisions at both levels. In this variant, the products are delivered to customers through distribution centers to meet demands of customers with minimum total cost, where the total cost is composed by ordering costs, inventory costs, vehicle fixed costs and transportation costs. We introduce a mathematical formulation and a branch-and-cut algorithm to solve the proposed Two-Echelon Inventory Routing Problem for different inventory policies and routing configurations. Intrinsic new valid inequalities to the two-echelon system are introduced. We analyze the impact of the new inequalities, as well as the already known valid inequalities on the efficiency of the proposed algorithm. Computational experiments are presented for a set of randomly generated instances. The results show that, for the simplest variant of the problem, the proposed method is able to solve medium-scale instances up to the problem optimality.

Key words: Inventory routing, two-echelon, valid inequalities, branch-and-cut

1. Introduction

Large companies tend to manage their activities in a coordinated way. Accordingly, an integrated supply chain has proven to have a major impact on overall success in terms of cost, quality and service time. The vast majority of company activities are interconnected and affect the operation of the others companies that belong to the same supply chain. In this context, two logistics operations are referred in the literature as key points to achieve an efficient supply chain: inventory and transportation. The coordination of these two operations is often known as the Inventory Routing Problem (IRP) (see Coelho et al. (2013) for an brief overview of IRPs). The IRP aims at determining simultaneously the delivery quantities and the routes in order to minimize the inventory holding and transportation costs.

Due to the growth of e-commerce and the imposition of regulatory traffic restrictions in urban areas, a new supply dynamics has become necessary in recent years. According to Crainic et al. (2009), city logistics appears to be one of the main points for managing the urban distribution and for supporting economic, environmental and social aspects. Thus, consolidation facilities, referred to as Distribution Centers (DCs), are introduced to coordinate freight flows inside and outside the urban areas. The DCs, usually located on the outskirts of the cities, receive products from a supplier through large vehicles, stock these products and deliver them with smaller vehicles to the final customers located in urban areas.

In this context, the adoption of DCs has proven to be an effective alternative to meet the deadline of customers and reduce traffic, pollution and noise in urban areas. Trentini et al. (2015) present several deployments of DC in Europe. According to Interface Transport (2004), in the cities of Monaco and La Rochelle, France, the use of DCs contributed to a reduction of about 38% and 48% in the emission of polluting gases and 40% and 61% in the noise levels, respectively.

In this work, we introduce a Two-Echelon Inventory Routing Problem (2E-IRP), which is an extension of the IRP with a finite horizon. In this problem, a set of customers must be served by a supplier strictly through DCs and routes must be defined in both echelons over a given time horizon. We address some variants of 2E-IRP in the context of replenishment policies and routing configuration. The following replenishment policies are modeled: Maximum Level (ML) and Order-Up-To Level (OU) policies, often considered in the IRPs; and we also consider the Order Fixed Quantity (OFQ), apparently not yet considered for this class of problems. In the ML policy, there are no additional constraints on the delivery quantities, provided that the capacities are respected. In the OU policy, whenever a DC or a customer is replenished its inventory level is raised to its maximum capacity. In the OFQ policy, a fixed delivery quantity must be defined for each echelon. The OFQ policy seems interesting because in practice, the inventory logistic management can be simplified when the delivery quantity is known whenever a replenishment is required.

We also consider two other routing configurations in addition to the multi-vehicle case: the single vehicle configuration (1-vehicle) and a configuration that allows multi-tours on the second echelon (∞^2 -vehicle). On the 1-vehicle configuration, one vehicle is available for the supplier and one for each DC. The ∞^2 -vehicle case is applied to urban centers where the traffic of large and/or heavy vehicles is limited. In this case, city centers are usually small and the distances between customers are short, which allows several routes to be made by the same vehicle at the second echelon in a time period.

We propose a Mixed Integer Linear Programming (MILP) formulation for solving the introduced 2E-IRPs variants. In addition, we adapt several known valid inequalities of the IRP and introduce new ones to the 2E-IRP. An exact Branch-and-Cut (B&C) algorithm is also developed to solve

the problems. The efficiency of the proposed methods is evaluated on a set of randomly generated instances. The paper is organized as follows. Section 2 presents a review of relevant works in the context of the 2E-IRP. The 2E-IRP is described and modeled in Section 3, while Section 4 presents adapted and new valid inequalities. The B&C algorithm is introduced in Section 5. The computational results are presented in Section 6. Finally, Section 7 provides the concluding remarks and further research directions.

2. Literature review

Since the 2E-IRP is a new variant of the IRP, we focus on works that deal with: IRPs with focus on classical versions of the problem; Two-Echelon Capacitated Vehicle Routing Problems (2E-CVRPs) considered as a subproblem of the 2E-IRP; and inventory and routing problems on two-echelon systems.

A large variety of IRPs are often addressed in the literature (see, e.g., the survey of Coelho et al. 2013). Several heuristics are available for the classical IRP under the OU policy with deterministic demands and a finite time horizon. For this problem, where a minimum inventory level for each customer should be respected, Bertazzi et al. (2002) proposed a two-step heuristic in which a feasible solution is built by an iterative procedure that inserts customers in a route for one preselected period by solving a shortest-path problem. In the second step, the current solution is improved by using a procedure that removes a pair of customers and reinsert it using the first step method.

Coelho et al. (2012) studied a multi-vehicle IRP (MIRP) considering a consistency requirement related to the quality of service. They proposed a matheuristic that applies an Adaptive Large Neighborhood Search (ALNS) in which two subproblems are solved by MILP formulations. The first subproblem concerns the determination of delivery quantities for a given set of routes, while the second one is responsible for determining the best solution after removal and reinsertion operations. Santos et al. (2016) presented a hybrid algorithm based on an iterated local search metaheuristic with a randomized variable neighborhood descent to solve the MIRP under the ML policy. In their proposed approach, an initial solution is created based on replenishment services to prevent stockout. Then, the quantities delivered are increased to expand the number of periods served by the deliveries, as long as the capacities are respected. To reduce inventory costs, they combined the proposed hybrid algorithm with an exact formulation to solve an inventory management subproblem.

Archetti et al. (2007) were the first authors to exactly solve the IRP under the OU policy by using a B&C algorithm. Solyalı and Süral (2011) reformulated this problem using a shortest-path network representation for decisions of customers inventory replenishment and solved it by a B&C algorithm. They proposed a priori tour heuristic to generate an initial solution. It combines

a partial solution obtained by a formulation without routing decisions with routes obtained by solving traveling salesman subproblems. Archetti et al. (2014) compared the performance of several formulations for the MIRP using B&C algorithms with valid inequalities added at the root node. Coelho and Laporte (2014) proposed a B&C algorithm and new valid inequalities to solve the MIRP under a ML policy. For the same problem, Desaulniers et al. (2016) proposed a branch-price-and-cut method and valid inequalities to strengthen the linear relaxation of the model.

Archetti et al. (2018) proposed a B&C algorithm to solve an IRP with pickups and deliveries, where a product has to be picked up from multiple origins and delivered to the customers over a planning horizon. Li et al. (2011) studied an infinite horizon IRP at two levels, considering a system with one supplier, one DC and a set of customers. They proposed a decomposition solution approach, where customers are initially partitioned into disjoint sets by a genetic algorithm and the resulting subproblems are solved separately. In the literature, many problems involving two-echelon systems concerns vehicle routing problems. To better analyze the characteristics of these systems, we extended the literature review on 2E-CVRPs.

The 2E-CVRP have been studied in the recent years, as evidenced in Cuda et al. (2015). According to the authors, the 2E-CVRP involves tactical planning decisions as the routing of freight and the assignment of customers to DCs with the objective of minimizing routing and handling costs. Thus, in a classical version of this problem, demands of each customer must be served by a single depot through DCs. Direct shippings and split deliveries on the second echelon are not allowed. Two different fleets of homogeneous capacitated vehicles are available, one at each echelon. Each DC has a storage capacity and a maximum number of vehicles that can be routed through it. Most of these characteristics are observed in the routing subproblem of the 2E-IRP.

Perboli et al. (2011) introduced a formal definition of a classical version of the 2E-CVRP. The authors proposed a flow-based model, valid inequalities derived from vehicle routing problem formulations, two matheuristics and a B&C to solve the problem. Jepsen et al. (2013), however, indicated that the formulation proposed by Perboli et al. (2011) may not provide correct upper bounds when the solution contains more than two DCs. Thus, they design a relaxation for the 2E-CVRP and proposed a B&C to solve the problem. Crainic et al. (2013) separated the problem into subproblems, one for each echelon, linked by an adjustment of DC loadings. Handling costs were not considered. They proposed a hybrid heuristic that combines a Greedy Randomized Adaptive Search Procedure (GRASP) and a path relinking algorithm. The GRASP is used to generate initial solutions that are improved by a local search procedure. Then, a path between the current and the best solution is built by means of a path relinking procedure.

Santos et al. (2015) presented a branch-and-cut-and-price which relies on a reformulation that overcomes symmetry issues. The formulation is strengthened with several classes of valid inequalities. More recently, Liu et al. (2018) studied a variant of the 2E-CVRP that considers grouping

constraints in the second echelon. In this case, the customers are divided into several disjoint groups and customers from the same group are served by vehicles from the same DC. The authors solved the problem using a B&C with several valid inequalities.

Most of works addressing two-echelon inventory and routing problems deal with a problem with specific characteristics. Jung and Mathur (2007) and Zhao et al. (2008), for example, studied a system with a supplier, a DC and a set of customers, with the aim of identifying a stationary reorder interval for the DC and each customer to minimize ordering, holding and transportation costs. They presented a fixed-partition strategy that defines clusters of customers and reorder intervals as a power-of-two multiple of a base period for customers in a cluster. Jung and Mathur (2007) proposed a heuristic that decomposes the problem into three sub-problems: the customers clustering; the customers sequencing for each subgroup; and the determination of the DC and inventory policies of customers. Zhao et al. (2008), in turn, developed a heuristic algorithm based on a variable neighborhood search.

