European Journal of Operational Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Distribution network deployment for omnichannel retailing

Ayşe N. Arslan[a,*], Walid Klibi[b,d], Benoit Montreuil[c,d]

[a] Univ Rennes, INSA Rennes, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France
[b] The Centre of Excellence for Supply Chain Innovation & Transportation (CESIT), Kedge Business School, Bordeaux, France
[c] H. Milton Stewart School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, United States
[d] Physical Internet Center, Supply Chain & Logistics Institute, Georgia Institute of Technology, Atlanta, United States

ARTICLE INFO

Article history:
Received 2 May 2019
Accepted 8 April 2020
Available online xxx

Keywords:
Omnichannel distribution
Advanced inventory deployment
Stochastic programming

ABSTRACT

This article studies the distribution problem of brick-and-mortar retailers aiming to integrate the online channel into their operations. The article presents an integrated modeling approach addressing the online channel-driven distribution network deployment (e-DND) problem under uncertainty. The e-DND involves decisions on operating fulfillment platforms and on assigning a fulfillment mission to those platforms, while anticipating the revenues and costs induced by order fulfillment, replenishment, delivery, and inventory holding. To model this problem while taking into account the uncertain nature of multi-item online orders, store sales, and capacities, a two-stage stochastic program with mixed-integer recourse is developed. Two alternative deployment strategies, characterized by allocation of orders, inventory positioning, delivery schema, and inbound flow pattern decisions, are investigated using this model. The first deployment strategy investigates the ship-from stores practice where the on-hand inventory is used for all sales channels. The second deployment strategy additionally considers the advanced positioning of inventory at a fulfillment center in the urban area where the online orders are requested. To solve the two-stage stochastic model with integer recourse, an exact solution approach combining scenario sampling and the integer L-shaped method is proposed. Numerical results, inspired by the case of a European retailer, are provided to evaluate the performance of the deployment strategies and the efficiency of the proposed solution approach.

1. Introduction

Tompkins and Ferrell (2016) reported that e-commerce already generates about 7% of total retail sales, and is expected to increase its market share in the upcoming years. Joerss, Schroder, Neuhau, Klink, and Mann (2016) investigated the strong fundamental drivers supporting the growth of same-day delivery. They reported that 23% of consumers are willing to pay extra for same-day delivery whereas 70% of consumers are content with the cheapest form of home delivery. In practice, some retailers have already started reshaping their distribution strategies to cope with omnichannel retailing, the main pillar being the integration of online and in-store inventories. World-class retailers such as Walmart, Amazon, and JD.com recently started engaging in advanced deployment strategies and/or structural adaptation strategies to improve their distribution networks. The advanced deployment strategies encompass innovative practices such as ship-from store, advanced stocks, anticipatory shipments, and smart urban fulfillment. For instance, Walmart has been practicing the ship-from store policy for a while and plans to convert twelve Sam's Club stores to e-commerce fulfillment centers (FCs) to support the rapid e-commerce growth Deborah (2018). Structural adaptation strategies consist of adapting the design of the network by deciding to operate additional urban consolidation or fulfillment centers. For instance, companies such as Flexe (www.flexe.com) and darkstore (www.darkstore.com) have recently specialized in offering flexible warehousing and fulfillment spaces for the retail sector in the US.

Despite these recent attempts at adaptation, with the advent of the online channel, some retailers have experienced critical operational inefficiencies and on-time delivery issues when using their as-is distribution networks. Inspecting the current distribution networks of several retailers, one might observe that, in general, they operate a single centralized warehouse or a set of market-dedicated continental/regional warehouses (Martel & Klibi, 2016). Usually, the management of these warehouses, in terms of throughput and inventory level, depends on the in-store demand dynamics, the inventory policies, and the replenishment cycles. With the emergence of the online channel, the managers of

* Corresponding author.
E-mail address: Ayse-Nur.Arslan@insa-rennes.fr (A.N. Arslan).

