E-commerce shipping through a third-party supply chain

Diego Ponce[a,b,c], Ivan Contreras[a,b], Gilbert Laporte[b,d]

[a]*Concordia University, Montréal, Canada*
[b]*Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT), Montréal, Canada*
[c]*Institute of Mathematics of the University of Seville (IMUS), Sevilla, Spain*
[d]*HEC Montréal, Montréal, Canada*

Abstract

We consider an e-commerce retailer who must ship orders from a warehouse to a set of customers with delivery deadlines. As is often the case, the retailer uses a third-party service provider to ensure its distribution. The retailer can enter the supply chain of the service provider at various levels. Entering it at a higher level entails lower sorting costs for the retailer, but higher delivery costs, and longer delivery times. The customer orders arrive at various moments over a rolling planning horizon. This means that the retailer must also make consolidation decisions. We model and solve the static and dynamic cases of this problem. The static case is modeled as an integer linear program and solved by CPLEX. We develop and compare four shipping policies for the dynamic case. Extensive computational results based on real location data from California and Texas are reported.

Keywords: e-commerce, third-party logistics, supply chain, dynamic shipping policies

1. Introduction

E-commerce has known a phenomenal growth in recent years, with worldwide annual sales of 1,336 trillion USD in 2014, 2,842 trillion USD in 2018, and a projected figure of 4,878 trillion USD for 2021 (Clement, 2019). In several parts of the world the annual growth lies between 10% and 20% (Chen, 2017). The three main e-retailers are currently Amazon, Alibaba and JingDong (https://www.top10onlinebuy.com/top-10-ecommerce-companies-in-the-world-2018/). The success of e-commerce is due to the development and integration of innovative and efficient practices in marketing, information systems and logistics (see, e.g., Tang and Veelenturf, 2019). For reviews of successful e-commerce practices and perspectives, see Yu et al. (2016) and Lafkihi et al. (2019).

Amazon is responsible for several of the logistics innovations now observed in e-commerce (Cohen, 2018). Two delivery options commonly used by e-retailers are drop-shipping by which customers' orders are fulfilled by manufacturers (see, e.g., Ayanso et al., 2006; Yu et al., 2017; Chen et al., 2018; Li et al., 2019), and crowd-shipping where last-mile deliveries are performed by members of the public (see, e.g., Buldeo Rai et al., 2017; Le et al., 2019; Cheng et al., 2019; Macrina et al., 2020). A practice employed by Amazon is to subcontract the delivery of several of their orders to a third-party logistics provider (3PL), such as FedEx, UPS, DHL, or

Email addresses: dponce@us.es (Diego Ponce), icontrer@encs.concordia.ca (Ivan Contreras), gilbert.laporte@cirrelt.net (Gilbert Laporte)

some of the national postal services. The need for Amazon to make use of 3PL providers is particularly acute in peak periods such as the end-of-year holiday season and Prime Day in July. During these peak periods, Amazon deals with a wide range of distributors to ensure speedy delivery. Smaller operators who do not possess the necessary infrastructure to perform their own deliveries, also rely on 3PL providers on a regular basis.

We consider a general setting in which customers place orders with strict delivery deadlines to the e-retailer over a planning horizon consisting of several periods. At the end of each period the retailer must decide for each undelivered order whether to keep it on inventory for at least one extra period or to ship it immediately. In the latter case it must decide at which level it should enter the supply chain of the 3PL. As usual, the supply chain consists of several levels containing sorting centers (SCs) and final delivery points (FDPs). The FDPs at level 0 are the final points of the supply chain considered in this study (Figure 1). The SCs at level 1 can be post offices or retail outlets which deliver to the customers or to which the customers go (see, e.g., Landete and Laporte, 2019). The retailer has the option of entering the supply chain at any SC, but we exclude the possibility of making direct deliveries to the FDPs without loss of generality. We assume that the retailer can adopt a mixed policy in the sense that depending on the customers, the supply chain of the 3PL may be entered at different levels. Note, however, that such a policy is customer related and not product related. The higher the level at which the retailer enters the supply chain, the lower are the sorting costs for the retailer, the higher are its delivery costs, and the longer are the delivery times to the FDPs.

1.1. Scientific contribution

We call the problem the E-commerce Shipping Problem (ESP). We will successively consider 1) the static ESP in which all orders pertaining to a fixed planning horizon are known at the beginning of the horizon and must be delivered no later than the end of the planning horizon, and 2) the dynamic ESP in which the orders arrive in real time over a rolling planning horizon. In both cases, orders can be consolidated: i.e., at the end of each period a decision must be made whether to keep an order on inventory for additional periods, or to deliver it by entering the 3PL's supply chain at an appropriate level, considering the costs and the need to meet the customers' deadlines. We model the static case as an integer linear program (ILP) which is solved by CPLEX. We also develop and compare four shipping policies for the dynamic case.

1.2. Positioning of our study within the relevant literature

The basic problem underlying our study is the Periodic Vehicle Routing Problem (PVRP) which is defined over a fixed planning horizon consisting of several periods. In the static PVRP, at the beginning of the planning horizon a list of orders delivery deadlines arriving at each period is provided. The problem is then to decide when to deliver them to customers and how to plan the routing (see, e.g., Francis et al., 2008; Hemmelmayr et al., 2009; Archetti et al., 2015; Dayarian et al., 2015). In the dynamic PVRP, the orders are gradually revealed over a rolling planning horizon (see, e.g., Angelelli et al., 2009; Wen et al., 2010; Albareda-Sambola et al., 2014; Ulmer et al., 2018).

At the heart of the multi-period routing problem lies the question of freight consolidation. This issue calls for an equilibrium between inventory costs and freight costs, namely in contexts where lower unit transportation costs can be achieved through full vehicle shipments or larger shipments (see, e.g., Tyan et al., 2003; Satır et al., 2018; Hanbazazah et al., 2019). The need to make more frequent shipments arises namely in perishable goods logistics (Hu et al., 2018).

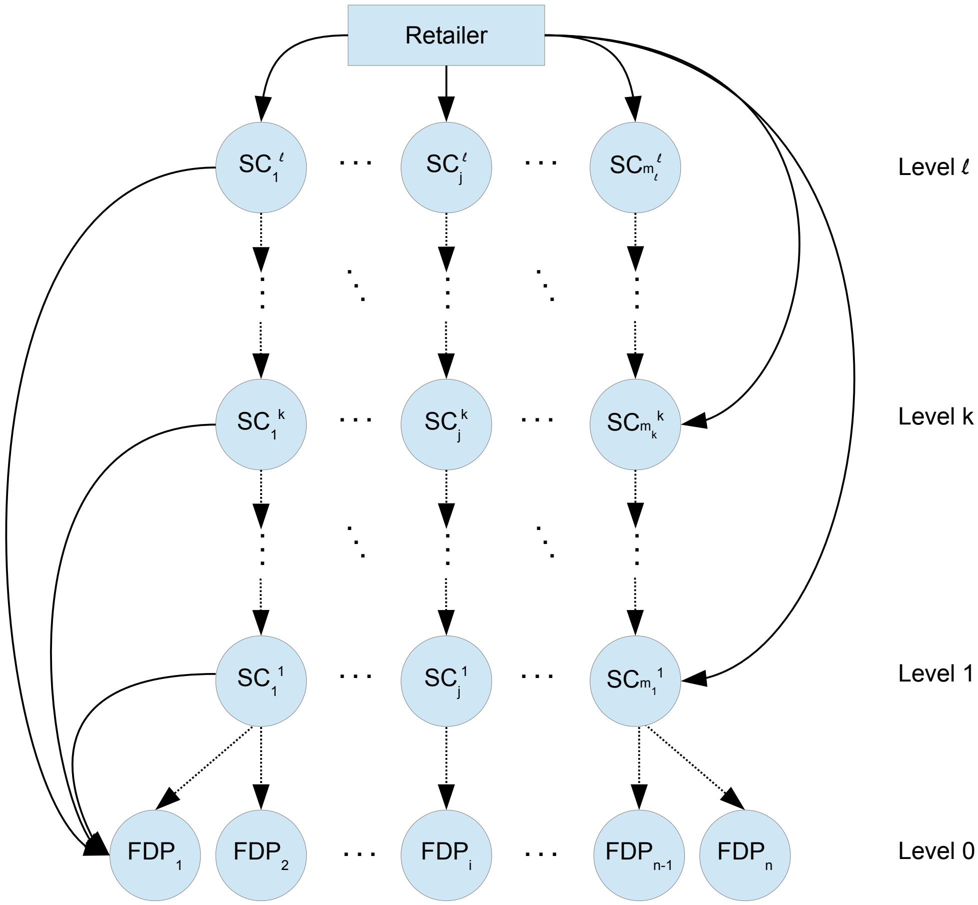


Figure 1: General supply chain with ℓ levels, m_k SCs at level each level k ($k = 1, \dots, \ell$), and n FDPs at level 0, showing only some of the potential arcs.

However, fewer shipments are preferable in order to reduce the amount of polluting emissions (Stenius et al., 2018; Ma and Tan, 2019).

Another important dimension of our problem is the use of a 3PL to handle the deliveries. According to Jayaram and Tan (2010) and to Giri and Sarker (2017), there is a significant advantage to be gained by using a 3PL to coordinate a supply chain. Qin et al. (2014) describe the interesting case of a 3PL that itself outsources its deliveries to express delivery companies such as FedEx and UPS. Lee et al. (2003) present a mathematical model for an integrated inventory and shipment scheduling policy in the context of a dynamic demand problem motivated by a 3PL application. Aguezzoul (2014) presents a review of the 3PL logistics selection problem. Our problem extends the literature on the use of 3PL in that, in addition to the inventory consolidation policy and the dispatch schedule, we also seek to determine at which level to enter the 3PL's supply chain.

As far as we are aware, the problem considered in this paper has never been discussed in the relevant scientific literature, which makes our contribution original.

1.3. Organization of the paper

The remainder of this paper is organized as follows. We first provide a mathematical model for the static case in Section 2. Section 3 describes four policies for the dynamic case. In Section 4 we perform extensive computational experiments on data from California and Texas in order to assess these policies. Conclusions follow in Section 5.

2. Mathematical model for the static ESP

In order to formally model the static ESP we need to introduce two graphs. The first one is the *supply chain graph*. The second graph, called the *decision graph*, represents the transportation and inventory holding decisions made by the retailer.