Nambirajan et al. (2016) studied a system in which multiple manufacturers must supply multi products to a set of customers through a central depot. Thus, the central depot places orders to the manufacturers and assembles the products to supply them to the customers to meet their demands. The authors proposed a integer linear programming formulation and a two steps heuristic to solve the problem. The first step consists of defining the replenishment policy from the manufacturers to the central depot by using a dynamic programming. The second step determines delivering decisions between the central depot and the customers.

Saragih et al. (2018) addressed a two-echelon location inventory routing problem involving one supplier, a set of DCs and a set of customers. The main objective is to decide, based on stochastic demands, the opening of DCs, assignment of customers to opened DCs, order sizes and reorder points for the supplier, DCs and customers. They proposed a two steps heuristic. The first step uses a constructive approach to generate an initial solution and the second step uses a simulate annealing method to improve the solution.

Recently, Guimarães et al. (2019) addressed a multi-depot two-echelon IRP, which is the closest related problem to the one studied by us among those identified in the literature. They studied a system with several suppliers, DCs and customers under the ML and the OU policies. Each DC has a fleet of homogeneous vehicles that pickup products from a supplier and deliver to customers. A vehicle can perform pickup and deliveries in a single time period and there are no routing decisions between suppliers and DCs. To solve this problem, the authors proposed a B&C approach and a matheuristic algorithm that integrates a subproblem formulation with an ALNS heuristic.

The 2E-IRP addressed in this work contains many characteristics of the problems discussed above. For each DC, an IRP is solved to meet the demands of customers. On the first echelon,

an IRP is managed by the supplier, which provides the DCs. A primary and a secondary fleet of homogeneous vehicles are available to supply DCs and customers, respectively. In addition to the ML and OU replenishment policies, we also propose the case with the OFQ policy, since delivery management could be simplified, resulting in lower costs by the use of a fixed quantity. Furthermore, as discussed in Section 1, the construction of DCs to serve customers in a geographic region and to optimize vehicle flow in urban centers has proved to be a current practice.

3. Problem description and mathematical formulation

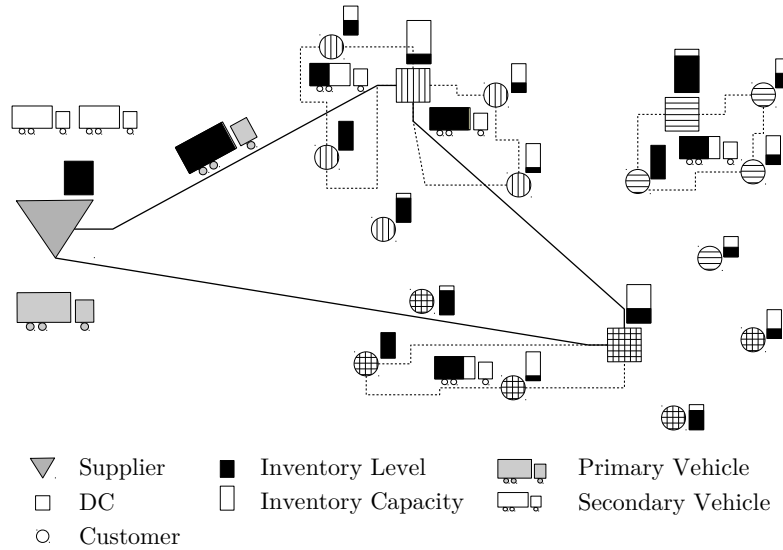
The 2E-IRP is an extension of the classical IRP where there are two echelons $e \in \{1, 2\}$. In the first echelon, the main supplier (represented by $\mathcal{V}_0 = \{0\}$) must meet the requirements of a set $\mathcal{V}_s = \{1, \dots, m\}$ of m DCs that, in turn, must serve a set $\mathcal{V}_c = \{m+1, \dots, m+n\}$ of n customers on the second echelon. Direct deliveries from the supplier to customers are not allowed. Each DC must meet the demands for a single product of a subset of customers, in such way that each DC supplies a subset $V_{c,u} \subseteq V_c$ of customers on the second echelon. The subset of customers assigned to each DC is known. A fleet of κ^1 homogeneous primary vehicles with a large capacity B^1 is available for deliveries from the supplier to DCs. A different fleet of κ^2 homogeneous secondary vehicles with capacity B^2 , where $B^2 \ll B^1$, is available for deliveries from the DCs to the customers. This configuration of vehicles is referred to as multi-vehicle (m-vehicle).

The problem is defined on an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \mathcal{V}_0 \cup \mathcal{V}_s \cup \mathcal{V}_c$ is a vertex set and $\mathcal{E} = \mathcal{E}^1 \cup \mathcal{E}^2$ is the edge set that contains the edges of the first (\mathcal{E}^1) and of the second echelons (\mathcal{E}^2), respectively. The edge set $\mathcal{E}^1 = \{(i, j) | i, j \in \mathcal{V}_0 \cup \mathcal{V}_s, j < i\}$ connects supplier and DCs on the first echelon. The edge set \mathcal{E}^2 is composed by the subsets of edges $\mathcal{E}_u^2 = \{(i, j) | i, j \in \mathcal{V}_{c,u} \cup \{u\}, j < i\}$, for all $u \in \mathcal{V}_s$, between each DC $u \in \mathcal{V}_s$ and its customers. An initial inventory level I_i^0 is associated to each vertex $i \in \mathcal{V}$. Inventory capacities C_i of DCs and customers $i \in \mathcal{V}_s \cup \mathcal{V}_c$ must be respected. The inventory capacity of the supplier is not restrictive. Each customer $i \in \mathcal{V}_c$ has a demand d_i^t at time period $t \in \mathcal{T} = \{1, \dots, \tau\}$. A quantity of product r^t is available at the supplier at time period $t \in \mathcal{T}$. We consider that routes managed by the supplier and by each DC are independent of each other.

Each vertex $(i, j) \in \mathcal{E}$ has a transportation cost c_{ij} associated with. The transportation cost is equal to the distance between two vertices multiplied by a cost coefficient defined per echelon. A unit of inventory held by the supplier, a DC or a customer $i \in \mathcal{V}$ incurs a holding cost h_i per time period $t \in \mathcal{T}'$, where $\mathcal{T}' = \mathcal{T} \cup \{0\}$. In addition, we consider fixed transportation costs f^e for each echelon $e = \{1, 2\}$ proportional to the vehicle capacity. We also consider a fixed order cost g_i for each DC and customer $i \in \mathcal{V}_c \cup \mathcal{V}_s$. A primary vehicle can visit several DCs on a route starting and ending at the supplier, as long as its capacity is respected. Similarly, a second vehicle can visit

several customers of a DC on a route starting and ending at the DC, respecting its capacity. The maximum inventory level at DCs and customers must be respected and no stockout is allowed. The objective of the problem is to minimize the total cost such as the final customer demands are met during each time period. Therefore, delivery quantities and routes at each time period should be defined in both echelons. The problem can be shown to be NP-hard via reduction to the IRP, which is a special case of 2E-IRP when there is only one DC. Figure 1 illustrates the 2E-IRP.

Figure 1 A schema of the 2E-IRP



In the following, we provide an extension of the mathematical formulation proposed by Archetti et al. (2007) and Coelho and Laporte (2013) to the IRP. Let us consider $\mathcal{K}^1 = \{1, \dots, \kappa^1\}$ and $\mathcal{K}^2 = \{1, \dots, \kappa^2\}$ the sets of possible routes starting from the supplier and from DCs per time period, respectively. Let us also define the integer variable $x_{i,j}^{k,t}$ that represents the number of times the edge $(i, j) \in \mathcal{E}^e$ of the echelon $e = \{1, 2\}$ is traversed on the route $k \in \mathcal{K}^e$ at time period $t \in \mathcal{T}$. The binary variable $y_i^{k,t}$ assumes value 1 when the vertex $i \in \mathcal{V}_0 \cup \mathcal{V}_s$ is present on route $k \in \mathcal{K}^1$ or when $i \in \mathcal{V}_s \cup \mathcal{V}_c$ is on route $k \in \mathcal{K}^2$ at time period $t \in \mathcal{T}$, 0 otherwise. The inventory level of the vertex $i \in \mathcal{V}$ at the end of the time period $t \in \mathcal{T}'$ is denoted by the continuous variable I_i^t . Therefore, an inventory level of $i \in \mathcal{V}$ at a time period $t \in \mathcal{T}$ takes into account all products that arrive in i at period $t \in \mathcal{T}$ and that are delivered from i at $t \in \mathcal{T}$ in the case of $i \in \mathcal{V}_0 \cup \mathcal{V}_s$, as well as the demands of i at $t \in \mathcal{T}$ in the case of $i \in \mathcal{V}_c$. The continuous variable $q_i^{k,t}$ denotes the quantity of product delivered to DC $i \in \mathcal{V}_s$ via route $k \in \mathcal{K}^1$ and to customer $i \in \mathcal{V}_c$ via route $k \in \mathcal{K}^2$ at time period $t \in \mathcal{T}$.