https://doi.org/10.1016/j.ejor.2020.04.016

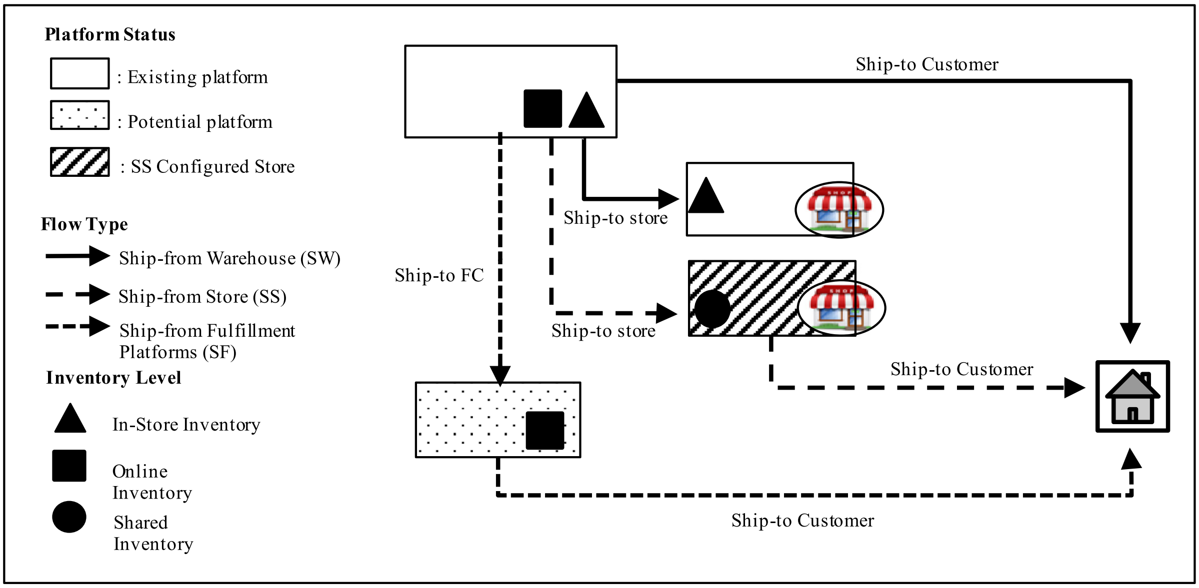


Fig. 1. Potential network structure with alternative deployment strategies.

these warehouses are often asked to provide e-fulfillment capabilities. However, these storage platforms are not designed and located to operate as FCs, and are not specifically capable of covering next-day and/or same-day deliveries to urban areas in a reliable manner. Even if some companies manage to operate next-day deliveries in dense markets, ensuring the deliveries in an efficient and reliable way still poses a challenge in several business contexts. This could be due to travel distances in large markets, the incapability of service providers in some regions, or the synchronization between fulfillment and transportation operations. Such centralized distribution network structures clearly reduce retailers' capabilities to capture online demand requesting fast delivery services in most of the cases. This is accentuated by the two-speed operations required by physical and online sales. The former is usually regulated by store replenishment, scheduled at a frequency ranging from weekly to daily. The latter is an on-demand fulfillment and delivery service that typically usually requires a shorter response time, usually ranging from a few hours to a few days.

Recently, some papers have focused on identifying new challenges and opportunities created by omnichannel retailing (Bell, Gallino, & Moreno, 2014; Chopra, 2018; Hübner, Holzapfel, & Kuhn, 2016). In this context, Bektaş, Crainic, and Van Woensel (2017) and Mancini, Gonzalez-Feliu, and Crainic (2014) reviewed multi-tier network structures in a city logistics context and highlighted inventory, fulfillment, and delivery planning among the areas of further research identified. Savelsbergh and Van Woensel (2016) discussed the challenges behind the expansion to multi-tier distribution systems in terms of their prohibitive costs. In the same way, the concept of hyperconnected distribution that promotes the shift to a distributed multi-segment distribution approach is appended to a conceptual urban context in Crainic and Montreuil (2016), to an omnichannel B-to-C context in Montreuil (2016) and, to a hyperconnected urban fulfillment and transportation system in Kim, Montreuil, Klibi, and Kholgade (2020). Despite a pronounced potential value, there is, to the best of our knowledge, little work on the modeling of these innovative practices and on their improvement of current distribution systems in terms of cost and delivery speed.