2.1. Description of the supply chain graph

The supply chain graph (Figure 1) is a directed graph consisting of a retailer and $\ell+1$ levels k indexed from 0 to ℓ. Each level $k = 1, \dots, \ell$ contains m_k SCs, denoted by SC_j^k, $j = 1, \dots, m_k$. There are n FDP_i at level 0, each of which can be reached from a single SC at any level, but an SC can reach several FDPs. We define $j(i,k)$ as the SC_j^k used to deliver orders to FDP_i. Let F_j^k be the set of FDPs that can be reached if the retailer enters the supply chain through SC_j^k. Our problem is defined over p time periods and, for simplicity of notation, we assume that moving goods between two successive levels takes one time period. The general case is easily modeled but involves a more complicated notation. Hence a demand having a delivery deadline of d periods cannot enter the supply chain at a level k higher than d (assuming that same-period delivery corresponds to a one-period deadline). No horizontal movements are allowed.

2.2. Description of the decision graph

The decision graph is also a directed graph denoted by $G = (W, A)$, where W is a node set partitioned into three sets U, V, and $\{\omega\}$, and A is an arc set partitioned into four sets A_x, A_y, A_w, and A_z.

The set U is made up of *source* nodes u_{it}^k $(k = 1, \dots, \ell, i = 1, \dots, n, t = 1, \dots, p-k+1)$. Each node $u_{it}^k \in U$ supplies the demand q_{it}^k ordered by FDP_i made known at period t and having a delivery deadline of k periods. Without loss of generality, we assume that q_{it}^k is integer. The set V is made up of *transfer* nodes v_{jt}^k $(k = 1, \dots, \ell, j = 1, \dots, m_k, t = 1, \dots, p-k+1)$. Each node $v_{jt}^k \in V$ receives the orders shipped to FDP_i at period t entering the supply chain at level k. The node ω is a *sink* at which all delivered demands arrive. Let

$$A_x = \left\{(u_{it}^k, v_{jt}^k) : k = 1, \dots, \ell, j = 1, \dots, m_k, i \in F_j^k, t = 1, \dots, p-k+1\right\},$$

be the set of *shipping arcs*. Using such an arc corresponds to shipping some demand of FDP_i at period t by entering the supply chain at level k. Let c_{it}^k $(k = 1, \dots, \ell; i = 1, \dots, n; t = 1, \dots, p-k+1)$ denote the *variable* cost of shipping one unit of demand to FDP_i at period t if the supply chain is entered at $\mathrm{SC}_{j(i,k)}^k$, which is the unique SC serving FDP_i at level k. c_{it}^k is defined as the sum of the sorting and outsourcing costs. The former cost is incurred directly by the retailer when sorting demand to be shipped to the supply chain at level k. The latter cost is the fare charged by the 3PL when entering their supply chain at level k. We assume the outsourcing cost to be proportional to the length of the unique path from $\mathrm{SC}_{j(i,k)}^k$ to FDP_i in the supply chain. There are no capacity limitations on these arcs. Let

$$A_y = \left\{(u_{it}^k, u_{it}^{k-1}) : k = 2, \dots, \ell, i = 1, \dots, n, t = 1, \dots, p-k+1\right\},$$

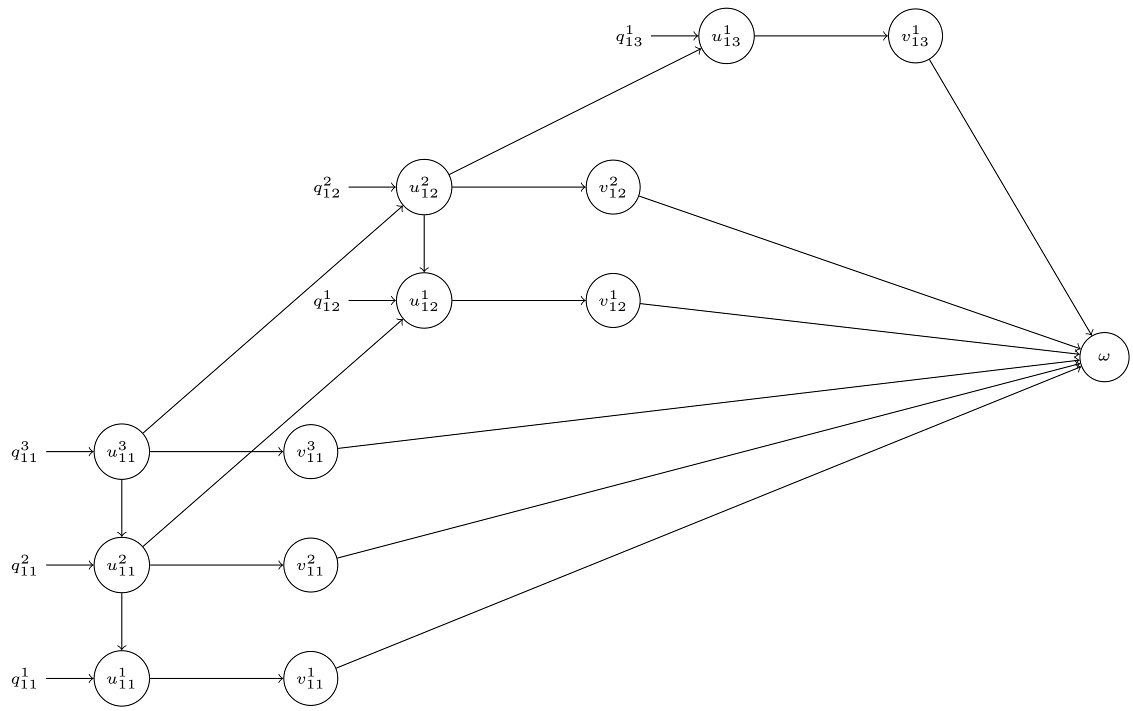


Figure 2: Decision graph $G = (W, A)$ for $n = 1, p = 3$ and $m_k = 1, F_1^k = \{1\}$ for $k = 1, 2, 3$.

be the set of *transfer arcs*, each of which corresponds to the fact that some demand will enter the supply chain at a level lower than k in either the current period t or a subsequent one. This implies that this demand will be consolidated with another demand having a shorter deadline. There are no costs or capacity limitations on these arcs. Let

$$A_w \left\{ (u_{it}^k, u_{it+1}^{k-1}) : k = 2, \ldots, \ell, i = 1, \ldots, n, t = 1, \ldots, p-k+1 \right\},$$

be the set of *waiting arcs*. Using such an arc correspond to holding some of the demand of FDP$_i$ for at least one more period. Let h_i $(i = 1, \ldots, n)$ denote the cost of holding one unit of the demand of FDP$_i$ for one period. There are no capacity limitations on these arcs. Finally, let

$$A_z = \left\{ (v_{jt}^k, \omega) : k = 1, \ldots, \ell, j = 1, \ldots, m_k, t = 1, \ldots, p-k+1 \right\},$$

be the set of *vehicle arcs*. Let g_{jt}^k $(k = 1, \ldots, \ell; j = 1, \ldots, m_k; t = 1, \ldots, p-k+1)$ be the *fixed* cost of using a vehicle from the retailer to an SC$_j^k$ at period t. This cost includes the delivery cost from SC$_j^k$ to the FDPs in the set F_j^k which can be approximated, as we will explain in Section 4.2.3. Let Q_j^k $(k = 1, \ldots, \ell, j = 1, \ldots, m_k)$ denote the capacity of the vehicle sent to sorting center j at level k. Figure 2 depicts a simple decision graph.

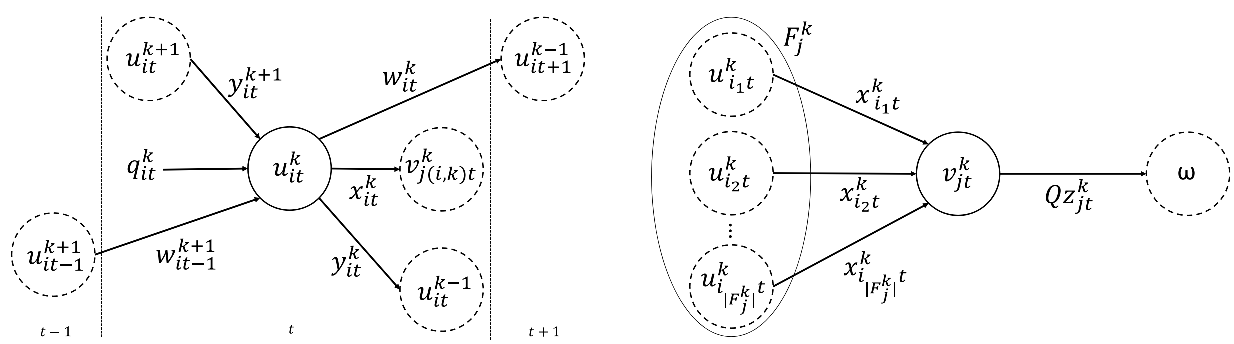
2.3. Mathematical model

We will use the following variables in our mathematical model. We recall that ℓ is the number of levels in the supply chain excluding the FDPs, p is the number of periods, and m_k is the number of SCs at level k.

Decision variables:

- x_{it}^k ($(i,t,k) \in A_x$): amount of demand shipped to FDP_i at period t entering the supply chain at level k;
- y_{it}^k ($(i,t,k) \in A_y$): amount of demand of FDP_i with delivery deadline k consolidated at period t with demand having delivery deadline $k-1$;
- w_{it}^k ($(i,t,k) \in A_w$): amount of demand of FDP_i in waiting, i.e., moved from level k at period t to level $k-1$ at period $t+1$;
- z_{jt}^k ($(j,t,k) \in A_z$): number of vehicles sent to sorting center j at period t at level k.

We note that the variables x_{it}^k associated with arcs in A_x model the flow routed from the retailer to different sorting centers. In contrast, the variables y_{it}^k and w_{it}^k associated with arcs in A_y and A_w, respectively, model the flow moved by the retailer from one level to another. This flow will be consolidated with other demand having a shorter deadline to enter the supply chain, using some arcs in A_x, in the current period or a subsequent one. Figures 3a and 3b depict decision subgraphs incident to nodes u_{it}^k and v_{jt}^k, respectively, and their associated decision variables. As can be seen from Figure 3a, node u_{it}^k consolidates the demand of FDP_i ordered at period t and having a delivery deadline of k periods with the demand from the previous period $t-1$ having delivery deadline $k+1$, and the demand from the same period t having delivery deadline $k+1$. This demand will be either shipped to enter the supply chain using sorting center $\text{SC}_{j(i,k)t}^k$, or moved to a lower level and be eventually consolidated with demand having a shorter delivery deadline in the current period t or in the subsequent one $t+1$.