Firstly, we consider the ML replenishment policy. It is assumed that the time periods have the same duration at both echelons and routes start and finish in the same period. Hence, the 2E-IRP with ML replenishment policy can be formulated as follows:

$$\begin{aligned} \min \quad & \sum_{k \in \mathcal{K}^1} \sum_{t \in \mathcal{T}} f^1 y_0^{k,t} + \sum_{u \in \mathcal{V}_s} \sum_{k \in \mathcal{K}^2} \sum_{t \in \mathcal{T}} f^2 y_u^{k,t} + \sum_{e \in \{1,2\}} \sum_{(i,j) \in \mathcal{E}^e} \sum_{k \in \mathcal{K}^e} \sum_{t \in \mathcal{T}} c_{i,j} x_{i,j}^{k,t} \\ & \sum_{u \in \mathcal{V}_s} \sum_{k \in \mathcal{K}^1} \sum_{t \in \mathcal{T}} g_u y_u^{k,t} + \sum_{i \in \mathcal{V}_c} \sum_{k \in \mathcal{K}^2} \sum_{t \in \mathcal{T}} g_i y_i^{k,t} + \sum_{t \in \mathcal{T}'} \sum_{i \in \mathcal{V}} h_i I_i^t \end{aligned} \quad (1)$$

$$\text{subject to } I_0^t = I_0^{t-1} + r^t - \sum_{k \in \mathcal{K}^1} \sum_{u \in \mathcal{V}_s} q_u^{k,t} \quad \forall t \in \mathcal{T} \quad (2)$$

$$I_u^t = I_u^{t-1} + \sum_{k \in \mathcal{K}^1} q_u^{k,t} - \sum_{k \in \mathcal{K}^2} \sum_{i \in \mathcal{V}_{c,u}} q_i^{k,t} \quad \forall u \in \mathcal{V}_s, t \in \mathcal{T} \quad (3)$$

$$I_i^t = I_i^{t-1} + \sum_{k \in \mathcal{K}^2} q_i^{k,t} - d_i^t \quad \forall i \in \mathcal{V}_c, t \in \mathcal{T} \quad (4)$$

$$\sum_{k \in \mathcal{K}^e} q_i^{k,t} \leq C_i - I_i^{t-1} \quad \forall \begin{cases} e = 1, i \in \mathcal{V}_s, t \in \mathcal{T} \\ e = 2, i \in \mathcal{V}_c, t \in \mathcal{T} \end{cases} \quad (5)$$

$$q_i^{k,t} \leq \min\{B^e, C_i\} y_i^{k,t} \quad \forall \begin{cases} e = 1, i \in \mathcal{V}_s, k \in \mathcal{K}^1, t \in \mathcal{T} \\ e = 2, i \in \mathcal{V}_c, k \in \mathcal{K}^2, t \in \mathcal{T} \end{cases} \quad (6)$$

$$\sum_{u \in \mathcal{V}_s} q_u^{k,t} \leq B^1 y_0^{k,t} \quad \forall k \in \mathcal{K}^1, t \in \mathcal{T} \quad (7)$$

$$\sum_{i \in \mathcal{V}_{c,u}} q_i^{k,t} \leq B^2 y_u^{k,t} \quad \forall u \in \mathcal{V}_s, k \in \mathcal{K}^2, t \in \mathcal{T} \quad (8)$$

$$\sum_{\substack{v \in \mathcal{V}_s \cup \mathcal{V}_0 \\ v < u}} x_{u,v}^{k,t} + \sum_{\substack{v \in \mathcal{V}_s \\ v > u}} x_{v,u}^{k,t} = 2y_u^{k,t} \quad \forall u \in \mathcal{V}_s, k \in \mathcal{K}^1, t \in \mathcal{T} \quad (9)$$

$$\sum_{\substack{j \in \mathcal{V}_{c,u} \cup \{u\} \\ j < i}} x_{i,j}^{k,t} + \sum_{\substack{j \in \mathcal{V}_{c,u} \\ j > i}} x_{j,i}^{k,t} = 2y_i^{k,t} \quad \forall u \in \mathcal{V}_s, i \in \mathcal{V}_{c,u}, k \in \mathcal{K}^2, t \in \mathcal{T} \quad (10)$$

$$\sum_{i \in S} \sum_{\substack{j \in S \\ j < i}} x_{i,j}^{k,t} \leq \sum_{i \in S} y_i^{k,t} - y_m^{k,t} \quad \forall \begin{cases} S \subseteq \mathcal{V}_s, m \in S, k \in \mathcal{K}^1, t \in \mathcal{T} \\ u \in \mathcal{V}_s, S \subseteq \mathcal{V}_{c,u}, m \in S, k \in \mathcal{K}^2, t \in \mathcal{T} \end{cases} \quad (11)$$

$$\sum_{k \in \mathcal{K}^e} y_i^{k,t} \leq 1 \quad \forall \begin{cases} e = 1, i \in \mathcal{V}_s, t \in \mathcal{T} \\ e = 2, i \in \mathcal{V}_c, t \in \mathcal{T} \end{cases} \quad (12)$$

$$\sum_{k \in \mathcal{K}^1} y_0^{k,t} \leq \kappa^1 \quad \forall t \in \mathcal{T} \quad (13)$$

$$\sum_{u \in \mathcal{V}_s} \sum_{k \in \mathcal{K}^2} y_u^{k,t} \leq \kappa^2 \quad \forall t \in \mathcal{T} \quad (14)$$

$$x_{i,j}^{k,t} \in \{0, 1\} \quad \forall \begin{cases} i, j \in \mathcal{V}_s, j < i, k \in \mathcal{K}^1, t \in \mathcal{T} \\ u \in \mathcal{V}_s, i, j \in \mathcal{V}_{c,u}, j < i, k \in \mathcal{K}^2, t \in \mathcal{T} \end{cases} \quad (15)$$

$$x_{i,u}^{k,t} \in \{0, 1, 2\} \quad \forall \begin{cases} u \in \mathcal{V}_s, i \in \mathcal{V}_s, k \in \mathcal{K}^1, t \in \mathcal{T} \\ u \in \mathcal{V}_s, i \in \mathcal{V}_{c,u}, k \in \mathcal{K}^2, t \in \mathcal{T} \end{cases} \quad (16)$$

$$y_i^{k,t} \in \{0, 1\} \quad \forall \begin{cases} i \in \mathcal{V}_0 \cup \mathcal{V}_s, k \in \mathcal{K}^1, t \in \mathcal{T} \\ i \in \mathcal{V}_s \cup \mathcal{V}_c, k \in \mathcal{K}^2, t \in \mathcal{T} \end{cases} \quad (17)$$

$$q_i^{k,t} \geq 0 \quad \forall \begin{cases} i \in \mathcal{V}_s, k \in \mathcal{K}^1, t \in \mathcal{T} \\ i \in \mathcal{V}_c, k \in \mathcal{K}^2, t \in \mathcal{T} \end{cases} \quad (18)$$

$$I_i^t \geq 0 \quad \forall i \in \mathcal{V}, t \in \mathcal{T}' \quad (19)$$

The objective function (1) is to minimize the sum of fixed primary and secondary vehicle costs, transportation costs at both echelons, order costs of DCs and customers and inventory holding costs at both echelons. Constraints (2) define the inventory level at the supplier at each time period, taking into account quantities delivered to DCs and the products available at the supplier. Similarly, constraints (3) determine for each DC the inventory over the time horizon, based on the products provided by the supplier and quantities delivered to the subset of customers assigned to the DC. Constraints (4) define the inventory level of the customers at each time period. Constraints (5) impose that the quantities delivered to each DC and customer do not exceed their remaining capacities at each time period. Constraints (6) link the quantity delivery variables to routing variables in both echelons. Constraints (7) and (8) reinforce that quantities delivered by a route at the first and second echelon respect the primary and secondary vehicles capacity, respectively. Constraints (9)–(11) are degree and subtour elimination constraints. Constraints (12) ensure that each DC and customer is visited at most once per time period. The number of primary and secondary vehicles available is respected through constraints (13) and (14), respectively. Constraints (15)–(17) and (18)–(19) impose the integrality and non-negative conditions of the variables.

3.1. Problem variants

We model some variants of the 2E-IRP by considering particular replenishment policies and routing configurations. In addition to the ML policy, we consider the OU and the OFQ policies. In order to consider the OU policy applied to both echelons, constraints (20) must be considered in the model (1)–(19). In the OFQ policy, the quantities delivered must be fixed by echelon. Thus, all quantities delivered to DCs and to customers must be equal to a fixed quantity to be defined for the first and second echelon, Q^1 and Q^2 , respectively. To consider the OFQ policy, constraints (21)–(23) must be added to the formulation (1)–(19).

$$\sum_{k \in \mathcal{K}^e} q_i^{k,t} \geq \sum_{k \in \mathcal{K}^e} C_i y_i^{k,t} - I_i^{t-1} \quad \forall \begin{cases} e = 1, u \in \mathcal{V}_s, t \in \mathcal{T} \\ e = 2, i \in \mathcal{V}_c, t \in \mathcal{T} \end{cases} \quad (20)$$

$$\sum_{k \in \mathcal{K}^e} q_i^{k,t} \leq Q^e \quad \forall \begin{cases} e = 1, u \in \mathcal{V}_s, t \in \mathcal{T} \\ e = 2, i \in \mathcal{V}_c, t \in \mathcal{T} \end{cases} \quad (21)$$

$$\sum_{k \in \mathcal{K}^e} q_i^{k,t} - Q^e \geq C_i \left(\sum_{k \in \mathcal{K}^e} y_i^{k,t} - 1 \right) \quad \forall \begin{cases} e = 1, u \in \mathcal{V}_s, t \in \mathcal{T} \\ e = 2, i \in \mathcal{V}_c, t \in \mathcal{T} \end{cases} \quad (22)$$

$$Q^e \geq 0 \quad \forall e \in \{1, 2\} \quad (23)$$

Regarding the availability and use of vehicles, two other cases are considered: the 1-vehicle, in which one vehicle is available for the supplier and one for each DC, and the multi-tours ∞^2 -vehicle

case, where secondary vehicles can perform several routes at the same period. In the first case, each vehicle is able to perform at most one tour per time period ($\mathcal{K}^1 = \{1\}$ and $\mathcal{K}^2 = \{1\}$). Thus, constraints (13) and (14) can be removed from the model.