Inspired by the case of a large retailer in the cosmetics industry that sells a broad catalog of items through physical and online channels, this article studies the integration of in-store and online channels from a distribution system perspective. Hereafter, the notation e-DND stands for the distribution network deployment (DND) problem integrating the e-commerce channel to physical sales. Fig. 1 illustrates the distribution network structure and the potential deployment strategies that could be employed. This figure highlights all the considered potential ship-to allocation flows and fulfillment inventory locations in a typical customer zone covered by two existing stores, and one potential urban fulfillment center. The tactical decisions involve the selection of the subset of stores to use as omnichannel-stores during the planning horizon and the decision to operate a subset of urban fulfillment platforms to support online order fulfillment. A key feature is that the mission of the selected stores and urban fulfillment platforms must remain the same for each operational day of the planning horizon. The modeling approach applied to the e-DND problem relies on the location-allocation model, which is widely used to formulate production-distribution networks (Martel & Klibi, 2016). However, it is clear that the anticipation of the order fulfillment problem (OFP) in the e-DND decision model requires the consideration of a more refined time granularity, the modeling of an uncertain multi-item order process and also the inclusion of inventory and transportation specifications arising in an urban context. For this reason, the decision model is built as a hierarchical decision problem due to the temporal hierarchy between the deployment decisions and the order fulfillment decisions. One of the main difficulties underlying the e-DND problem is that, at the time the deployment decisions are being made, several crucial parameters, such as store capacities to online fulfillment, in-store demand, and online orders, are not known with certainty. It is important to take this uncertainty into account in the mathematical model, as operations of the system such as assignment of online orders to ship-from platforms and replenishment quantities are inherently dependent on these parameters.

In this work, three deployment strategies are investigated for the e-DND problem, namely, ship-from warehouse (SW), ship-from stores (SS), and ship-from urban fulfillment platforms (SF). Each deployment strategy is represented by a set of allocation decisions, selecting to keep fulfilling online orders from the warehouse, to operate a dedicated urban FC, and/or to assign the fulfillment of online orders to a subset of the stores located in the city. The first strategy, referred to as ship-from warehouse (SW), corresponds to the baseline strategy in which the online channel fulfillment is managed at the central warehouse. This strategy is distinguished in Fig. 1 by the solid arcs separating the store replenishment flow from the ship-to consumer flow for the in-store and online channels, respectively. The second deployment strategy, referred to as ship-from stores (SS), takes advantage of the existing physical network by turning certain store locations to ship-from points for the online channel. In Fig. 1, SS deployment strategy is represented by the dashed arcs underlining the joint replenishment of the selected store and the shared inventory. The third deployment strategy considers an external set of urban fulfillment platforms and will be referred to as ship-from fulfillment platforms (SF). In the SF strategy, the stores (as in the SS strategy) and additionally a set of dedicated-to-online-channel urban FCs are considered as potential fulfillment platforms. In Fig. 1, the SF deployment strategy is highlighted by the dotted arcs in that some online orders are satisfied via an urban fulfillment platform with dedicated inventory.

The contribution of this article is threefold. First, we provide a description of the e-DND problem under uncertainty and formulate it as a two-stage stochastic program with mixed-integer recourse. A scenario-based approach is used to characterize the uncertain environment in terms of the multi-item ordering process, in-store demands, and the fulfillment capacity. We consider the integration of online and physical inventories in a stochastic setting and propose an approximation of the continuous order fulfillment problem with the treatment of multi-item stochastic orders. To the best of our knowledge, this article is the first to consider omnichannel store selection and urban fulfillment decisions to operate as ship-from locations under response time constraints, and to integrate multi-period order fulfillment and inventory assignment decisions under demand and capacity uncertainty. Second, we propose a solution approach based on the integer L-shaped method using a branch-and-cut procedure (see Laporte & Louveaux, 1993) with an alternating cut strategy (see Angulo, Ahmed, & Dey, 2016) and improved optimality cuts (see Birge & Louveaux, 2011). Numerical results that show the efficiency of this approach in solving realistically sized instances are presented. Third, we study two advanced deployment strategies that are ship-from stores and ship-from urban fulfillment platforms in comparison to the ship-from warehouse. We provide managerial insights on the capabilities of these deployment strategies in increasing profitability, reducing lost sales, and responding to fast delivery demand. A key observation in our results is that the introduction of the ship-from stores (SS) strategy immediately leads to a shift in the fulfillment of online orders. Another important insight is that the opening of a second tier of urban fulfillment platforms is highly dependent on the associated usage cost, customer shift to faster delivery, the warehouse-city proximity to offer fast response capability, and the availability of transportation capacity.