(a) Decision subgraph incident to node u_{it}^k

(b) Decision subgraph incident to node v_{jt}^k

Figure 3: Nodes of decision graph $G = (W, A)$.

Using these variables, the model for the static problem (SESP) is as follows:

$$(SESP) \text{ minimize } \sum_{(i,t,k)\in A_x} c_{it}^k x_{it}^k + \sum_{(i,t,k)\in A_w} h_i w_{it}^k + \sum_{(j,t,k)\in A_z} g_{jt}^k z_{jt}^k \tag{1}$$

$$\text{subject to } \quad y_{it}^k + w_{it}^k + x_{it}^k = q_{it}^d + y_{it}^{k+1} + w_{i,t-1}^{k+1} \qquad u_{it}^k \in U \tag{2}$$

$$\sum_{i\in F_j^k} x_{it}^k \le Q_j^k z_{jt}^k \qquad v_{jt}^k \in V \tag{3}$$

$$x_{it}^k \in \mathbb{Z}^+ \qquad (i,t,k) \in A_x \tag{4}$$

$$y_{it}^k \in \mathbb{Z}^+ \qquad (i,t,k) \in A_y \tag{5}$$

$$w_{it}^k \in \mathbb{Z}^+ \qquad (i,t,k) \in A_w \tag{6}$$

$$z_{jt}^k \in \mathbb{Z}^+ \qquad (j,t,k) \in A_z. \tag{7}$$

The objective function (1) contains two terms corresponding to the shipping and holding costs, and to the vehicle cost, respectively. Constraints (2) ensure flow conservation in the decision graph. We note that these flow conservation constraints are stated only in their most general form, i.e., assuming that the arcs associated with all variables in (2) exist in A (see Figure 3a). Depending on the particular values of (i,t,k), some of the variables may not exist. For example, when $t = 1$ and $k = 1$ the associated constraints reduce to $x_{i1}^1 = q_{i1}^1 + y_{i1}^2$, for $i = 1,\dots,n$. For ease of exposition, we avoid writing each combination of (i,t,k) considering a different susbset of variables appearing in (2). Constraints (3) guarantee that the vehicle capacity is always sufficient to carry the demand entering the supply chain at any level at any time period. Constraints (4)–(7) define the domains of the variables. The SESP is a multi-facility extension of the well-known single-commodity, uncapacitated, fixed-charge network flow problem (Ortega and Wolsey, 2003), which is known to be NP-Hard.

3. Policies for the dynamic ESP

As observed by Berbeglia et al. (2010), solving a dynamic problem does not consist in optimizing a static input, but rather in developing a policy that specifies actions to be taken in the presence of information revealed over time. This is the approach taken in this study.

When the retailer receives an order it can either decide to ship it immediately, in which case the level at which the supply chain is entered must also be determined, or to postpone the decision regarding this order until the next period. Of course, the delivery deadline associated with the order has to be taken into account when making this decision. We have developed four policies for the dynamic ESP: a threshold policy, an optimistic policy, a pessimistic policy, and a minimax regret policy. The first of these policies applies a simple decision rule based on a vehicle filling threshold θ_k for each level $k \ge 1$, and uses a ranking mechanism to prioritize some orders. In contrast, the last three policies optimize the cost by solving a mathematical model.

3.1. Threshold policy

At each period $t = 1,\dots,p$, the threshold policy applies a greedy criterion to fill and send vehicles to each potential SC_j^k, for $k = 1,\dots,\ell$, $j = 1,\dots,m_k$, starting from the lowest to the highest level of the supply chain. Recall that any order with a deadline of d periods has to enter the supply chain at a level k not exceeding d.

Algorithm 1 Threshold policy.

for $k = 1, \dots, \ell$ **do**
  for $j = 1, \dots, m_k$ **do**
    if $k = 1$ **then**
      $z_{jt}^1 \leftarrow \left\lfloor \left(\sum_{i \in F_j^1} q_{it}^1 + \sum_{d=2}^{\ell} \sum_{i \in O_{jd}^1} q_{it}^d \right) / Q_j^1 \right\rfloor$
      if $\left(\sum_{i \in F_j^1} q_{it}^1 + \sum_{d=2}^{\ell} \sum_{i \in O_{jd}^1} q_{it}^d \right) - Q_j^1 z_{jt}^1 \geq \theta_1$ **or** $\sum_{i \in F_j^1} q_{it}^1 - Q_j^1 z_{jt}^1 > 0$ **then**
        $z_{jt}^1 \leftarrow z_{jt}^1 + 1$
      end if
      Fill z_{jt}^1 vehicles with orders q_{it}^1 and send them to SC_j^1
      for $d = 2, \dots, \ell$ **do**
        Fill z_{jt}^1 vehicles with orders based on Δ_{id}^1 values and send them to SC_j^1
      end for
      Update satisfied orders
    else
      $z_{jt}^k \leftarrow \left\lfloor \sum_{d=k}^{\ell} \sum_{i \in O_{jd}^k} q_{it}^d / Q_j^k \right\rfloor$
      if $\sum_{d=k}^{\ell} \sum_{i \in O_{jd}^k} q_{it}^d - Q_j^k z_{jt}^k \geq \theta_k$ **then**
        $z_{jt}^k \leftarrow z_{jt}^k + 1$
      end if
      for $d = k, \dots, \ell$ **do**
        Fill z_{jt}^k vehicles with orders based on Δ_{id}^k values and send them to SC_j^k
      end for
      Update satisfied orders
    end if
  end for
end for

This policy is based on two features. The first is that for each SC_j^k, it identifies a promising set of orders O_{jd}^k, $d = k, \dots, \ell$, to be sent in period t using as many fully loaded vehicles as possible. In particular, we define

$$O_{jd}^k = \left\{ i \in F_j^k : k \in \arg \min_{k'=1,\dots,d} \left\{ g_{jt}^{k'} / Q_j^{k'} + c_{it}^{k'} \right\} \right\}. \tag{8}$$

Assuming that the contribution of each order to the cost of using a vehicle is proportional to the amount of capacity consumed by q_{it}^d, we note that O_{jd}^k contains the set of orders q_{it}^d with delivery deadline d having k as their best level to enter the supply chain. If one of the vehicles open at level k is not full, it can also be dispatched immediately with its partial load or wait until the next period in the expectation that new orders will arrive to fill it more. This decision is made by using a filling threshold θ_k, $k = 1, \dots, \ell$ which must be attained in order to dispatch the vehicle. The values of θ_k are user defined and can be determined empirically. This process is applied iteratively until level ℓ is reached.

The second feature is that the policy uses a filling strategy of vehicles in which the elements of O_{jd}^k are ranked in non-increasing order of the weighted cost difference

$$\Delta_{id}^k = \min_{k'=1,\dots,d-1 : k' \neq k} \left\{ g_{jt}^{k'} / Q_j^{k'} + c_{it}^{k'} \right\} - g_{jt}^k / Q_j^k + c_{it}^k, \tag{9}$$

$k = 2, \dots, \ell, d = k, \dots, \ell, i \in O^k_{jd}$. This ranking system fills vehicles using orders that have an expected higher increase in cost going from period t to $t+1$. The process just described is summarized in Algorithm 1.

3.2. Optimistic policy

The optimistic policy works on the expectation that the delivery cost of the orders that do not completely fill a vehicle at the current period will be lower at a subsequent period, considering the holding, shipping and vehicle costs. This policy is optimistic in that it assumes that these delivery costs will be minimized by using only full vehicle loads at subsequent periods. The optimized cost is estimated by solving an optimistic problem OP_t at period t. The model is then:

$$(OP_t) \text{ minimize} \sum_{(i,k)\in A^t_x} c^k_{it} x^k_{it} + \sum_{(i,k)\in A^t_w} \left(h_i + c^{k-1}_{i,t+1} + \frac{g^{k-1}_{j(i,k),t+1}}{Q^{k-1}_{j(i,k)}} \right) w^k_{it} + \sum_{(j,k)\in A^t_z} g^k_{jt} z^k_{jt} \tag{10}$$

$$\text{subject to } y^k_{it} + w^k_{it} + x^k_{it} = q^d_{it} + y^{k+1}_{it} + w^{k+1}_{i,t-1} \qquad u^k_{it} \in U^t \tag{11}$$

$$\sum_{i\in F^k_j} x^k_{it} \le Q^k_j z^k_{jt} \qquad v^k_{jt} \in V^t \tag{12}$$

$$x^k_{it} \in \mathbb{Z}^+ \qquad (i,t,k) \in A^t_x \tag{13}$$

$$y^k_{it} \in \mathbb{Z}^+ \qquad (i,t,k) \in A^t_y \tag{14}$$

$$w^k_{it} \in \mathbb{Z}^+ \qquad (i,t,k) \in A^t_w \tag{15}$$

$$z^k_{jt} \in \mathbb{Z}^+ \qquad (j,t,k) \in A^t_z, \tag{16}$$

where U^t, V^t, A^t_x, A^t_y, A^t_w, and A^t_z denote the nodes and edges of $G^t = (W^t, A^t)$ which is a subgraph of $G(W, A)$ that contains all the nodes and edges associated with periods $t-1$, t, and $t+1$.

The objective function (10) contains three terms. The first one consists of the shipping cost of the demand satisfied at period t. The second represents the holding, shipping and estimated unit vehicle cost of the demand held until the next period. The third term corresponds to the vehicle cost incurred at period t. The remainder of the model is similar to SESP.

3.3. Pessimistic policy

The pessimistic policy works on the expectation that the delivery cost of the orders that do not completely fill a vehicle at the current period will be higher at a subsequent period and hence, these orders should be delivered immediately. This policy is pessimistic in that it assumes that no new demand will arrive in subsequent periods to share the vehicle cost.