In the ∞^2 -vehicle case, we consider that the set of routes \mathcal{K}_u^2 starting at each DC $u \in \mathcal{V}_s$ per time period visits at most $|\mathcal{V}_{c,u}|$ vertices, since it comprises the case in which all customers of the DC are visited at a time period, each one being visited on a different route. As the set of routes on the second echelon for each DC may be different one from each other, the model must consider one set \mathcal{K}_u^2 for each $u \in \mathcal{V}_s$ instead \mathcal{K}^2 . In this case, if the number of secondary vehicles available (κ^2) is less than the number of DCs, only κ^2 DCs can perform routes at each time period. Thus, by introducing a new binary variable K_u^t , for all $u \in \mathcal{V}_s$ and $t \in \mathcal{T}$, assuming value one if at least one route is performed by a secondary vehicle from the DC $u \in \mathcal{V}_s$ at time $t \in \mathcal{T}$, zero otherwise, and by replacing constraints (14) by (24)–(26) the ∞^2 -vehicle policy is modeled.

$$B^2 K_u^t \geq \sum_{k \in \mathcal{K}_u^2} y_u^{k,t} \quad \forall u \in \mathcal{V}_s, t \in \mathcal{T} \quad (24)$$

$$\sum_{u \in \mathcal{V}_s} K_u^t \leq \kappa^2 \quad \forall t \in \mathcal{T} \quad (25)$$

$$K_u^t \in \{0, 1\} \quad \forall u \in \mathcal{V}_s, t \in \mathcal{T} \quad (26)$$

Constraints (24)–(26) together define the bounds on the number of available secondary vehicles per time period.

4. Valid inequalities

In this section we discuss about valid inequalities for the 2E-IRP. Firstly, we adapt to our problem some valid inequalities available for the IRP. Secondly, we devise some new valid inequalities for the 2E-IRP.

4.1. Adapted valid Inequalities from the literature

We adapt to our problem the inequalities (27), (28) and (30) proposed by Archetti et al. (2007) and (29) and (31) presented in Archetti et al. (2014).

$$y_i^{k,t} \leq y_u^{k,t} \quad \forall \begin{cases} u=0, i \in \mathcal{V}_s, k \in \mathcal{K}^1, t \in \mathcal{T} \\ u \in \mathcal{V}_s, i \in \mathcal{V}_{c,u}, k \in \mathcal{K}^2, t \in \mathcal{T} \end{cases} \quad (27)$$

$$x_{i,u}^{k,t} \leq 2y_i^{k,t} \quad \forall \begin{cases} u=0, i \in \mathcal{V}_s, k \in \mathcal{K}^1, t \in \mathcal{T} \\ u \in \mathcal{V}_s, i \in \mathcal{V}_{c,u}, k \in \mathcal{K}^2, t \in \mathcal{T} \end{cases} \quad (28)$$

$$x_{i,u}^{k,t} \leq 2y_u^{k,t} \quad \forall \begin{cases} u=0, i \in \mathcal{V}_s, k \in \mathcal{K}^1, t \in \mathcal{T} \\ u \in \mathcal{V}_s, i \in \mathcal{V}_{c,u}, k \in \mathcal{K}^2, t \in \mathcal{T} \end{cases} \quad (29)$$

$$x_{i,j}^{k,t} \leq y_i^{k,t} \quad \forall \begin{cases} i, j \in \mathcal{V}_s, j < i, k \in \mathcal{K}^1, t \in \mathcal{T} \\ u \in \mathcal{V}_s, i, j \in \mathcal{V}_{c,u}, j < i, k \in \mathcal{K}^2, t \in \mathcal{T} \end{cases} \quad (30)$$

$$x_{i,j}^{k,t} \leq y_j^{k,t} \quad \forall \begin{cases} i, j \in \mathcal{V}_s, j < i, k \in \mathcal{K}^1, t \in \mathcal{T} \\ u \in \mathcal{V}_s, i, j \in \mathcal{V}_{c,u}, j < i, k \in \mathcal{K}^2, t \in \mathcal{T} \end{cases} \quad (31)$$

Constraints (27) enforce, on the first echelon, the supplier to be part of a route at a time period if at least one DC is visited in the same route and period. On the second echelon, constraints (27) impose a DC to be part of a route if at least one of its customers is in the route at the same period. Constraints (28)–(31) strengthen the link between route and assignment variables. Constraints (28) enforce that if a DC or customer is the first or last in a route at a time period, the corresponding assignment variable must be equal to one. Constraints (29) are similar, but applied to assignment variables of depots, i.e., the respective DC on the second echelon and the supplier on the first echelon. In fact, constraints (28) and (29) are weaker forms of the degree constraints with only one x variable in the left-hand-side. Constraints (30) and (31) are equivalent to the subtour elimination constraints (11) for pairs of nodes and link route variables between two DCs or two customers to their respective assignment variable.

For vehicle/route indexed formulations with a homogeneous vehicle fleet, Coelho and Laporte (2013) introduced the following symmetry breaking constraints, which can also be adapted to the 2E-IRP.

$$y_u^{k,t} \leq y_u^{k-1,t} \quad \forall \begin{cases} u = 0, k \in \mathcal{K}^1 \setminus \{1\}, t \in \mathcal{T} \\ u \in \mathcal{V}_s, k \in \mathcal{K}^2 \setminus \{1\}, t \in \mathcal{T} \end{cases} \quad (32)$$

$$y_u^{k,t} \leq \sum_{\substack{j \in \mathcal{V}_s \\ v < u}} y_v^{k-1,t} \quad \forall u \in \mathcal{V}_s, k \in \mathcal{K}^1 \setminus \{1\}, t \in \mathcal{T} \quad (33)$$

$$y_i^{k,t} \leq \sum_{\substack{j \in \mathcal{V}_{c,u} \\ j < i}} y_j^{k-1,t} \quad \forall u \in \mathcal{V}_s, i \in \mathcal{V}_{c,u}, k \in \mathcal{K}^2 \setminus \{1\}, t \in \mathcal{T} \quad (34)$$

Constraints (32) impose that a route k from the supplier on the first echelon, or from a DC on the second echelon, is used only if all possible routes with indexes less than k are used. Similarly, constraints (33) and (34) are applied to DCs on the first echelon and to customers on the second echelon. Coelho and Laporte (2014) presented constraints to impose a minimum number of necessary deliveries to meet demands of a time window, which are valid for the second echelon in our problem. These constraints are:

$$\sum_{t_1 \leq t \leq t_2} \sum_{k \in \mathcal{K}^2} y_i^{k,t} \geq \left\lceil \frac{\sum_{t_1 \leq t \leq t_2} d_i^t - C_i}{\min\{B^2, C_i\}} \right\rceil \quad \forall i \in \mathcal{V}_c, t_1, t_2 \in \mathcal{T}, t_2 \geq t_1 \quad (35)$$

$$\sum_{t_1 \leq t \leq t_2} \sum_{k \in \mathcal{K}^2} y_i^{k,t} \geq \frac{\sum_{t_1 \leq t \leq t_2} d_i^t - I_i^{t_1-1}}{\min\{B^2, C_i\}} \quad \forall i \in \mathcal{V}_c, t_1, t_2 \in \mathcal{T}, t_2 \geq t_1 \quad (36)$$

$$\sum_{t_1 \leq t \leq t_2} \sum_{k \in \mathcal{K}^2} y_i^{k,t} \geq 1 - \frac{I_i^{t_1-1}}{\sum_{t_1 \leq t \leq t_2} d_i^t} \quad \forall i \in \mathcal{V}_c, t_1, t_2 \in \mathcal{T}, t_2 \geq t_1 \quad (37)$$

Constraints (35) enforce, for each customer, that if the sum of the demands from period t_1 to t_2 are greater than or equal to its capacity, then the number of deliveries to be made to this

customer in this time window is at least equal to the number of times that its inventory capacity must be reached to meet these demands. In constraints (36), the right hand side is tightened by considering the inventory level at the preceding time period. Constraints (35) and (36), in turn, are strengthened by considering the delivery quantities limited by the minimum between the inventory and secondary vehicle capacity.

Constraints (37) define that at least one visit to a customer is necessary in the interval $[t_1, t_2]$ if its inventory level at the preceding time is not sufficient to meet the customer demands over $[t_1, t_2]$. An adjustment was made in constraints (36) and (37) compared to the ones presented by Coelho and Laporte (2014). We consider the inventory variables at time $t - 1$ instead of time t , as in our proposed formulation, the inventory level is updated at the end of the time period. Thus, the inventory in stock at time $t - 1$ is available to meet the demands at time t and the inventory level at t takes the demand of period t into account. Only the right hand side of constraints (35) can be rounded up because there are no variables on the right-hand side. Based on the ideas considered in constraints (35)–(37), we propose new valid inequalities to the 2E-IRP. These new valid inequalities are applied to the first echelon and are discussed in Section 4.2.

4.2. New valid inequalities for the 2E-IRP

In this section we present new valid inequalities that explore features of the 2E-IRP. Constraints (38), (39) and (40) are therefore respectively equivalent to (35), (36) and (37), but applied to the first echelon. To do this, we replace the customers demand by the delivery quantities to customers, that represents DCs demand. Recently, a constraint similar to (39) was presented in Guimarães et al. (2019). Because constraints (40) are nonlinear, we do not consider them in our experiment tests.

$$\sum_{t_1 \leq t \leq t_2} \sum_{k \in \mathcal{K}^1} y_u^{k,t} \geq \frac{\sum_{t_1 \leq t \leq t_2} \sum_{k \in \mathcal{K}^2} \sum_{i \in \mathcal{V}_{c,u}} q_i^{k,t} - C_u}{\min\{B^1, C_u\}} \quad \forall u \in \mathcal{V}_s, t_1, t_2 \in \mathcal{T}, t_2 \geq t_1 \quad (38)$$

$$\sum_{t_1 \leq t \leq t_2} \sum_{k \in \mathcal{K}^1} y_u^{k,t} \geq \frac{\sum_{t_1 \leq t \leq t_2} \sum_{k \in \mathcal{K}^2} \sum_{i \in \mathcal{V}_{c,u}} q_i^{k,t} - I_u^{t_1-1}}{\min\{B^1, C_u\}} \quad \forall u \in \mathcal{V}_s, t_1, t_2 \in \mathcal{T}, t_2 \geq t_1 \quad (39)$$

$$\sum_{t_1 \leq t \leq t_2} \sum_{k \in \mathcal{K}^1} y_u^{k,t} \geq 1 - \frac{I_u^{t_1-1}}{\sum_{t_1 \leq t \leq t_2} \sum_{k \in \mathcal{K}^2} \sum_{i \in \mathcal{V}_{c,u}} q_i^{k,t}} \quad \forall u \in \mathcal{V}_s, t_1, t_2 \in \mathcal{T}, t_2 \geq t_1 \quad (40)$$