The remainder of the article is organized as follows. In Section 2 we present related work from the literature highlighting the contribution of the current work. In Section 3, we detail our modeling approach to uncertainty and the deployment strategies proposed. We then present a mathematical model to optimize the deployment of a distribution network as well as the order fulfillment and inventory management under uncertainty. In Section 4, we present the solution approach and implementation details, and finally in Section 5 we present computational results and managerial insights. We conclude in Section 6 with perspectives on future research opportunities.

2. Literature review

The omnichannel operations research stream is still at its infancy. Based on a customer-focused perspective Bell et al. (2014) introduced four alternatives in omnichannel retail: traditional retail, shopping and delivery hybrid, online retail plus showrooms, and pure-play e-commerce. Gallino and Moreno (2014) presented the integration of online and offline channels in retail. The authors used an empirical analysis to identify patterns in channel-shift behavior that arise from the integration of the online sales with in-store sales. Chopra (2018) defined omnichannel retailing as the use of a variety of channels to interact with customers and fulfill their orders. The author discussed relative costs for the omnichannel alternatives, such as in-store channel, customer delivery, and "buy online and pick at store" (BOPS) channel. Agatz, Fleischmann, and Van Nunen (2008) presented three e-fulfillment structures: integrated fulfillment, dedicated fulfillment, and store fulfillment. The last refers to picking online orders from regular retail shelves for a separate, dedicated delivery. A recent review of e-fulfillment and distribution in omnichannel retailing is found in Melacini, Perotti, Rasini, and Tappia (2018). This review highlights the lack of literature on two key aspects of omnichannel operations: the evolution of retail distribution networks and the logistics role played by stores in the delivery process.

Many studies in the literature rely on empirical studies to explore this novel topic. An early-stage survey-based analysis of the distribution strategy to adopt for the online channel is provided by De Koster (2003), who discussed the opportunity to use existing stores. The author mentioned key factors determining the distribution strategy, such as the assortment type and width, the delivery time and area, and the transport policy (in-house versus outsourcing). He argued that the distribution channel for internet customers should be integrated with existing operations, but for larger internet order volumes, direct-delivery distribution centers are more appropriate. A recent exploratory study, conducted by Hübner et al. (2016) with 33 retailers, reported the five most important areas for fulfilling distribution requirements: developing and optimizing modes of delivery, increasing delivery speed, inventory transparency, optimizing the cross-channel processes in distribution centers and stores, and inventory integration and allocation. The qualitative review proposed by Ishfaq, Defee, Gibson, and Raja (2016) ranks, additionally, the store fulfillment/ship-from store among the top five areas listed by practitioners. In the context of store-based retailing, the authors mentioned the impact of the configuration and capabilities of the distribution network on the omnichannel fulfillment strategy. This latter would be enabled in their analysis by forward placement of inventory and a store-replenishment schedule. In the same way, Hübner, Wollenburg, and Holzapfel (2016) studied empirical data from German retailers to investigate specifically the transition from multi-channel to omnichannel logistics. A similar study conducted by Lang and Bressolles (2013) with French retailers added the perspective of contrasting the economic performance with the customer expectation. All these empirical investigations appeal to the need to integrate online and in-store operations and to investigate the ship-from store option. They also highlight the importance of cost-service trade-offs, the dynamics of online retailing, and jointly considering order fulfillment, inventory assignment, and delivery decisions.