To model the problem we define new nodes and arcs:

- $\bar{V}^t$ is made up of *transfer* nodes $\bar{v}^k_{jt}$ $(k = 2, \dots, \ell, j = 1, \dots, m_k : k \le p - t + 1)$.
- $\bar{A}^t_w = \left\{ (u^k_{it}, \bar{v}^k_{jt}) : k = 2, \dots, \ell, j = 1, \dots, m_k, i \in F^{k-1}_j, k \le p - t + 1 \right\}$,
- $\bar{A}^t_z = \left\{ (\bar{v}^k_{jt}, \omega_t) : k = 2, \dots, \ell, j = 1, \dots, m_k, k \le p - t + 1 \right\}$

The optimal delivery plan can be determined by solving a pessimistic problem at period t (PP_t):

$$(PP_t) \text{ minimize } \sum_{(i,k)\in A_x^t} c_{it}^k x_{it}^k + \sum_{(i,k)\in \bar{A}_w^t} (h_i + c_{i,t+1}^{k-1}) w_{it}^k + \sum_{(j,k)\in A_z^t} g_{jt}^k z_{jt}^k + \sum_{(j,k)\in \bar{A}_z^t} g_{j,t+1}^{k-1} \bar{z}_{jt}^k \tag{17}$$

$$\text{subject to } (11) - (16)$$

$$\sum_{i\in F_j^{k-1}} w_{it}^k \leq Q_j^{k-1} \bar{z}_{jt}^k \qquad \bar{v}_{jt}^k \in \bar{V}^t \tag{18}$$

$$\bar{z}_{jt}^k \in \mathbb{Z}^+ \qquad (j,t,k) \in \bar{A}_z^t. \tag{19}$$

The objective function of this model is similar to that of OP_t except that the vehicle cost of period $t+1$ is included in the last term by means of new variables $\bar{z}_{jt}^k (k = 2, \dots, \ell, j = 1, \dots, m_k)$. These are the number of vehicles that would be send to sorting center j at level $k-1$ at the next period because of the holding demand of current period. Constraints (18) ensure that the vehicle capacity will be sufficient to carry the waiting demand entering the supply chain at level $k-1$ during the next period.

3.4. Minimax regret policy

At period t this policy minimizes a regret R_t computed by comparing the value of a solution to the optimal values $v(OP_t)$ and $v(PP_t)$ of the models OP_t and PP_t, respectively. This is achieved by solving a minimax regret problem at period t (RP_t):

$$(RP_t) \quad \text{minimize} \quad R_t$$

$$\text{subject to } R_t \geq \sum_{(i,k)\in A_x^t} c_{it}^k x_{it}^k + \sum_{(i,k)\in \bar{A}_w^t} \left(h_i + c_{i,t+1}^{k-1} + \frac{g_{j(i,k),t+1}^{k-1}}{Q_{j(i,k)}^{k-1}} \right) w_{it}^k + \sum_{(j,k)\in A_z^t} g_{jt}^k z_{jt}^k - v(OP_t) \tag{20}$$

$$R_t \geq \sum_{(i,k)\in A_x^t} c_{it}^k x_{it}^k + \sum_{(i,k)\in \bar{A}_w^t} (h_i + c_{i,t+1}^{k-1}) w_{it}^k + \sum_{(j,k)\in A_z^t} g_{jt}^k z_{jt}^k + \sum_{(j,k)\in \bar{A}_z^t} g_{j,t+1}^{k-1} \bar{z}_{jt}^k - v(PP_t) \tag{21}$$

$$(11) - (16), (18), (19).$$

Constraint (20) states that the regret is at least equal to the difference between the optimistic problem's objective function (evaluated on the minimax regret solution) and $v(OP_t)$. Similarly, constraint (21) states that the regret is at least equal to the difference between the pessimistic problem's objective function (evaluated on the minimax regret solution) and $v(PP_t)$.

4. Computational experiments

We have conducted an extensive computational study in order to assess the empirical performance of the four policies described in Section 3. We next provide the details of the experimental design used to carry out our analyses.

4.1. Experimental design and evaluation

All algorithms were implemented in C and the ILPs were solved using the callable library of CPLEX 12.8.0 and run on an Intel Xeon CPU E5-2687W v3 processor at 3.10 GHz and 750 GB of RAM under a Linux environment. We used the default optimality gap of 0.01% in CPLEX for all of our experiments with a time limit of 3,600 seconds. In Section 4.2, we describe in detail the set of benchmark instances we have generated using real location data from the states of California and Texas. We generated a wide range of instances (over 1,300) to represent different scenarios of network topologies, cost structure, length of the planning horizon, and demand distributions.

In Section 4.3, we first provide the results of computational analyses of the threshold policy on the California and Texas networks. The goal of these experiments is to identify the best values of the vehicle filling threshold θ_k for each level $k \geq 1$. For these experiments, we used a training set containing a total of 648 instances. We then compare the four policies for the dynamic ESP using a larger test set having a total of 1,296 instances. Finally, we provide the results of the experiments we have carried out to assess the sensitivity of optimal solutions to changes on each of the cost and demand parameters considered in the ESP.

4.2. Instance generation

The instances we have generated differ according to the network topology, the length of the planning horizon, the cost structure, and the demand distribution.

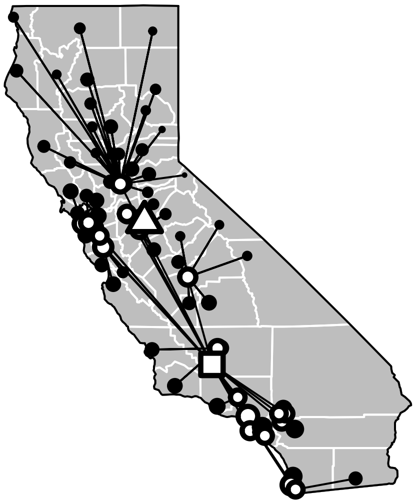
4.2.1. Network topologies

All networks used for our experiments contain three levels ($\ell = 3$) in addition to the retailer and the FDPs, as in Figure 1. We used real location data from the states of California and Texas to generate the networks. The population of California is largely concentrated in the south and in several large cities. This is therefore a good example to assess how our algorithms behave under this type of asymmetry. Texas is the state with the largest number of counties (used to generate the FDP locations) in the United States and is therefore interesting to assess the behavior of our algorithms on large-scale instances.

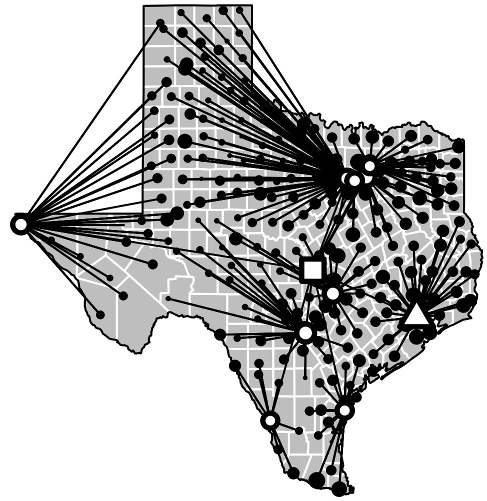
There are 58 counties in California and 254 in Texas. We located an FDP at each county seat, i.e., $n \in \{58, 254\}$. Since $m_1 = m_2$ and $m_3 = 1$, from now on we use $m = m_1 = m_2$. The number of SCs at levels $k = 1$ and $k = 2$ were set at $m \in \{5, 10, 20\}$. We located these at the centroid of the most populated cities of the state (case A) or at the most populated counties of the state (case B). The SC at level $k = 3$ was located at the center of gravity of the SCs at levels 1 and 2 weighted by their population (case C) or not (case D). Hence, there are four network topologies: AC, AD, BC, and BD. The retailer was either located at the center of gravity of the FDPs weighted by their population (case E) or not (case F), or at the center of gravity in the most populated city of the state (case G). To complete the network, each FDP was assigned to its closest SC at levels 1 or 2. Hence, $F_j^1 = F_j^2$ and $F_j^3 = \{1, \ldots, n\}$, $j = 1, \ldots, m$. In the following, $j(i) = j(i,1) = j(i,2)$. Figure 4 depicts six networks for California and Texas, and several topologies.

4.2.2. Length of the planning horizon

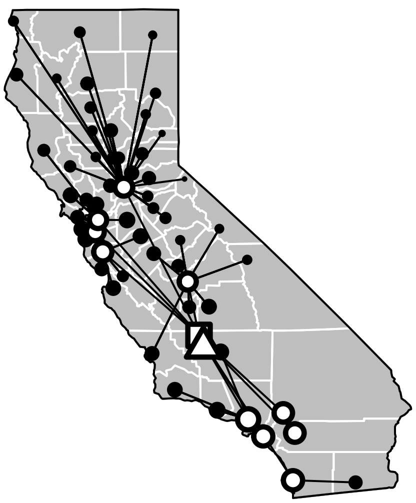
The length p of the planning horizon is set at six, 30, 60, 90, 180, and 365 in our experiments.



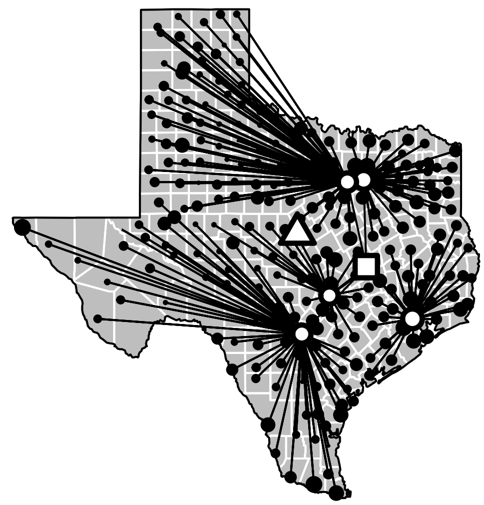
(a) California, cases AC, F, and $m = 20$



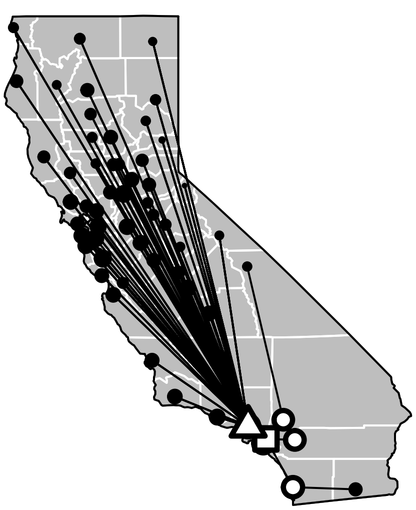
(b) Texas, cases AD, G, and $m = 10$



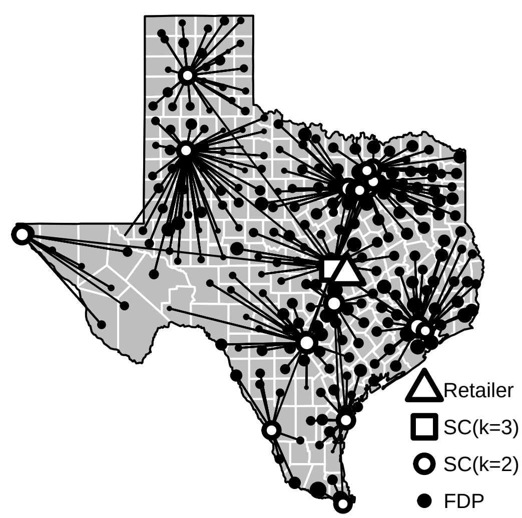
(c) California, cases BD, E, and $m = 10$



(d) Texas, cases BC, F, and $m = 5$



(e) California, cases BC, G, and $m = 5$



(f) Texas, cases AC, E, and $m = 20$

Figure 4: Six network instances drawn with the R library ggplot2 (Kahle and Wickham, 2013).