We introduce a new group of valid inequalities in the context of the two-echelon configuration. Due to the instance configurations described in Section 6.1, a route is performed on the first echelon only if the inventory level at one or more DCs is not enough to meet the demands of their customers. This is justified, among other factors, by the cost configuration used to generate the instances, in which we assume that inventory holding from the upper to lower echelons is a

value-added operation. Thus, the inventory holding cost of the supplier is lower than inventory holding costs at DCs that, in turn, are lower than at customers. Hence, constraints (41) impose that, if a delivery is made from the supplier to a DC, at least one delivery will be made to one of its customers at a subsequent time period. Constraints (42) apply the same logic to the quantities shipped, but are restricted to the ML policy.

$$\sum_{k \in \mathcal{K}^1} y_u^{k,t} \leq \sum_{\substack{t_1 \in \mathcal{T} \\ t_1 \geq t}} \sum_{k \in \mathcal{K}^2} \sum_{i \in \mathcal{V}_{c,u}} y_i^{k,t_1} \quad \forall u \in \mathcal{V}_s, t \in \mathcal{T} \quad (41)$$

$$\sum_{k \in \mathcal{K}^1} q_u^{k,t} \leq \sum_{\substack{t_1 \in \mathcal{T} \\ t_1 \geq t}} \sum_{k \in \mathcal{K}^2} \sum_{i \in \mathcal{V}_{c,u}} q_i^{k,t_1} \quad \forall u \in \mathcal{V}_s, t \in \mathcal{T} \quad (42)$$

5. Solving method

We propose an exact B&C algorithm combined to a two-phase heuristic to generate initial feasible solutions. The B&C algorithm is described in Section 5.1, while the heuristic method is detailed in Section 5.2.

5.1. Branch-and-Cut algorithm

We develop a B&C algorithm based on the MILP proposed in Section 3 and on some of the valid inequalities discussed in Section 4. Initially, at each node of the Branch-and-Bound (B&B) tree, a Linear Program (LP) is solved. The LP corresponds to the continuous relaxation of the MILP, which is defined by (1)–(14),(18) and (19), without the subtour elimination constraints (11). In addition, the LP contains some of the constraints (20)–(26), depending on the inventory policy and routing configuration considered, and a subset of valid inequalities among (27)–(42). Then, violated subtour elimination constraints with $m = \arg \max_{j \in S} \{y_j^{kt}\}$ are iteratively identified, by using a separation algorithm of Padberg and Rinaldi (1991), and then added to the model. The resulting model is reoptimized until no violated constraints are found. If the incumbent solution is non-integer, and the node Lower Bound (LB) is not greater or equal than the current Upper Bound (UB), branching occurs, otherwise, either the node is pruned because a feasible solution is found or the node LB is greater than the UB. The process iterates until the UB is equal to the tree LB. The subset of valid inequalities in the LP are chosen based on preliminary experiments discussed in Section 6.2. All the selected valid inequalities are included at the root node.

5.2. Two-phase heuristic

We propose a two-phase heuristic to generate initial feasible solutions to the 2E-IRP. Similarly to Solyalı and Süral (2011), the main idea is to simplify routing decisions. In the first phase, we optimally solve a partial problem that does not involve routing decisions, but penalizes if a route is used in the solution. Thus, knowing delivery quantities and time periods for each customer and

each DC obtained by solving the first phase, we solve several TSPs to define the routes in the second phase. This partial formulation (called PF) defines how much and at which time period products will be delivered from the supplier to DCs and from each DC to its customers to meet the demands of customers, respecting inventory and vehicle capacities and the conditions imposed by constraints (12)–(14). Some of the constraints (20)–(26) can be added to the model depending on the inventory policy and routing configuration considered.

In addition, to minimize inventory and order costs, the use of routes is penalized aiming the reduction on the number of routes in a solution. Whenever a route is allocated to the solution, a cost C_{\max} equal to the highest inventory capacity among the ones of the DCs is added to the total cost. This is justified by the impact of adding new routes on the total cost. Preliminary experiments show that penalizing all routes with C_{\max} produces, on average, better results than the ones obtained by penalizing only the routes on the first echelon or by the ones where only fixed transportation costs are considered. Then, a TSP is optimally solved for each situation where a route is required involving deliveries from the supplier to DCs and from one DC to its customers at a specific time period.

$$\begin{aligned}
 \text{(PF)} \quad \min \quad & \sum_{k \in \mathcal{K}^1} \sum_{t \in \mathcal{T}} C_{\max} y_0^{k,t} + \sum_{u \in \mathcal{V}_s} \sum_{k \in \mathcal{K}^2} C_{\max} y_u^{k,t} + \sum_{i \in \mathcal{V}_c} \sum_{k \in \mathcal{K}^2} \sum_{t \in \mathcal{T}} g_i y_i^{k,t} + \\
 & \sum_{u \in \mathcal{V}_s} \sum_{k \in \mathcal{K}^1} \sum_{t \in \mathcal{T}} g_u y_u^{k,t} + \sum_{i \in \mathcal{V}_0 \cup \mathcal{V}_s \cup \mathcal{V}_c} \sum_{t \in \mathcal{T}'} h_i I_i^t
 \end{aligned} \tag{43}$$

subject to (2) – (8), (12) – (14), (17) – (19)

6. Computational experiments

In this section, we detail the computational experiments conducted to evaluate our method and discuss the obtained results.

All experiments are conducted on a computer equipped with an Intel Core i5-6300U, 2.40 GHz and 8 GB RAM running under ubuntu 16.04 LTS operating system and with a time limit of one hour for each execution of the exact method. For the two-phase heuristic, we impose a maximum running time of five minutes. All the proposed methods are implemented in C++ language and we use CPLEX 12.7.1 with one thread as mixed-integer programming solver. For the B&C, we use the default setting in CPLEX for branching variable selection and for the node selection rule (best-bound-first strategy). We use Concorde (Applegate et al. 2007) to solve TSP instances to the optimality.

In Section 6.1 we detail the instances generation procedure. A computational experiment conducted to analyze the effectiveness of inequalities is presented in Section 6.2. And the computational results for each replenishment policy are detailed in Section 6.3.

6.1. Generation of instances

As benchmark instances for the 2E-IRP were not identified in the literature, we generate a set of instances based on the parameters proposed by Archetti et al. (2007). However, it is necessary to adapt some of these parameters to apply the OFQ inventory policy. Moreover, as in our problem, there are two echelons with dependent demands, we consider different parameters and data scales. We suppose that the first echelon is characterized by larger distances and DCs with higher capacities than customers. This configuration is presented in supply chains where a supplier is located far away from its customers. So, in order to reduce costs, there are intermediate depots (DCs) to store products.

A set of 5 test instances are randomly generated for each combination of time horizon $\mathcal{T} \in \{3, 6, 10\}$ and number of DCs $m \in \{3, 4, 5\}$ yielding a total of 45 instances. Instances with a time horizon equal to 3, 6 and 10 are respectively called small, medium and large-scale instances. For each instance, the number of customers for each DC, n_u , $u \in \mathcal{V}_s$, is randomly generated in the interval $[5, 10]$, where $n = \sum_{u \in \mathcal{V}_s} n_u$. The demand d_i^t of customer $i \in \mathcal{V}_c$ at time $t \in \mathcal{T}$ is constant over time, i.e., $d_i^t = d_i$, for all $i \in \mathcal{V}_c$, and is randomly generated in the interval $[50, 100]$. The quantity r^t available at the supplier at each time period $t \in \mathcal{T}$ is equal to the sum of all customers demand ($\sum_{i \in \mathcal{V}_c} d_i$).

We generate the maximum inventory level C_i of the customer i by multiplying its demand d_i by a randomly selected number from a set $\{2, 3\}$ plus a margin defined as the fixed quantity $\lceil (\max_i d_i)/2 \rceil$, where $\max_i d_i$ is the largest demand among all customers. This margin is added to reduce the probability of generating infeasible problems in the case of the inventory OFQ policy is considered. The initial inventory level of customer i is defined as $I_i^0 = (C_i - d_i)$ and the initial inventory level of the supplier is $I_0^0 = \sum_{i \in \mathcal{V}_c} C_i$. The maximum inventory level C_u for each DC $u \in \mathcal{V}_s$ is equal to $\sum_{i \in \mathcal{V}_{c,u}} C_i$. We define $I_u^0 = \lceil C_u/l_u \rceil$, where l_u is randomly selected from $\{3, 6\}$.

For the m-vehicle case, there is one primary vehicle available and the number of secondary vehicles is equal to the number of DCs. As originally proposed by Archetti et al. (2007), the primary vehicle capacity B^1 is equal to $\lceil \frac{3}{2} \sum_{i \in \mathcal{V}_c} d_i \rceil$. However, on the second echelon, we consider a homogeneous fleet of vehicles with $B^2 = \lceil \frac{3}{2} \max_{u \in \mathcal{V}_s} \sum_{i \in \mathcal{V}_{c,u}} d_i \rceil$. When multi-tours are allowed (∞ -vehicle configuration), the capacity of secondary vehicles is set to one-third of the initially generated capacity.

We opt for a spatial configuration where the supplier is located far away from the cities and each city has one DC and a subset of customers. The cities are not overlapped. Therefore, the coordinates (X_i, Y_i) , $i \in \mathcal{V}$, of the supplier, the DCs and the customers are randomly generated in the interval $[0, 10000]$. Each customer is within a radius of 500 from its DC. The Euclidean distances are equal to $\text{int}(\sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2})$. The unit transportation cost for the first

α^1 and second α^2 echelon are generated respectively in the intervals $[0.01, 0.05]$ and $[0.06, 0.10]$. The fixed transportation costs for each echelon (f^1 and f^2) are proportional to the primary and secondary vehicle capacities: $f^1 = [50, 100]$ and $f^2 = [10, 50]$.