Despite the empirical studies highlighted above, only few quantitative models were proposed in the omnichannel distribution and multi-channel distribution network design contexts (Agatz et al., 2008; Hübner et al., 2016). Only in recent years, some quantitative models were developed, aiming to tackle certain features of omnichannel retailing. These models can be categorized as fulfillment/delivery models and inventory models, and focus mainly on inventory aggregation effects and rely on multi-echelon inventory theory (Govindarajan, Sinha, & Uichanco, 2018;Harsha, Subramanian, & Uichanco, 2019).

When decisions on opening or building new distribution/fulfillment centers are involved, location decisions need to be considered. In this context, two-echelon location-allocation models

Please cite this article as: A.N. Arslan, W. Klibi and B. Montreuil, Distribution network deployment for omnichannel retailing, European Journal of Operational Research, https://doi.org/10.1016/j.ejor.2020.04.016

(Ben Mohamed, Klibi, & Vanderbeck, 2020) and two-echelon routing models (Cuda, Guastaroba, & Speranza, 2015) cope with the current expansion of distribution operations in retail and/or urban contexts. However, these extended models rely mostly on inter-facility flow decisions, and/or do not cope with the notion of inventory policy; neither consider the dynamics of fulfillment operations as would be required by an omnichannel context. The emergence of multi-tier network systems was reviewed in Bektaş et al. (2017). Savelsbergh and Van Woensel (2016) discussed the prohibitive cost of expansion to multi-tier networks but underlined that the strategy of making several short trips may be preferable to long trips. In the same way, Crainic and Montreuil (2016), introduced the concept of hyperconnected city logistics to promote the shift to a distributed multi-segment distribution approach, which builds on the availability of a shared network of logistics resources. This concept is further discussed in an omnichannel B-to-C context in Montreuil (2016). Kim et al. (2020) studied last-mile delivery of large items and demonstrated the potential by concurrently improving often opposing measures: economic efficiency, service capability, and sustainability (undesirable atmospheric emissions are reduced by 45% to 50%).

When location decisions are not considered, deemed as pre-set and fixed, then assignment decisions are the main choices to consider. Torabi, Hassini, and Jeihoonian (2015), studied an inventory fulfillment-allocation problem with transshipment decisions between a predetermined set of FCs for the online channel. They developed a mixed-integer programming model to optimally fulfill customer online orders for a fixed time window under a deterministic setting. Paul, Agatz, Spliet, and de Koster (2017) studied the routing problem of a grocery retailer in which goods that are ordered online can be picked up from the stores. The authors modeled and solved the related vehicle routing problem with order consolidation but did not investigate the decisions of store selection and inventory positioning. Acimovic and Graves (2014) studied the order fulfillment problem (OFP) and stated that this complex problem should be modeled in theory as a continuous-time dynamic program. They then developed a heuristic for the OFP, embedding a linear program for the transportation problem to provide an approximation of the immediate and the expected outbound shipping cost. They considered multi-item orders but assumed that inventory positioning cannot be revoked and that FCs are predetermined. Further references dealing with multi-echelon networks and dynamic delivery systems are found in Savelsbergh and Van Woensel (2016).

A key component here is the modeling approach employed in order to integrate the inventory policy or to mimic the inventory consumption and replenishment. The modeling approach proposed by Alptekinoğlu and Tang (2005) for multi-channel distribution systems is based on the base stock model. The authors consider a multi-depot assignment problem that integrates a near-optimal approximation of the inventory allocation to the depots. This approximation is based on an equal-fractile allocation rule in which the problem is decomposed to a set of single depots. However, the proposed approach is suitable for the distribution facilities echelon, but is less convenient for store-based delivery and storage, especially when shared inventory among channels need to be considered. Bendoly, Blocher, Bretthauer, and Venkataramanan (2007) propose a static model minimizing the inventory costs of a two-echelon system that determines whether it is beneficial to realize inventory-pooling when using a centralized facility and/or decentralized store satellites. The authors assume that inventory intended for both in-store and online sales is combined at a central facility or store satellites and that demand is normally distributed. Under these assumptions, they prove the existence of a threshold online demand level determining when decentralization of online inventory to the retail stores is preferable. In extension, a cost-minimization dual-channel assignment problem is formulated in Bretthauer, Mahar, and Venakataramanan (2010), which consists of finding the optimal inventory level at the central warehouse and deciding locations that will fulfill each demand. The authors consider the modeling of an order-up-to-level inventory policy at the warehouse. However, the notion of service level is not integrated and their static model does not allow capturing inventory management at stores. In addition, the proposed model assumes that only demand is uncertain and is normally distributed. Recently, Ishfaq and Bajwa (2019) proposed a profit maximization model for the online sales fulfillment problem. The model explicitly considers profitability affecting factors such as the choice of fulfillment option and sale price and the possibility to drop ship-from vendors. The authors proposed a non-linear deterministic model with several seasonal replenishment periods and solved it heuristically. Finally, another class of models builds on the newsvendor problem with the surrounding objective of inventory optimization for one period under a stochastic demand (see Li, Lu, & Talebian, 2015; Gao & Su, 2016; Govindarajan et al., 2018), which is out of the scope of this article.