4.2.3. Cost structures

For any $t = 1, \ldots, p$, the shipping costs are calculated as follows: c_{it}^1, $(i = 1, \ldots, n)$ is a fixed fare including sorting and outsourcing costs; c_{it}^2 is calculated by means of a linear function with fixed sorting cost and outsourcing cost directly proportional to the length $d(i, j(i))$ of the unique path in the supply chain from the corresponding $\text{SC}_{j(i,2)}^2$ to FDP_i; to set c_{it}^3 the distances between post office i and the corresponding city sorting center and from that to the main sorting center are aggregated and taken into account. We set the holding costs to $h_i = 0$ for each FDP_i.

For each SC_j^k, $k = 1, 2, 3$, $j = 1, \ldots, m_k$, the transportation cost g_{jt}^k is proportional to the distance covered by the vehicle. To compute g_j^1, we assume that a vehicle starting at SC_j^1 will deliver to a set F_j^1 of FDPs by means of a tour. As in Daganzo (1984), (see also Franceschetti et al., 2017) the length of this tour is approximated by $0.57\sqrt{n_j B_j}$, where $n_j = |F_j^1|$ and B_j is the area of the convex hull of the points of F_j^1. The total distance traveled is approximately $2d(0, j) + 0.57\sqrt{n_j B_j}$ to account for twice the shortest path length between the retailer 0 and SC_j^1, plus the approximate tour length. For the computation of g_{jt}^2 and g_{jt}^3, we only consider twice the shortest path length.

In order to generate a diverse set of instances in which each type of cost dominates more or less the total cost, we considered three different structures: H, I, and J. These cost structures were designed with the objective of having target percentages of sorting costs, outsourcing costs, and transportation costs incurred directly by the retailer in an optimal solution. H is a structure in which 60% of the total cost corresponds to the transportation cost and the remaining 40% is equality split between the sorting and outsourcing costs (i.e., 20-20-60 %). I is a structure in which 50% of the total cost is sorting cost and the remaining 50% is equality split between the outsourcing and transportation costs (i.e., 50-25-25 %). Finally, J is a structure where 60% of the total cost corresponds to the outsourcing cost and the remaining 40% is equality split between the sorting and transportation costs (i.e., 20-60-20 %). For each one of these cost structures, we generated sorting costs that are larger when entering the supply chain at a lower level and smaller when entering it at a higher level. As in Ortiz-Astorquiza et al. (2019), we generated the outsourcing costs by considering economies of scale in transportation costs perceived by the 3PL when using sorting centers at higher levels of the supply chain. Here follow the details of the three cost structures:

Cost structure H:

$$
\begin{aligned}
c_{it}^1 &= 30 + 15; \\
c_{it}^2 &= 10 + 0.3d(i, j(i)); \\
c_{it}^3 &= 2.5 + 0.1(d(i, j(i)) + d(j(i), \text{SC}_1^3)); \\
g_{jt}^1 &= 5.3\left(2d(0, j) + 0.57\sqrt{n_j B_j}\right); \\
g_{jt}^2 &= 5.3(2d(0, j)); \\
g_{1t}^3 &= 5.3(2d(0, \text{SC}_1^3)).
\end{aligned}
$$

Cost structure I:

$$
\begin{aligned}
c_{it}^1 &= 30 + 15;\\
c_{it}^2 &= 20 + 2.0d(i, j(i));\\
c_{it}^3 &= 5.0 + 1.0(d(i, j(i)) + d(j(i), \mathrm{SC}_1^3));\\
g_{jt}^1 &= 1.0\left(2d(0, j) + 0.57\sqrt{n_j B_j}\right);\\
g_{jt}^2 &= 1.0(2d(0, j));\\
g_{1t}^3 &= 1.0(2d(0, \mathrm{SC}_1^3)).
\end{aligned}
$$

Cost structure J:

$$
\begin{aligned}
c_{it}^1 &= 30 + 120;\\
c_{it}^2 &= 10 + 15.0d(i, j(i));\\
c_{it}^3 &= 2.5 + 8.0(d(i, j(i)) + d(j(i), \mathrm{SC}_1^3));\\
g_{jt}^1 &= 2.0\left(2d(0, j) + 0.57\sqrt{n_j B_j}\right);\\
g_{jt}^2 &= 2.0(2d(0, j));\\
g_{1t}^3 &= 2.0(2d(0, \mathrm{SC}_1^3)).
\end{aligned}
$$

4.2.4. Demand distributions

The vehicle capacity was set to $Q_j^k = 100$ for all j, k. The demand was generated according to three different distributions K, L, and M. Distribution K is Poisson with a parameter λ equal to $0.00003 \times P_i$, where P_i is the population fo the county corresponding to FDP_i. Distributions L and M use a definition of λ that depends on k. Here follows the description of the three demand distributions:

Demand distribution K:

$$
\begin{aligned}
q_{it}^1 &\sim \mathrm{Poisson}(\lambda = 0.000030 \times P_i);\\
q_{it}^2 &\sim \mathrm{Poisson}(\lambda = 0.000030 \times P_i);\\
q_{it}^3 &\sim \mathrm{Poisson}(\lambda = 0.000030 \times P_i).
\end{aligned}
$$

Demand distribution L:

$$
\begin{aligned}
q_{it}^1 &\sim \mathrm{Poisson}(\lambda = 0.000015 \times P_i);\\
q_{it}^2 &\sim \mathrm{Poisson}(\lambda = 0.000030 \times P_i);\\
q_{it}^3 &\sim \mathrm{Poisson}(\lambda = 0.000060 \times P_i).
\end{aligned}
$$

Demand distribution M:

$$
\begin{aligned}
q_{it}^1 &\sim \mathrm{Poisson}(\lambda = 0.000060 \times P_i);\\
q_{it}^2 &\sim \mathrm{Poisson}(\lambda = 0.000030 \times P_i);\\
q_{it}^3 &\sim \mathrm{Poisson}(\lambda = 0.000015 \times P_i).
\end{aligned}
$$

4.3. Computational analyses

In this section we provide two sets of analyses carried out for the dynamic ESP. We first present results for the threshold policy on the California and Texas networks. We then compare the four policies presented in Section 3.

4.3.1. Analysis of the threshold policy on the California and Texas networks

Tables 1 and 2 provide computational results for the California and Texas network for 12 network topologies, $p = 30$, three values of m, and three cost structures. Each line represents a mean value over three demand distributions K, L, and M. Hence, 324 instances were solved and reported in each table. The values under the heading θ_1^*, θ_2^* and θ_3^* correspond to the best values of θ_1, θ_2 and θ_3 over all possible 1,331 combinations for each value of $(\theta_1, \theta_2, \theta_3) \in \{0, 10, \dots, 100\}^3$. The columns Dev(%) represent the percentage deviation of the solution cost over the optimal cost that would have been obtained if the instance had been solved optimally in a static environment, i.e., with all information available at the beginning of the planning horizon.

The results show that when the dominant component of the cost is sorting or outsourcing related, then the threshold heuristic policy yields a good approximation of the static cost. In contrast, when the transportation cost become more important, then the approximation deteriorates. Overall, the results are slightly better for California than for Texas which exhibits a more balanced demand distribution. In most cases, $\theta_1^* \geq \theta_2^* \geq \theta_3^*$, especially for cost structures I and J. This means that vehicles should be more filled when entering at a lower level of the supply chain.

Out of the 1,331 tested combinations for each value of $(\theta_1, \theta_2, \theta_3)$, the combination $(\bar{\theta}_1, \bar{\theta}_2, \bar{\theta}_3) = (80, 50, 50)$ minimizes the average Dev (%) over all 648 instances considered in Tables 1 and 2. This combination will be used to further assess the performance of the threshold policy over a different set of instances.

4.3.2. Comparison of the four policies for the dynamic ESP

We compare in Tables 3 to 6 the computational results obtained under the four policies described in Section 3 for the dynamic ESP: threshold, optimistic, pessimistic, and minimax regret. For the threshold policy, we provide the results obtained using either the best combination $\theta^* = (\theta_1^*, \theta_2^*, \theta_3^*)$ for each instance, or the same combination $\bar{\theta} = (\bar{\theta}_1, \bar{\theta}_2, \bar{\theta}_3) = (80, 50, 50)$ for all instances.

Tables 3 to 5 correspond to the three cost structures H, I, and J. In all cases, the length of the planning horizon is $p = 30$. We consider 12 network topologies, three values of m, two states, three demand distribution, yielding 216 instances per table, each solved under four policies. The first three lines of these tables provide a breakdown of the total cost into sorting, outsourcing, and transportation. The next three lines are: the percentage of the demand transferred to a lower level during the same period; the percentage of the demand waiting an extra period to be moved to the next level; and the average filling level of the vehicles. The last two lines are the maximal and average values of the percentage deviation Dev (%) with respect to the optimal static solution value.