Due to the added value of the product as it approaches the end customer, we consider that the inventory holding cost of the supplier is always smaller than the holding costs of DCs, which in turn is always smaller than the holding costs of the customers, i.e., $h_0 < h_u < h_i$, $u \in \mathcal{V}_s$ and $i \in \mathcal{V}_c$. For the instances, we define $h_0 = 0.01$, $h_u = [0.02, 0.10]$ and $h_i = [0.11, 0.20]$. In contrast, the fixed order costs are higher on the first echelon. The order cost g_u for each DC $u \in \mathcal{V}_s$ and the order cost g_i for each customer $i \in \mathcal{V}_c$ are generated as follows: $g_u = 5l_u$ and $g_i = 5l_i$, where l_u and l_i are randomly generated from $\{6, 7, 8, 9, 10\}$ and $\{1, 2, 3, 4\}$, respectively. All generated instances are available at <https://sites.google.com/view/katyannefarias/>.

6.2. Analysis of valid inequalities

To analyze the effectiveness of each family of valid inequalities, we present a detailed experiment on a subset of 10 instances with 3 DCs, in which 5 instances have time horizon equal to 3 and 5 instances have time horizon equal to 6. For each instance, tests are performed for all possible 9 combinations of the 3 inventory policies (ML, OU and OFQ) and the 3 routing configurations (1-vehicle, m-vehicle and ∞^2 -vehicle), with some exceptions. This results in a maximum of 90 tests for each family of valid inequalities. The basic model is defined by (1)–(19) and, depending on the inventory policy and routing configuration considered, some of the constraints (20)–(26).

Table 1 shows a comparison between the basic model and improved models that include one family of valid inequalities. The comparison is made in terms of UB, LB, lower bound at the root node (LB_R) and running time. For each of these characteristics, we give the number of times that the obtained value by the improved model is better ($\#$ improved), equal ($\#$ equaled) or worse ($\#$ worsened) than the ones obtained by the basic model. If an optimal solution is found in both cases, then the LBs are considered equal. The first row shows the number of performed tests ($\#$ tests) for each group of valid inequalities. In some cases, the family of valid inequalities can not be applied to one or more inventory policy or routing configuration, which explains a number of tests smaller than 90.

To observe how much the basic model has been improved or worsened, we compute the average percentage gaps between the values obtained by the improved model (including valid inequalities) and the basic model, in particular the gaps of UB, LB, LB_R and running time. We consider a positive gap when the result of the improved model is better and a negative one otherwise. On the one hand, the percentage gap of UB or running time is the value obtained by the basic model minus the value by the improved model, divided by the value of the basic model. On the other

Table 1 Impact of valid inequalities on the quality of solutions

	Additional Valid Inequalities										
	(27)	(28)	(30)	(29)	(31)	(32)	(33) (34)	(35) (38)	(36) (39)	(41)	(42)
#Tests	90	90	90	90	90	60	60	90	90	90	30
#UB improved	19	21	1	19	1	8	5	14	23	21	5
#UB equaled	52	50	86	56	84	31	28	55	50	51	20
#UB worsened	19	19	3	15	5	21	27	21	17	18	5
AM of UB Gap (%)	0.1	0.3	-1.1	-1.3	-1.1	-22.2	-9.6	-7.7	0.8	-0.6	1.1
#LB improved	30	24	12	35	15	18	12	18	27	27	6
#LB equaled	38	36	68	38	65	17	15	35	36	35	20
#LB worsened	22	30	10	17	10	25	33	37	27	28	4
AM of LB Gap (%)	0.2	-0.2	-0.1	0.3	-0.1	-0.7	-0.9	-0.3	0.2	0.2	0.0
#LB _R improved	62	43	0	39	1	39	36	37	31	53	18
#LB _R equaled	3	3	88	10	86	1	1	5	0	6	4
#LB _R worsened	25	44	2	41	3	20	23	48	59	31	8
AM of LB _R Gap (%)	0.7	0.1	0.0	0.0	0.0	0.3	0.4	-0.3	-0.8	0.6	0.3
#Time improved	26	20	26	24	25	4	5	12	15	13	8
#Time equaled	48	51	52	51	52	42	42	52	51	52	10
#Time worsened	16	19	12	15	13	14	13	26	24	25	12
PM of Time Gap (%)	0.5	-3.1	0.0	-1.2	0.0	-3.0	-4.6	-4.8	-1.2	-1.6	-1.0

hand, the percentage gap of LB or LB_R is the value obtained by the improved model minus the value by the basic model, divided by the value of the basic model.

For time gaps, when an execution of the improved model is interrupted because of memory usage and not because of time limit, we consider a negative time gap of 100%. In this case, the running time of improved model is considered worse than the running time of the basic model. In contrast, if the execution of the basic model is interrupted because of memory limit, we consider a positive time gap of 100%. If the basic model and the compared improved one are, both, interrupted by the one hour running time limit, the execution times are considered equal and the time gap is set to zero.

As costs have approximately the same order of magnitude in the tests, we compute Arithmetic Means (AM) to UB, LB and LB_R gaps. We opt to Pondered Means (PM) for time gaps in order to give more importance to improvement percentages on longer running times. Valid inequalities (32) and (33)–(34) are only applied to the m-vehicle and the ∞^2 -vehicle cases and inequalities (42) are restricted to the ML policy.

Regarding the results in Table 1, it can be noticed that valid inequalities (27) are effective in all evaluated characteristics. We consider that inequalities (28) are effective because they improved and achieved positive gaps of UB and also LB_R in a considerable number of instances, at the expense of an increasing in the running times, on average. The negative value of this time gap is mainly due to 6 tests out of 90 tests. If we disregard these cases, the PM of the time gap becomes

0.20%. For valid inequalities (30) and (31), the improved model has similar results when compared to the basic model. Therefore, we do not consider them in our proposed B&C algorithm. The use of constraints (29) does not improve UBs, on average. An additional test was performed by adding to the model inequalities (28) and (29), since they are similar constraints. However, the use of these two groups of inequalities simultaneously provide a bad AM concerning the UBs.

Constraints (32) and (33)–(34) are not considered in the B&C because they provide, on average, a negative impact on UBs, LBs and on execution times. Inequalities (35) and (38) have the same function as inequalities (36) and (39). As the first two are not effective in all evaluated items, only (36) and (39) are considered in the B&C framework. These inequalities also provide negative means of LB_R and time gaps, but this negative result is softened by the improvement in UB and LB. A negative PM of time gaps also occurs when inequalities (42) are included. On the other hand, they provide, in general, good results in terms of UB, LB and LB_R . Inequalities (41) have a great negative effect on UB gaps for 4 tests. In fact, for two tests the basic model finds feasible solutions, but the improved model with constraints (41) is not able to find a simple feasible solution in the running time limit of one hour. The AM of UB gap is equal to 4.00% if these cases are disregarded.

Accordingly, we insert at the root node of the proposed B&C algorithm valid inequalities (27), (28), (36), (39), (41) and (42). The last 4 groups of inequalities provide a negative mean of the time gap, which is relatively negligible compared to the improvements achieved in terms of UB and LB.

6.3. Computational results

In this section we evaluate the performance of the proposed B&C algorithm on the variants of the 2E-IRP. Figures 2, 3 and 4 present a comparison between the gaps obtained by the B&C algorithm and the MILP basic model for each combination of replenishment policies and routing configurations. In particular, Figure 2 shows the average gaps obtained for small-scale instances.

Both methods are able to solve to the optimality all small-scale instances when 1-vehicle configuration with ML or OU policy and m-vehicle with ML policy are considered. However, for these cases, the B&C algorithm is faster than the basic model in almost all executions. When 1-vehicle and OQF policy are considered, the basic model provides a better average gap than the B&C. In all other cases, the proposed B&C algorithm provides lower gaps. When a more complex replenishment policy and routing configuration are considered, the gaps achieved by both methods have a higher difference.

The average gaps obtained by the B&C algorithm and the MILP basic model for medium-scale instances are shown in Figure 3. In the case of the 1-vehicle configuration, the basic model provides a slightly better performance when compared to other routing configurations. In other cases, the