To the best of our knowledge, no explicit modeling approach exists in the literature that considers omnichannel store selection and urban fulfillment decisions to operate as ship-from locations under response time constraints and that integrates multi-period order fulfillment and inventory assignment decisions under demand and capacity uncertainty.

3. Modeling approach

The e-DND is a hierarchical decision problem due to the temporal hierarchy between the tactical deployment decisions and the operational order fulfillment decisions. The tactical decisions need to be made before the realization of uncertain parameters, whereas the operational decisions can be adapted to the observed realization. In order to account for the nature of the problem in the optimization framework, we consider a two-stage scenario-based stochastic model in which the scenarios represent realizations of uncertain parameters, namely, a set of orders, physical demands, and store capacities over the planning horizon. In this section, we first elaborate on the assumptions underlying our modeling approach. We then discuss the modeling of uncertainty to obtain the scenarios used in our model and present our mathematical model.

3.1. Description of deployment strategies and modeling assumptions

As illustrated in Fig. 1, our modeling framework for the distribution network builds on the inclusion of three complementary deployment strategies that are hereafter compared.

Ship-from warehouse (SW) is the baseline strategy in which the online channel fulfillment is managed at the existing central warehouse. Although this strategy benefits from economies of scale and does not incur additional investment costs, it tends to increase the long freight traffic toward the city centers. Additionally, SW establishes a priori that the inventory positioning will not cover future online orders requiring a short response time, mainly due to the distance between the warehouse and the city. In implementing this strategy in our mathematical model, we assume that the central warehouse has unlimited inventory. This assumption is reasonable because the warehouse has a much larger capacity to stock items and generally relies on an adequate replenishment policy with suppliers. We assume that a fixed charge composed of fixed, replenishment, and holding costs applies to each order shipped from the warehouse, as well as a per-item picking cost and a transportation cost.

Ship-from stores (SS) is an emerging deployment strategy aimed at taking advantage of the existing network of stores. The

Please cite this article as: A.N. Arslan, W. Klibi and B. Montreuil, Distribution network deployment for omnichannel retailing, European Journal of Operational Research, https://doi.org/10.1016/j.ejor.2020.04.016

selection of a store to be part of the fulfillment tier of the network requires an investment related to the store backroom layout and the setting up of the inventory management system. This is reflected in a fixed cost in our mathematical model. Additionally, the capacity of physical stores to stock and pick and pack online orders is limited, expressed in number of orders treated per day. These capacities are estimated based on historical fulfillment data and are not subject to change since the physical space, the number of employees, and the fulfillment technologies are fixed. On a day-to-day basis, the order management system integrates the inventory of the warehouse as well as the inventory of the selected subset of stores before allocating the order to the most profitable location. When it is profitable, a given order is assigned to a store, the ordered items are fulfilled (picked and packed) by the store employees, and on-demand transportation is requested to deliver the order to the customer location. We assume that per-product replenishment, holding, and backlogging costs apply to each order shipped from stores, as well as a per-item picking cost and a transportation cost. In implementing this strategy in our mathematical model, we assume that the replenishment plan from the central warehouse is predetermined for the entire horizon according to a weekly calendar such that each store knows in advance its replenishment periods. However, the replenishment quantity of each item at each replenishment period is to be optimized. This assumption reflects the current contract setting between the retailer and the carrier. In our numerical study, we conduct a sensitivity analysis experiment to explore the effects of relaxing this assumption.