The line Dev (%) of Tables 3 to 5 (in bold) shows that the optimistic policy is always the best whereas the threshold policy using $\bar{\theta}$ is the worst. The optimistic and minimax regret policies tend to be always very good irrespective of the cost structure. The threshold policy has a slightly better performance when using the best θ^* values as compared to the average best $\bar{\theta}$. The distribution of the total cost is about the same for the optimistic, pessimistic and minimax regret policies under the three cost structures. It is particularly interesting to observe, for cost structure H, that this cost distribution is very similar to that of the static case for the last three policies. In contrast, the threshold policy yields a slightly more different distribution from the optimal one. The threshold policy is also the one having the smallest filling rates, which is due to the fact that it gives higher priority to early shipping as opposed to holding the demand for

Topologies			Cost structure H				Cost structure I				Cost structure J			
SCs	Retailer	m	θ_1^*	θ_2^*	θ_3^*	Dev (%)	θ_1^*	θ_2^*	θ_3^*	Dev (%)	θ_1^*	θ_2^*	θ_3^*	Dev (%)
AC	E	5	50	80	60	3.53	70	30	10	0.91	70	30	10	0.69
AC	E	10	80	60	20	6.98	70	40	10	1.36	50	30	10	0.84
AC	E	20	80	60	10	5.40	60	50	10	1.46	60	30	10	1.09
AC	F	5	40	80	50	2.92	60	30	10	1.14	50	20	10	0.91
AC	F	10	80	80	70	6.30	40	50	10	1.61	40	30	10	1.10
AC	F	20	80	80	80	6.39	40	30	10	2.08	40	30	10	1.44
AC	G	5	40	80	40	8.26	80	30	10	0.82	70	20	10	0.56
AC	G	10	90	70	50	8.80	60	50	10	1.25	80	40	10	0.89
AC	G	20	70	70	100	10.33	60	50	10	1.24	60	40	10	1.00
AD	E	5	40	70	70	4.26	60	30	10	0.89	60	20	10	0.66
AD	E	10	50	50	70	6.24	60	40	10	1.28	60	30	10	0.89
AD	E	20	40	60	10	5.36	60	40	70	1.49	60	30	10	1.06
AD	F	5	50	80	90	2.91	60	30	10	1.25	50	20	10	0.87
AD	F	10	40	80	70	7.49	40	40	10	1.63	50	20	10	1.11
AD	F	20	50	70	70	7.19	40	30	70	2.02	50	20	10	1.42
AD	G	5	60	80	70	4.70	70	40	10	0.87	80	30	10	0.56
AD	G	10	80	80	80	7.11	70	50	10	1.19	70	30	10	0.81
AD	G	20	90	70	100	11.13	60	50	10	1.22	60	40	10	1.00
BC	E	5	70	20	60	19.24	100	60	10	0.43	80	50	10	0.40
BC	E	10	50	60	80	5.10	80	40	10	1.08	70	40	10	0.99
BC	E	20	30	60	20	5.98	80	30	10	1.12	90	20	10	1.00
BC	F	5	80	80	80	16.15	100	70	30	0.54	70	60	10	0.47
BC	F	10	80	80	100	4.83	80	40	10	1.14	80	40	10	1.07
BC	F	20	100	70	100	5.81	100	30	10	1.25	80	20	10	1.16
BC	G	5	90	10	40	23.15	80	40	20	0.31	60	50	10	0.29
BC	G	10	50	50	30	8.86	70	50	10	1.12	70	40	10	1.09
BC	G	20	30	60	50	8.68	90	50	10	1.20	100	20	10	1.06
BD	E	5	80	40	80	22.43	70	70	10	0.42	80	50	10	0.39
BD	E	10	40	60	30	4.30	70	40	10	1.10	80	40	10	1.05
BD	E	20	30	60	50	6.89	100	30	10	1.12	100	10	10	1.07
BD	F	5	80	80	60	18.64	100	70	10	0.51	80	60	10	0.49
BD	F	10	50	70	70	4.89	80	50	10	1.09	70	40	10	1.08
BD	F	20	60	70	50	6.24	70	30	10	1.30	90	20	10	1.33
BD	G	5	80	10	40	28.29	70	50	10	0.34	70	50	10	0.31
BD	G	10	60	50	80	7.52	80	50	10	1.09	80	40	10	1.13
BD	G	20	20	70	60	8.11	90	40	10	1.23	80	20	10	1.05
						8.90				**1.11**				**0.90**

Table 1: Best θ_1, θ_2 and θ_3 and corresponding Dev (%) value for the threshold policy for 12 network topologies, three values of m, three cost structures, and $p = 30$ for the state of California.

at least one more period.

Table 6 provides an overall comparison of the four policies under a wide variety of settings for cost structure H: 12 network topologies, three values of m, two states, three demand distributions, and six planning horizons for a total of 1,296 instances, each solved under four policies. We note that the threshold policy exhibits a very stable behavior with respect to an increase in

Topologies			Cost structure H				Cost structure I				Cost structure J			
SCs	Retailer	m	θ_1^*	θ_2^*	θ_3^*	Dev (%)	θ_1^*	θ_2^*	θ_3^*	Dev (%)	θ_1^*	θ_2^*	θ_3^*	Dev (%)
AC	E	5	90	40	10	10.10	90	50	10	1.54	90	40	10	0.64
AC	E	10	30	50	70	6.93	80	40	10	1.74	90	20	10	1.32
AC	E	20	20	60	30	7.17	80	70	0	1.42	60	30	10	1.06
AC	F	5	80	60	50	12.09	90	50	10	1.52	80	40	10	0.65
AC	F	10	30	60	100	7.99	90	40	10	1.69	80	30	10	1.34
AC	F	20	20	60	20	8.04	70	80	20	1.42	80	30	10	1.18
AC	G	5	70	50	70	13.59	100	50	10	1.67	90	50	10	0.63
AC	G	10	40	50	100	8.71	80	50	10	1.87	80	30	10	1.52
AC	G	20	20	70	100	8.15	80	80	30	1.37	70	50	10	1.14
AD	E	5	90	60	10	8.85	80	40	10	1.62	90	40	10	0.59
AD	E	10	30	60	90	6.91	80	50	10	1.76	80	30	10	1.28
AD	E	20	20	60	50	7.20	80	70	10	1.34	80	30	10	1.01
AD	F	5	100	60	80	10.83	90	40	10	1.66	90	40	10	0.73
AD	F	10	40	50	70	7.79	90	40	10	1.74	90	30	10	1.36
AD	F	20	20	60	30	8.25	80	70	10	1.39	70	30	10	1.25
AD	G	5	70	60	50	11.90	90	40	10	1.59	90	50	10	0.68
AD	G	10	30	50	100	9.46	90	50	10	1.91	80	20	10	1.50
AD	G	20	20	70	100	8.04	70	80	10	1.38	70	40	10	1.11
BC	E	5	70	40	20	7.28	80	40	10	0.98	90	20	10	0.84
BC	E	10	40	50	0	8.42	80	20	10	1.67	100	20	10	0.99
BC	E	20	40	70	20	7.57	80	30	10	1.99	90	10	10	1.31
BC	F	5	80	80	30	10.46	90	30	10	1.14	90	40	10	0.93
BC	F	10	30	60	40	10.37	80	30	10	1.73	90	20	10	1.06
BC	F	20	40	90	50	7.31	90	40	10	1.72	80	20	10	1.36
BC	G	5	70	60	20	11.42	90	40	10	1.08	90	30	10	0.80
BC	G	10	30	60	70	13.38	80	40	10	1.86	90	20	10	1.12
BC	G	20	100	80	80	11.46	90	30	10	2.41	100	10	10	1.59
BD	E	5	90	50	20	7.02	90	40	10	0.96	80	40	10	0.84
BD	E	10	30	50	60	8.71	90	30	10	1.56	90	20	10	0.98
BD	E	20	60	80	30	7.30	90	30	100	2.00	100	10	10	1.37
BD	F	5	90	70	0	7.80	80	40	10	1.02	80	40	10	0.87
BD	F	10	30	50	60	9.87	90	30	10	1.66	90	20	10	1.05
BD	F	20	60	80	60	7.22	90	40	90	1.79	90	20	10	1.42
BD	G	5	70	50	40	12.29	90	40	10	0.93	90	30	10	0.87
BD	G	10	30	60	100	10.81	90	20	10	2.00	90	20	10	1.17
BD	G	20	80	80	90	12.01	90	30	90	2.51	100	10	10	1.54
						9.24				**1.60**				**1.09**

Table 2: Best θ_1, θ_2 and θ_3 and corresponding Dev (%) value for the threshold policy for 12 network topologies, three values of m, three cost structures, and $p = 30$ for the state of Texas.

the length of the planning horizon. However, it can yield rather expensive solutions for some settings. The remaining three policies are near-optimal under all settings and improve as the planning horizon becomes longer.

We summarize in Table 7 the optimality gap and running time (in seconds) obtained for different lengths of the planning horizon under cost structure H for the 1,296 considered instances

	Static	Dynamic				
		Threshold (θ^*)	Threshold ($\bar{\theta}$)	Optimistic	Pessimistic	Minimax regret
Sorting (%)	22.32	19.61	19.14	22.08	21.68	21.81
Outsourcing (%)	19.27	16.31	16.17	19.12	19.06	18.99
Transportation (%)	58.41	64.08	64.69	58.80	59.26	59.20
Transferred (%)	24.21	23.43	25.53	36.69	36.75	39.75
Waiting (%)	22.23	16.14	10.88	9.05	6.22	4.35
Truckload (%)	90.72	83.18	81.83	89.68	88.39	88.82
Max Dev(%)	0.00	48.47	49.79	2.16	3.02	2.64
Dev(%)	**0.00**	**9.21**	**9.64**	**0.61**	**1.07**	**0.87**

Table 3: Comparison of the four policies for the dynamic ESP under cost structure H, and $p = 30$.

	Static	Dynamic				
		Threshold (θ^*)	Threshold ($\bar{\theta}$)	Optimistic	Pessimistic	Minimax regret
Sorting (%)	53.01	52.24	52.29	52.93	52.17	52.73
Outsourcing (%)	21.87	21.67	21.90	21.80	22.18	21.99
Transportation (%)	25.12	26.09	25.81	25.27	25.65	25.28
Transferred (%)	48.58	78.31	76.42	65.89	62.80	66.81
Waiting (%)	46.28	17.72	21.32	29.09	31.49	27.63
Truckload (%)	87.96	84.33	85.81	87.38	83.97	86.73
Max Dev(%)	0.00	2.84	4.37	0.90	3.76	1.29
Dev(%)	**0.00**	**1.37**	**1.58**	**0.17**	**1.54**	**0.44**

Table 4: Comparison of the four policies for the dynamic ESP under cost structure I, and $p = 30$.