Figure 2 Average gap (%) for small-scale instances

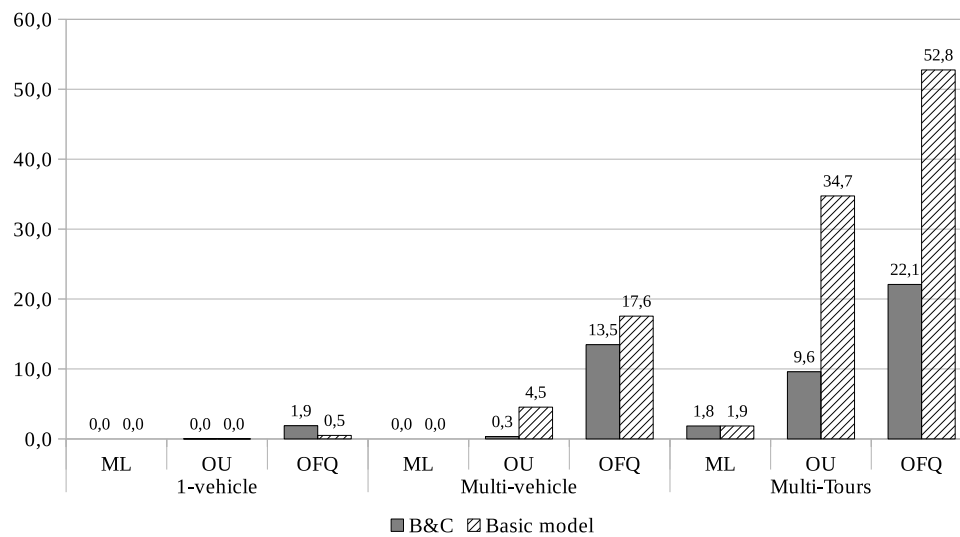
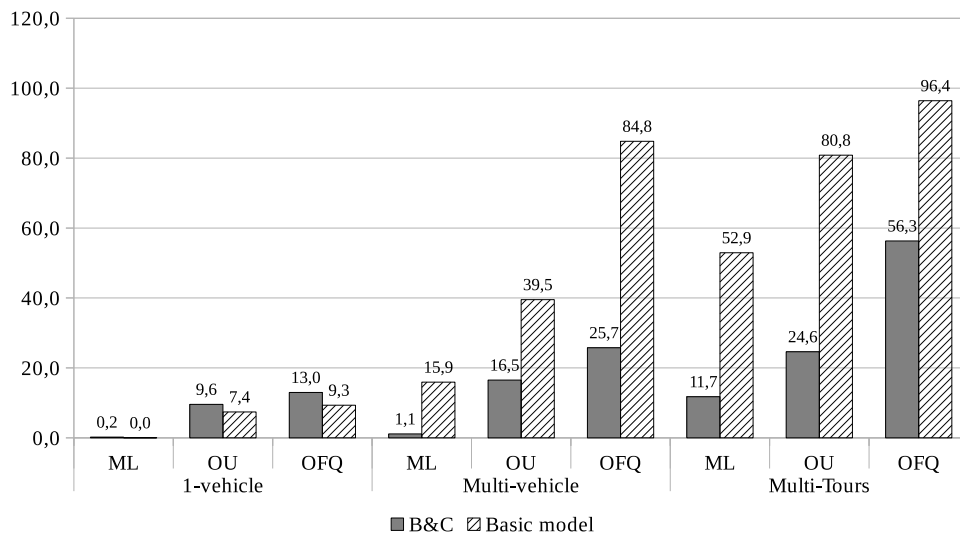


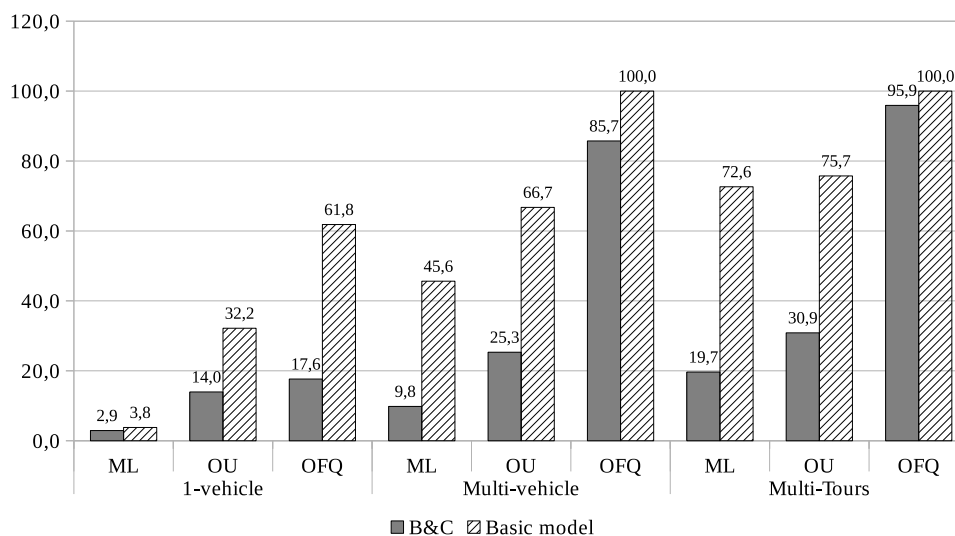
Figure 3 Average gap (%) for medium-scale instances



B&C gives gaps of higher quality. As presented in Figure 4, for large-scale instances, the B&C algorithm always provides better average gaps than the basic model.

For each replenishment policy, we opt to show the simulation results for the m-vehicle and ∞ -vehicle configurations, which tends to be more difficult to solve than the 1-vehicle case because of a larger number of combinatorial parts in variables and constraints. Results for all vehicle configurations are available at <https://sites.google.com/view/katyannefarias/>.

A summary of the computational results for the ML, OU and OFQ policies is presented respectively in Tables 2, 3 and 4. We evaluate a comparison between the proposed B&C and the basic model. We show for each routing configuration and time horizon:

Figure 4 Average gap (%) for large-scale instances

- the number of instances solved to optimality (#Optimal) by the B&C;
- the number of instances where a feasible solution is found (#Feasible) by the B&C;
- the number of instances solved to optimality only by the B&C (#New optimal);
- the number of instances solved with a feasible solution found only by the B&C (#New feasible);
- the number of instances that the B&C does not prove an optimal solution proven by the basic model (#Optimal not proven);
- the number of UB, LB and time improved, equaled or worsened compared to the basic model;
- and average gaps between B&C and basic model results.

The gaps between the B&C and the basic model are calculated as explained in Section 6.2, but also with AMs of the time gap, since these analysis are performed by instance size with running times in the same order of magnitude.

Table 2 shows results for the ML policy. When the time horizon is equal to 3 for the m-vehicle configuration, the proposed B&C algorithm achieves the same results as the basic model, but in a shorter execution time. For the ∞ -vehicle configuration, in small-scale instances, the B&C improves UBs of 6 instances. However, most of the LBs are slightly worse than the ones by the basic model. In general, the B&C provides better solutions when the time horizon is equal to 6 and 10, finding new feasible solutions and improving UBs and LBs for both routing configurations. Negative AMs of time gaps and the majority of worst running times are mainly due to limit memory.

In the case of the OU replenishment policy, for a time horizon equal to 3, the UB and LB are improved on several instances, as shown in Table 3. On these instances, running times are also in average improved for the m-vehicle configuration and, in the ∞ -vehicle configuration, the average time gap is negative due to memory limit. Similarly to the ML policy, for a time horizon of 6 and

Table 2 Summary for ML policy of the B&C compared to the basic model

	m-vehicle config.			∞ -vehicle config.		
	$T=3$	$T=6$	$T=10$	$T=3$	$T=6$	$T=10$
#Optimal	15	3	0	2	0	0
#New optimal	0	1	0	0	0	0
#Optimal not proven	0	2	0	4	0	0
#Feasible	15	15	15	15	15	15
#New feasible	0	2	6	0	4	8
#Feasible not found	0	0	0	0	0	0
#UB improved	0	10	11	6	12	14
#UB equaled	15	4	0	7	0	0
#UB worsened	0	1	4	2	3	1
AM of UB Gap (%)	0.00	14.52	40.57	0.12	47.17	66.86
#LB improved	0	10	7	3	9	7
#LB equaled	15	2	0	2	0	0
#LB worsened	0	3	8	10	6	8
AM of LB Gap (%)	0.00	1.36	1.42	-0.06	0.83	1.35
#Time improved	15	2	0	1	0	0
#Time equaled	0	8	15	0	8	15
#Time worsened	0	5	0	14	7	0
AM of Time Gap (%)	86.01	-57.84	0.00	-267.21	-46.67	0.00

10, feasible solutions not found by the basic model are found by the B&C and, in general, better bounds are achieved.

Table 3 Summary for OU policy of the B&C compared to the basic model

	m-vehicle config.			∞ -vehicle config.		
	$T=3$	$T=6$	$T=10$	$T=3$	$T=6$	$T=10$
#Optimal	9	0	0	0	0	0
#New optimal	1	0	0	0	0	0
#Optimal not proven	1	0	0	0	0	0
#Feasible	15	15	15	15	15	15
#New feasible	0	2	7	0	9	7
#Feasible not found	0	0	0	0	0	0
#UB improved	4	14	14	13	15	15
#UB equaled	11	0	0	2	0	0
#UB worsened	0	1	1	0	0	0
AM of UB Gap (%)	1.07	25.13	56.46	19.14	74.22	64.95
#LB improved	5	14	13	13	15	15
#LB equaled	8	0	0	0	0	0
#LB worsened	2	1	2	2	0	0
AM of LB Gap (%)	3.81	4.70	3.57	14.46	5.24	5.48
#Time improved	9	0	0	0	0	0
#Time equaled	0	14	15	1	15	15
#Time worsened	6	1	0	14	0	0
AM of Time Gap (%)	14.72	-6.67	0.00	-93.33	0.00	0.00

On the m-vehicle configuration, when an optimal solution is found by both methods, the proposed B&C is always faster. Only in one instance with 6 and one with 10 time periods, the basic model provides a better gap. On the ∞ -vehicle configuration, the execution of almost all instances with 3

time periods is limited by the use of memory. However, except in two instances, including instances with 6 and 10 time periods, the B&C provides better gaps than the basic model.

It is important to emphasize that if we compare the performance of the B&C with that of the basic model when adding the OU replenishment policy, we verify that the increase in the gaps obtained by the basic model occurs in a greater proportion than the gaps reached by the B&C algorithm. In other words, the negative impact on the performance of the proposed B&C algorithm when we consider the OU policy is slighter than the impact on the basic model. The same occurs when, given an inventory policy, we compare the m-vehicle with the ∞ -vehicle configuration.

Table 4 shows that, when the OFQ policy is considered the problem seems to be come more difficult to solve. In this case, no optimal solutions are found or proven by the B&C algorithm. However, in all groups of instances, the proposed B&C algorithm is able to find new feasible solutions and to considerably improve the UB for several instances. Most of the best LBs are also improved in the case of the ∞ -vehicle configuration. However, in the case of the m-vehicle configuration, the best LBs have, on average, a lower quality. Even in this configuration, the B&C is not able to find/prove 4 optimal solutions and, for some small-scale instances, the execution is interrupted due to memory limit, which justifies the negative average time gap. In all other cases, both methods reach the running time limit. Similarly to the OU policy, the OFQ policy has a smaller impact on the performance of the B&C than on the performance of the basic model.