Ship-from urban fulfillment platforms (SF) is an alternative deployment strategy that considers an intermediate tier of urban fulfillment platforms added to the network structure. The SF strategy is more comprehensive because it complements the SW and SS options with a set of dedicated-to-online-channel urban FCs as potential fulfillment locations. Here, the new urban fulfillment tier is devoted to receiving advanced stocks coming from the central warehouse on large trucks and to fulfilling online orders daily with smaller and environment-friendly vehicles suitable for city distribution. This strategy builds on exploiting on-demand storage services (Montreuil, 2017) and thus involves operating urban FCs shared with other retailers or cooperating with a third-party logistics partner already deploying urban FCs. This assumes that additional costs associated with exploiting the FCs are added in the form of a usage cost, as provided for instance by Flexe (www.flexe.com). These FC usage costs are sensitive to the location of the platforms (city center versus peri-urban), but they are not assimilated into facility opening costs. It is clear that relying on this additional fulfillment capability offers less capacity constraints than stores and lower picking and packing unitary costs. We assume that per-product replenishment and holding costs apply to each order shipped from an FC, as well as a per-item picking cost and a transportation cost. In implementing this strategy in our mathematical model, we assume that the dedicated capacity to the retailer in the urban FC is fixed. Indeed, the urban FCs are dedicated to picking and packing of online orders. In this sense, their capacity can be considered deterministic as it is not affected by the physical sales. We remark that, in this work, we exclude from investigation the strategy of operating a dedicated e-commerce distribution center that is not supplied from the cross-channel warehouse as considered in a multi-channel context (see Hübner et al., 2016).

In implementing the deployment strategies just outlined, our mathematical model builds on the following additional assumptions:

- The inbound transportation is managed by a contracted carrier and is capacitated, expressed in items shipped per day. We assume that the available transportation capacity dedicated by the carrier is fixed but the used capacity (i.e., truck fill rate) may vary on a daily basis. Given that dedicated store replenishment quantities are not in full truckloads, we consider that the online channel piggybacks on the residual transportation capacity, reflecting cost benefits from joint replenishment of both channels.
- Orders fulfilled from ship-from locations are delivered with on-demand delivery, which comes at a higher transportation cost. This assumption reflects the presumed operational characteristics of the on-demand transportation in a last-mile context.
- Only completed orders are shipped to the customer, which means that all the items of a given order should be in stock in the allocated ship-from location. More specifically, order splitting is not possible, and transshipment between stores or from urban FCs to stores is not permitted.
- Online orders that cannot be fulfilled within the requested response time will be lost (*i.e.*, no postponement). However, a physical demand for a given product could be backlogged if a stock-out is experienced on that day. This assumption is based on an observed practice in which store employees can offer customers the opportunity to order the backlogged product to be picked up from the store at a later date.

3.2. Modeling uncertainty

One of the main difficulties behind e-DND is that, at the time the deployment decisions are made, several crucial parameters such as capacities and demand are not known with certainty. It is important to take this uncertainty into account in the mathematical model, which is done in the form of a set of potential scenarios. These scenarios are defined a priori from the probability distributions of all the associated random processes and the correlations between them.

Let $\mathcal{S}$ denote the set of existing stores indexed by s. The first uncertain parameter we consider is the daily capacity of a given store s, denoted by b_{st}, for online order fulfillment. This capacity is dependent on the availability of the store workers on a daily basis to pick and pack the online orders, which itself depends on the store traffic induced by physical shoppers. To mimic this process, we categorize the physical stores into two groups: ones with high physical customer traffic and ones with low physical customer traffic. We then assume that the average fulfillment capacity, b_s, dependent on the store category, is known and is adjusted as a function of the physical customer traffic.

Let $\mathcal{P}$ denote the set of products. We secondly consider the physical demand d_{st}^p for each product p, at each store s, in each period t to be uncertain. Expected demand λ_s^p, expressed as the number of products sold per day, can be derived for each product p at each store s from historical data. We assume that a Poisson distribution based on this parameter governs the demand for each product. However, the actual realization of demand is once again considered to be dependent on the category of the store as well as the physical customer traffic during the period.