	Static	Dynamic				
		Threshold (θ^*)	Threshold ($\bar{\theta}$)	Optimistic	Pessimistic	Minimax regret
Sorting (%)	18.13	17.91	18.01	18.11	17.92	18.08
Outsourcing (%)	61.25	60.64	61.45	61.14	60.66	61.05
Transportation (%)	20.62	21.45	20.53	20.75	21.42	20.87
Transferred (%)	48.35	82.07	78.28	66.23	61.94	69.36
Waiting (%)	46.69	14.10	21.40	28.76	33.14	25.55
Truckload (%)	86.31	82.13	85.97	85.91	82.16	85.35
Max Dev (%)	0.00	2.30	8.60	0.54	2.93	0.85
Dev (%)	**0.00**	**1.01**	**2.01**	**0.11**	**1.22**	**0.24**

Table 5: Comparison of the four policies for the dynamic ESP under cost structure J, and $p = 30$.

of Table 6. We recall that the time limit was set to one hour.

In the case of the dynamic ESP, the MILPs of all four policies can be optimally solved (using such precision) in a few seconds and as a consequence, we do not report the optimality gap. In

	Threshold (θ^*)	Threshold ($\bar{\theta}$)	Optimistic	Pessimistic	Minimax regret
California	9.05	9.41	0.52	1.06	0.83
Texas	9.31	9.84	0.70	1.12	0.95
SC topology AC	7.97	8.31	0.56	1.11	0.88
SC topology AD	7.68	7.99	0.57	1.10	0.89
SC topology BC	10.41	10.95	0.68	1.12	0.91
SC topology BD	10.66	11.23	0.65	1.03	0.86
Retailer E	7.98	8.25	0.67	1.24	1.01
Retailer F	8.38	8.80	0.63	1.21	0.92
Retailer G	11.18	11.82	0.54	0.82	0.72
Demand K	8.68	9.06	0.68	1.14	0.94
Demand L	12.70	13.80	0.48	0.90	0.71
Demand M	6.17	6.00	0.68	1.22	1.01
$m = 5$	11.72	11.90	0.44	0.64	0.57
$m = 10$	7.95	8.08	0.85	1.42	1.22
$m = 20$	7.87	8.88	0.54	1.20	0.87
$p = 6$	8.56	8.82	1.00	1.54	1.31
$p = 30$	9.21	9.64	0.61	1.07	0.87
$p = 60$	9.30	9.77	0.55	1.02	0.82
$p = 90$	9.34	9.82	0.53	0.99	0.80
$p = 180$	9.34	9.83	0.51	0.97	0.78
$p = 365$	9.33	9.83	0.48	0.94	0.74
Average	**9.18**	**9.62**	**0.61**	**1.09**	**0.89**

Table 6: Average value of Dev (%) under cost structure H.

	Static			Threshold	Optimistic	Pessimistic	Minimax regret
	Max gap(%)	Avg gap(%)	Time	time	time	time	time
$p = 6$	0.01	0.01	8.83	0.00	0.10	0.12	0.48
$p = 30$	0.04	0.01	469.70	0.00	0.47	0.63	4.20
$p = 60$	0.16	0.01	1068.52	0.01	0.94	1.28	7.18
$p = 90$	0.22	0.02	1680.04	0.02	1.40	1.95	25.01
$p = 180$	0.30	0.03	2205.50	0.04	2.77	3.91	45.57
$p = 365$	1.04	0.06	2735.49	0.08	5.54	7.78	165.38

Table 7: Time in seconds comparing different policies by p (Cost structure H).

the static ESP, the average optimallity gap is only 0.023% for all 1,296 considered instances. The maximum % gap obtained in all these instances is only 1.04%.

4.3.3. Sensitivity analyses

In this section we provide the results of sensitivity analyses performed to evaluate the changes on the optimal solution when changing independently each of the costs and demand parameters of the ESP. Tables 8 to 11 provide the obtained results using the state of California, for cases BD, E, under cost structure H, demand K, $m = 10$, and $p = 30$. In Tables 8 to 11, we report the changes in the optimal solution of the static ESP when considering an increase or a decrease in demand, transportation cost, sorting cost, and outsourcing cost, respectively.

Demand variation (%)	Total cost (%)	Sorting(%)	Outsourcing (%)	Transportation (%)	Transferred (%)	Waiting (%)	Truckload (%)
−50	−43.67	28.37	23.02	48.61	19.64	11.29	89.11
−45	−40.53	28.40	23.51	48.09	11.83	16.15	90.38
−40	−36.31	28.50	22.72	48.78	13.40	13.61	91.67
−35	−32.63	28.30	21.90	49.80	11.72	13.51	91.81
−30	−28.63	28.27	21.03	50.69	11.70	12.44	93.06
−25	−24.69	28.19	19.78	52.03	11.12	12.85	93.60
−20	−20.99	27.83	19.66	52.50	11.41	10.07	93.65
−15	−17.40	27.85	20.23	51.93	9.85	9.93	94.41
−10	−13.53	27.78	21.61	50.60	8.36	8.81	95.39
−5	−9.53	27.71	22.96	49.33	7.74	7.32	96.06
0	0.00	27.31	23.65	49.04	7.11	7.49	96.37
5	3.32	27.32	24.12	48.56	7.36	6.74	96.28
10	9.40	27.49	24.03	48.48	7.85	7.65	95.84
15	14.99	27.72	24.78	47.50	7.49	8.34	96.32
20	20.04	27.88	25.37	46.74	7.60	8.08	97.08
25	24.92	27.86	25.47	46.66	7.79	7.54	96.97
30	29.03	27.81	25.47	46.72	6.77	8.03	97.07
35	33.72	27.72	25.65	46.63	6.15	7.93	96.98
40	37.95	27.63	25.66	46.72	6.68	6.46	96.86
45	41.86	27.53	25.61	46.86	5.72	6.72	96.60
50	47.02	27.23	26.13	46.64	4.87	5.72	96.62

Table 8: Sensitivity analysis: varying the demand q_{it}^{d}.

From Table 8, we note that when demand increases the percentage of the sorting, outsourcing and transportation costs remain more or less the same as well as the % truckload utilization. When demand decreases, the sorting costs slightly increase and the % truckload utilization decrease. From Table 9, we observe that when increasing the transportation cost, the % truckload utilization increases as expected. From Table 10, we note that when the sorting costs increase, the % truckload utilization slightly decreases. From Table 11, it can be seen that when the outsourcing cost increases, the % truckload utilization does not seem to have an impact.

5. Conclusions

We have introduced, modeled and solved a shipping problem faced by an e-retailer who uses a 3PL to ship orders to final delivery points. What distinguishes this work from the existing literature is that the retailer can enter the supply chain of the 3PL at any level. The higher the level at which the supply chain is entered, the lower the sorting costs and the higher are the delivery costs for the retailer. The problem can be cast in a static mode over a finite planning horizon, or in a dynamic mode over a rolling horizon. We have modeled the static problem as an ILP, and we have developed four policies for the dynamic case. These policies were extensively tested and compared over a large set of instances based on real location data extracted from a network in California and another one in Texas.

The four policies we have implemented for the dynamic case are called threshold, optimistic, pessimistic, and minimax regret. They mainly differ in the moments and levels at which the orders enter the 3PL supply chain. The problem is to achieve a suitable trade-off between loading orders on vehicles quickly, or holding them for later delivery. We found that for some cost structures with relatively low transportation costs, namely I and J, the threshold policy

Transportation cost variation (%)	Total cost (%)	Sorting(%)	Outsourcing (%)	Transportation (%)	Transferred (%)	Waiting (%)	Truckload (%)
−50	−27.07	39.65	20.50	39.85	16.31	14.86	93.95
−45	−24.18	38.01	20.19	41.80	15.74	14.64	94.04
−40	−21.32	36.45	20.10	43.45	15.03	14.22	94.04
−35	−18.48	35.33	19.78	44.89	15.86	13.30	94.68
−30	−15.68	34.18	19.69	46.13	14.48	14.17	95.05
−25	−12.91	33.10	19.40	47.50	14.13	14.21	95.24
−20	−10.16	32.04	19.04	48.92	14.75	13.21	95.33
−15	−7.48	29.83	22.79	47.38	9.54	8.50	95.71
−10	−4.94	28.69	24.14	47.16	7.55	7.59	95.90
−5	−2.46	27.93	23.98	48.09	7.51	7.13	95.99
0	0.00	27.31	23.65	49.04	7.11	7.49	96.37
5	2.45	26.66	23.38	49.96	6.63	7.72	96.57
10	4.87	26.12	23.36	50.52	6.81	7.32	97.05
15	7.27	25.55	22.98	51.47	7.16	6.88	97.14
20	9.66	25.05	22.68	52.27	6.96	7.08	97.44
25	12.04	24.68	22.49	52.83	7.29	7.04	98.12
30	14.40	24.22	22.19	53.59	6.86	7.52	98.32
35	16.76	23.78	21.79	54.43	7.44	7.08	98.52
40	19.11	23.34	21.34	55.33	7.19	7.52	98.62
45	21.46	22.85	20.95	56.20	7.42	7.11	98.52
50	23.82	22.42	20.55	57.03	7.63	6.90	98.52

Table 9: Sensitivity analysis: varying the transportation cost g^k_{jt}.