Table 4 Summary for OFQ policy of the B&C compared to the basic model

	m-vehicle config.			∞ -vehicle config.		
	$T=3$	$T=6$	$T=10$	$T=3$	$T=6$	$T=10$
#Optimal	0	0	0	0	0	0
#New optimal	0	0	0	0	0	0
#Optimal not proven	4	0	0	0	0	0
#Feasible	15	15	3	15	10	1
#New feasible	1	12	3	3	9	1
#Feasible not found	0	0	0	0	0	0
#UB improved	8	14	3	13	10	1
#UB equaled	1	0	12	0	5	14
#UB worsened	6	1	0	2	0	0
AM of UB Gap (%)	8.64	81.56	20.00	39.99	60.84	6.67
#LB improved	2	7	5	10	12	5
#LB equaled	0	0	0	0	0	0
#LB worsened	13	8	10	5	3	10
AM of LB Gap (%)	-3.26	-2.31	-0.92	1.41	1.54	0.13
#Time improved	0	0	0	0	0	0
#Time equaled	6	15	15	15	15	15
#Time worsened	9	0	0	0	0	0
AM of Time Gap (%)	-146.61	0.00	0.00	0.00	0.00	0.00

7. Conclusions and future work

We proposed a new extension of the IRP that considers a two echelon system, referred to as 2E-IRP, under several replenishment policies and routing configurations. We developed a MILP model and a B&C algorithm to solve the proposed variants of the 2E-IRP. The proposed B&C algorithm is able to solve the 2E-IRP for the 1-vehicle, m -vehicle or ∞^2 -vehicle configuration, and for the ML, OU or OFQ policy. We adapted valid inequalities available for the classical IRP and we also proposed new ones. To analyze the effectiveness of the valid inequalities and the performance of the methods, we generated a set of random instances, and performed an extensive computational experiment. We observed that if replenishment policies or more complex routing configurations are considered in the problem, the impact on the performance of the algorithms is smaller in the B&C than in the model. In addition, the B&C algorithm has presented, on average, better results in terms of objective function cost and running time than the ones by the model.

The OU policy turned out to be very strict, which implies in high costs. Depending on the company needs, an intermediate option is the OFQ policy. Substantial savings can be achieved using the ML policy. With regard to the routing configurations, the multi-tours presented the highest costs among all the tested routing configurations. However, it is important to consider this configuration because, in many urban cities, only small vehicles are allowed to make multi-tours in a period due to the short distances between customers.

As future work, we plan to investigate the use of column generation to solve the LP at each node of the B&B tree (Branch-and-Price) and heuristic methods to solve large instances. Another important direction is to add new features in the problem, such as transshipment between DCs or customers and uncertain data, to better fit with industrial company assumptions and constraints.

Acknowledgement

This work was partly supported by the research program of the region Auvergne-Rhône-Alpes, France: ARC 2016 N°08 – *Industrialisation et science du gouvernement*. This support is greatly appreciated.

References

- Applegate, David L., Robert E. Bixby, Vašek Chvátal, William J. Cook. 2007. *The Traveling Salesman Problem: A Computational Study*. Princeton University Press, Princeton, NJ, USA.
- Archetti, Claudia, Luca Bertazzi, Gilbert Laporte, Maria Grazia Speranza. 2007. A branch-and-cut algorithm for a vendor-managed inventory-routing problem. *Transportation Science* **41**(3) 382–391. doi:10.1287/trsc.1060.0188.
- Archetti, Claudia, Nicola Bianchessi, Stefan Irnich, M. Grazia Speranza. 2014. Formulations for an inventory routing problem. *International Transactions in Operational Research* **21**(3) 353–374. doi:10.1111/itor.12076.

- Archetti, Claudia, Marielle Christiansen, M. Grazia Speranza. 2018. Inventory routing with pickups and deliveries. *European Journal of Operational Research* **268**(1) 314–324. doi:10.1016/j.ejor.2018.01.010.
- Bertazzi, Luca, Giuseppe Paletta, M. Grazia Speranza. 2002. Deterministic order-up-to level policies in an inventory routing problem. *Transportation Science* **36**(1) 119–132. doi:10.1287/trsc.36.1.119.573.
- Coelho, Leandro C., Jean-François Cordeau, Gilbert Laporte. 2012. Consistency in multi-vehicle inventory-routing. *Transportation Research Part C: Emerging Technologies* **24** 270–287. doi:10.1016/j.trc.2012.03.007.
- Coelho, Leandro C., Jean-François Cordeau, Gilbert Laporte. 2013. Thirty years of inventory routing. *Transportation Science* **48**(1) 1–19. doi:10.1287/trsc.2013.0472.
- Coelho, Leandro C., Gilbert Laporte. 2013. The exact solution of several classes of inventory-routing problems. *Computers & Operations Research* **40**(2) 558–565. doi:10.1016/j.cor.2012.08.012.
- Coelho, Leandro C., Gilbert Laporte. 2014. Improved solutions for inventory-routing problems through valid inequalities and input ordering. *International Journal of Production Economics* **155**(Supplement C) 391–397. doi:10.1016/j.ijpe.2013.11.019.
- Crainic, Teodor Gabriel, Simona Mancini, Guido Perboli, Roberto Tadei. 2013. *GRASP with Path Relinking for the Two-Echelon Vehicle Routing Problem*. Springer New York, New York, NY, 113–125. doi:10.1007/978-1-4614-6322-1.7.
- Crainic, Teodor Gabriel, Nicoletta Ricciardi, Giovanni Storchi. 2009. Models for evaluating and planning city logistics systems. *Transportation Science* **43**(4) 432–454. doi:10.1287/trsc.1090.0279.
- Cuda, R., G. Guastaroba, M.G. Speranza. 2015. A survey on two-echelon routing problems. *Computers & Operations Research* **55** 185–199. doi:0.1016/j.cor.2014.06.008.
- Desaulniers, Guy, Jørgen G. Rakke, Leandro C. Coelho. 2016. A branch-price-and-cut algorithm for the inventory-routing problem. *Transportation Science* **50**(3) 1060–1076. doi:10.1287/trsc.2015.0635.
- Guimarães, Thiago A., Leandro C. Coelho, Cleder M. Schenekemberg, Cassius T. Scarpin. 2019. The two-echelon multi-depot inventory-routing problem. *Computers & Operations Research* **101** 220–233. doi:10.1016/j.cor.2018.07.024.
- Interface Transport. 2004. Espaces logistiques urbains de monaco et la rochelle : Analyse comparative des systèmes logistiques rochelais et monégasques. Tech. rep., Agence De l'Environnement et de la Maîtrise de l'Énergie.
- Jepsen, Mads, Simon Spoorendonk, Stefan Ropke. 2013. A branch-and-cut algorithm for the symmetric two-echelon capacitated vehicle routing problem. *Transportation Science* **47**(1) 23–37. doi:10.1287/trsc.1110.0399.
- Jung, Jaeheon, Kamlesh Mathur. 2007. An efficient heuristic algorithm for a two-echelon joint inventory and routing problem. *Transportation Science* **41**(1) 55–73. doi:10.1287/trsc.1060.0160.

- Li, Jianxiang, Feng Chu, Haoxun Chen. 2011. A solution approach to the inventory routing problem in a three-level distribution system. *European Journal of Operational Research* **210**(3) 736–744. doi:10.1016/j.ejor.2010.10.020.
- Liu, Tian, Zhixing Luo, Hu Qin, Andrew Lim. 2018. A branch-and-cut algorithm for the two-echelon capacitated vehicle routing problem with grouping constraints. *European Journal of Operational Research* **266**(2) 487–497. doi:10.1016/j.ejor.2017.10.017.
- Nambirajan, Ramkumar, Abraham Mendoza, Subramanian Pazhani, T.T. Narendran, K. Ganesh. 2016. Care: Heuristics for two-stage multi-product inventory routing problems with replenishments. *Computers & Industrial Engineering* **97** 41–57. doi:10.1016/j.cie.2016.04.004.
- Padberg, Manfred, Giovanni Rinaldi. 1991. A branch-and-cut algorithm for the resolution of large-scale symmetric traveling salesman problems. *SIAM Review* **33**(1) 60–100. doi:10.1137/1033004.
- Perboli, Guido, Roberto Tadei, Daniele Vigo. 2011. The two-echelon capacitated vehicle routing problem: Models and math-based heuristics. *Transportation Science* **45**(3) 364–380. doi:10.1287/trsc.1110.0368.
- Santos, Edcarlos, Luiz Satoru Ochi, Luidi Simonetti, Pedro Henrique González. 2016. A hybrid heuristic based on iterated local search for multivehicle inventory routing problem. *Electronic Notes in Discrete Mathematics* **52**(Supplement C) 197–204. doi:10.1016/j.endm.2016.03.026.
- Santos, Fernando Afonso, Geraldo Robson Mateus, Alexandre Salles da Cunha. 2015. A branch-and-cut-and-price algorithm for the two-echelon capacitated vehicle routing problem. *Transportation Science* **49**(2) 355–368. doi:10.1287/trsc.2013.0500.
- Saragih, Nova Indah, Senator Nur Bahagia, Suprayogi, Ibnu Syabri. 2018. A heuristic method for location-inventory-routing problem in a three-echelon supply chain system. *Computers & Industrial Engineering* doi:10.1016/j.cie.2018.11.026.
- Solyali, Oğuz, Haldun Süral. 2011. A branch-and-cut algorithm using a strong formulation and an a priori tour-based heuristic for an inventory-routing problem. *Transportation Science* **45**(3) 335–345. doi:10.1287/trsc.1100.0354.
- Trentini, Anna, Jesus Gonzales Feliu, Nicolas Malhéné. 2015. Developing urban logistics spaces: Ucc and pls in south-western europe. Working paper or preprint (halshs-01214749).
- Zhao, Qiu-Hong, Shuang Chen, Cun-Xun Zang. 2008. Model and algorithm for inventory/routing decision in a three-echelon logistics system. *European Journal of Operational Research* **191**(3) 623–635. doi:10.1016/j.ejor.2006.12.056.