Sorting cost variation (%)	Total cost (%)	Sorting(%)	Outsourcing (%)	Transportation (%)	Transferred (%)	Waiting (%)	Truckload (%)
−50	−13.95	17.11	22.67	60.22	14.65	13.47	98.42
−45	−12.49	18.38	22.37	59.24	14.55	12.79	98.03
−40	−11.05	18.92	25.89	55.19	8.29	9.62	97.83
−35	−9.65	20.12	25.86	54.02	8.26	9.01	97.83
−30	−8.26	21.20	25.58	53.23	7.17	9.27	97.44
−25	−6.87	22.29	25.42	52.30	8.46	7.49	97.34
−20	−5.49	23.38	24.98	51.64	8.23	7.58	97.14
−15	−4.11	24.38	24.28	51.34	7.54	8.15	96.66
−10	−2.74	25.40	24.08	50.52	6.96	8.36	96.57
−5	−1.37	26.32	23.83	49.85	7.30	7.52	96.28
0	0.00	27.31	23.65	49.04	7.11	7.49	96.37
5	1.36	28.18	23.32	48.50	6.71	7.50	96.09
10	2.72	28.98	23.14	47.88	6.69	6.90	95.80
15	4.08	29.76	23.04	47.19	6.23	6.74	95.52
20	5.42	30.64	22.64	46.72	6.87	6.12	95.43
25	6.76	31.48	22.31	46.20	6.63	6.31	95.33
30	8.11	32.26	22.19	45.56	6.47	6.15	95.15
35	9.45	33.08	21.92	45.00	6.35	6.26	95.15
40	10.79	33.74	21.62	44.64	6.39	5.84	94.78
45	12.12	34.45	21.28	44.28	6.00	6.11	94.59
50	13.44	34.98	21.26	43.77	6.22	5.09	94.23

Table 10: Sensitivity analysis: varying the sorting cost.

is not so much worse than some of the more sophisticated policies relying on the solution of ILPs. For these cost structures, this policy may be attractive to dispatchers in the sense that it is based on a simple decision rule and on a single parameter θ_k associated with each level k (to ship or not to ship depending on the threshold value θ_k), as opposed to being based on the solution of an ILP. This is particularly true if it is implemented with a standard combination $(\bar{\theta}_1, \ldots, \bar{\theta}_\ell)$ over all possible network topologies, lengths of the planning horizon, and demand structures. However, when the transportation costs predominate, as in the cost structure H, the

Outsourcing cost variation (%)	Total cost (%)	Sorting(%)	Outsourcing (%)	Transportation (%)	Transferred (%)	Waiting (%)	Truckload (%)
−50	−12.59	31.04	15.07	53.89	5.47	6.28	97.14
−45	−11.27	30.68	16.19	53.13	6.20	5.94	97.24
−40	−9.98	30.32	17.22	52.46	6.43	6.19	97.44
−35	−8.69	29.97	18.23	51.80	6.44	6.59	97.53
−30	−7.42	29.50	19.28	51.22	6.43	6.50	97.34
−25	−6.15	29.04	20.18	50.78	6.58	6.37	97.05
−20	−4.89	28.59	21.09	50.32	6.81	6.08	96.76
−15	−3.64	28.19	21.90	49.91	6.31	6.68	96.57
−10	−2.41	27.88	22.70	49.42	6.83	6.55	96.57
−5	−1.19	27.56	23.25	49.19	6.89	6.95	96.47
0	0.00	27.31	23.65	49.04	7.11	7.49	96.37
5	1.17	27.01	24.15	48.84	8.56	6.47	96.18
10	2.32	26.87	24.53	48.60	7.79	8.26	96.37
15	3.36	28.17	19.37	52.46	14.42	14.46	96.57
20	4.21	27.90	19.57	52.53	15.32	13.77	96.18
25	5.06	27.74	19.92	52.33	14.64	14.99	96.18
30	5.89	27.53	20.43	52.04	14.72	15.05	96.18
35	6.71	27.38	20.63	51.98	13.91	16.45	96.09
40	7.51	27.26	20.84	51.90	15.33	15.67	96.09
45	8.31	27.15	21.13	51.72	16.38	15.19	96.09
50	9.09	27.03	21.45	51.52	15.99	16.13	96.18

Table 11: Sensitivity analysis: varying the outsourcing cost.

threshold policy is not competitive with the other policies. Overall, as we have demonstrated, the best choice is the optimistic policy which exhibits a stable and near-optimal behavior over a wide range of topologies, cost structures, and demand patterns.

Our computational experiments have shown that for all policies and instances, the ILP could be solved within less than one second, which means that the managers can solve the problem almost instantaneously. Furthermore, three of the four policies developed for the dynamic case yield deviation values around 1% with respect to optimal solution costs obtained from the static ESP.

Possible extensions of our model include the consideration of multiple objectives, for example customer satisfaction related to ahead of schedule deliveries. In the same vein, another meaningful variant could be to allow the possibility of not meeting deadlines and penalize lateness.

References

Aguezzoul A, 2014. Third-party logistics selection problem: A literature review on criteria and methods. *Omega* 49, 69–78.

Albareda-Sambola M, Fernández E, Laporte G, 2014. The dynamic multiperiod vehicle routing problem with probabilistic information. *Computers & Operations Research* 48, 31–39.

Angelelli E, Bianchessi N, Mansini R, Speranza MG, 2009. Short term strategies for a dynamic multi-period routing problem. *Transportation Research Part C: Emerging Technologies* 17, 106–119.

Archetti C, Jabali O, Speranza MG, 2015. Multi-period vehicle routing problem with due dates. *Computers & Operations Research* 61, 122–134.

Ayanso A, Diaby M, Nair SK, 2006. Inventory rationing via drop-shipping in Internet retailing: A sensitivity analysis. *European Journal of Operational Research* 171, 135–152.

Berbeglia G, Cordeau J-F, Laporte G, 2010. Dynamic pickup and delivery problems. *European Journal of Operational Research* 202, 8–15.

Buldeo Rai H, Verlinde S, Merckx J, Macharis C, 2017. Crowd logistics: an opportunity for more sustainable urban freight transport? *European Transport Research Review* 9:39.

Chen A, 2017. Large-scale optimization in online-retail inventory management. PhD thesis, Massachusetts Institute of Technology, Massachusetts, USA.

Chen Y-K, Chiu F-R, Lin W-H, Huang Y-C, 2018. An integrated model for online product placement and inventory control problem in a drop-shipping optional environment. *Computers & Industrial Engineering* 117, 71–80.

Cheng X, Gou Q, Yue J, Zhang Y, 2019. Equilibrium decisions for an innovation crowdsourcing platform. *Transportation Research Part E: Logistics and Transportation Review* 125, 241–260.

Clement, J. Retail e-commerce sales worldwide from 2014 to 2021 (last accessed July 18, 2019), https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/.

Cohen R, 2018. How Amazon's delivery logistics redefined retail supply chains. *Journal of Supply Chain Management, Logistics and Procurement* 1, 75–86.

Daganzo CF, 1984. The length of tours in zones of different shapes. *Transportation Research Part B: Methodological* 18:135–145.

Dayarian I, Crainic TG, Gendreau M, Rei W, 2015. A branch-and-price approach for a multi-period vehicle routing problem. *Computers & Operations Research* 55, 167–184.

Franceschetti A, Jabali O, Laporte G, 2017. Continuous approximation models in freight distribution management. *TOP* 25, 413–433.

Francis PM, Smilowitz KR, Tzur M, 2008. The period vehicle routing problem and its extensions. In: Golden BL, Raghavan S, Wasil EA (Eds.), The vehicle routing problem: Latest advances and new challenges, Springer, Boston, pp. 73–102.

Giri BC, Sarker BR, 2017. Improving performance by coordinating a supply chain with third party logistics outsourcing under production disruption. *Computers & Industrial Engineering* 103, 168–177.

Hanbazazah AS, Abril L, Erkoc M, Shaikh N, 2019. Freight consolidation with divisible shipments, delivery time windows, and piecewise transportation costs. *European Journal of Operational Research* 276, 187–201.

Hemmelmayr VC, Doerner KF, Hartl RF, 2009. A variable neighborhood search heuristic for periodic routing problems. *European Journal of Operational Research* 195, 791–802.

Hu W, Toriello A, Dessouky M, 2018. Integrated inventory routing and freight consolidation for perishable goods. *European Journal of Operational Research* 271, 548–560.

Jayaram J, Tan K-C, 2010. Supply chain integration with third party logistics providers. *International Journal of Production Economics* 125, 262–271.

Kahle D, Wickham H, 2013. ggmap: Spatial visualization with ggplot2. *The R Journal* 5, 144–161.

Lafkihi M, Pan S, Ballot E. 2019. Freight transportation service procurement: A literature review and future research opportunities in omnichannel E-commerce. *Transportation Research Part E: Logistics and Transportation Review* 125, 348–365.

Landete M, Laporte G, 2019. Facility location problems with user cooperation. *TOP* 27, 125–145.

Le TV, Stathopoulos A, Van Woensel T, Ukkusuri SV, 2019. Supply, demand, operations, and management of crowd-shipping services: A review and empirical evidence. *Transportation Research Part C: Emerging Technologies* 103, 83–103.

Lee C-Y, Çetinkaya S, Jaruphongsa W, 2003. A dynamic model for inventory lot sizing and outbound shipment scheduling at a third party warehouse. *Operations Research* 51, 735–747.

Li G, Zhang X, Liu M, 2019. E-tailer's procurement strategies for drop-shipping: Simultaneous vs. sequential approach to two manufacturers. *Transportation Research Part E: Logistics and Transportation Review* 130, 108–127.

Ma NL, Tan KW, 2019. Reducing carbon emission through container shipment consolidation and optimization. *Journal of Traffic and Transportation Engineering* 7, 111–121.

Macrina G, Di Puglia Pugliese L, Guerriero F, Laporte G, 2020. Crowd-shipping with time windows and multiple deliveries. *Computers & Operations Research* 113, 104806.

Ortega F, Wolsey LA, 2003. A branch-and-cut algorithm for the single-commodity, uncapacitated, fixed-charge network flow problem. *Networks* 41, 143–158.

Ortiz-Astorquiza C, Contreras I, Laporte G, 2019. An exact algorithm for multi-level uncapacitated facility location. *Transportation Science* 53, 1085–1106.

Qin H, Zhang Z, Qi Z, Lim A, 2014. The freight consolidation and containerization problem. *European Journal of Operational Research* 234, 37–48.

Satır B, Safa Erenay F, Bookbinder JH, 2018. Shipment consolidation with two demand classes: rationing the dispatch capacity. *European Journal of Operational Research* 270, 171–184.

Stenius O, Marklund J, Axsäter S, 2018. Sustainable multi-echelon inventory control with shipment consolidation and volume dependent freight costs. *European Journal of Operational Research* 267, 904–916.

Tang CS, Veelenturf LP, 2019. The strategic role of logistics in the industry 4.0 era. *Transportation Research Part E: Logistics and Transportation Review* 129, 1–11.

Tyan JC, Wang F-K, Du TC, 2003. An evaluation of freight consolidation policies in global third party logistics. *Omega* 31, 55–62.

Ulmer MW, Soeffker N, Mattfeld DC, 2018. Value function approximation for dynamic multi-period vehicle routing. *European Journal of Operational Research* 269, 883–899.

Wen M, Cordeau J-F, Laporte G, Larsen J, 2010. The dynamic multi-period vehicle routing problem. *Computers & Operations Research* 37, 1615–1623.

Yu DZ, Cheong T, Sun D, 2017. Impact of supply chain power and drop-shipping on a manufacturer's optimal distribution channel strategy. *European Journal of Operational Research* 259, 554–563.

Yu Y, Wang X, Zhong RY, Huang GQ, 2016. E-commerce logistics in supply chain management: Practice perspective. *Procedia CIRP* 52, 179–185.