As explained, the two uncertain parameters we have so far, that is, physical demands and store capacities, depend both on the physical customer traffic realization in the period and are correlated through this parameter. To model this correlation, we introduce another stochastic parameter called the market condition, denoted by m. The details on how the market condition is characterized and generated are given in Section A.1 of the Appendix A. We assume that there is a continuous probability distribution governing m, and we use this distribution to generate a market condition, m_t, in each period. The market condition, m_t, is then used to adjust the physical order arrival rates as $\lambda_s^p(m_t/\bar{\mu})$ and the store capacities as $b_s(\bar{\mu}/m_t)$, where $\bar{\mu}$ is the expectation of the market condition. Additionally, if store s is categorized as high physical customer traffic, and if $m_t > \bar{\mu}$, then the capacity is further multiplied by a factor $0 < f < 1$ to reflect the fact that popular stores are even more vulnerable to surges in the demand.

In addition to the physical demand and capacity processes just outlined, a compound stochastic process is considered to shape the uncertainty of the multi-item online order process. In our mathematical model, an order is characterized by its arrival time, its customer location, its requested response time, and the set of products within the order.

First, each identified customer location is characterized by an order arrival rate that expresses the frequency of the customers in the location to order items online. We assume that the arrival of online customers from each customer location $l \in \mathcal{L}$ is defined by a random variable, τ_l. This random variable follows a Poisson process with expected arrival rate λ_l. Therefore the inter-arrival time distribution between two consecutive orders in customer location l is exponential with parameter $\frac{1}{\lambda_l}$ (*e.g.*, expressed in hours). Next, when an order occurs, the expected response time is shaped by a random variable, denoted by θ with the associated discrete distribution function $F(\cdot)$. We consider the following three response time options : next-day delivery ($1D$, $24h$), two-day delivery ($2D$, $48h$), and three-day delivery ($3D$, $72h$). Together these processes define the arrival time, location, and requested response time of an order.

Second, the set of items, among the portfolio of products $\mathcal{P}$, included in the order and the quantity of each item, are characterized by a stochastic compound process. In this case, a multinomial distribution is combined with a normal distribution to define the products requested and their quantities, respectively. We assume that the random variable associated with including product p in an order takes the value 1 (success) with inclusion probability α_p, $p \in \mathcal{P}$. These probabilities can be estimated by sales and marketing experts for each product p, based on available data. Probabilities α_p could also be linked to an online sales-based ABC classification of products. Finally, for each item p in the order, the ordered quantity is characterized by a random variable n_p. This random quantity n_p is normally distributed with mean μ_p and standard deviation σ_p. Together these processes define the set of products requested within an online order.

All the uncertain parameters and distributions detailed here are provided to a Monte Carlo sampling procedure that is outlined in Section A.2 of the Appendix A. A single call to this procedure generates a single scenario ω, which characterizes a future realization with a set of online orders, physical demand quantities, and store capacities. Our mathematical model works with a set of scenarios, denoted Ω, generated in this manner by repeated calls to the Monte Carlo procedure.

3.3. Notation

Sets

- $\mathcal{T} = \{1, \ldots, T\}$: planning horizon indexed by t
- $\mathcal{S}$: Set of existing stores indexed by s
- $\mathcal{F}$: Set of potential fulfillment centers indexed by f
- $\mathcal{P}$: Set of products indexed by p
- $\mathcal{L}$: Set of customer locations indexed by l
- Ω: Set of scenarios indexed by ω
- $\vec{\mathcal{O}}(\omega)$: Set of online orders placed under scenario ω
- $\vec{\mathcal{O}}_t(\omega)$: Set of online orders placed on day t under scenario ω
- $\vec{\mathcal{O}}_{lt}(\omega)$: Set of online orders placed from customer location l on day t under scenario ω
- $\mathcal{P}(o)$: Set of products that make up online order o

In the following, we let the set of ship-from points $\mathcal{S} \cup \mathcal{F} = \mathcal{SF}$ and index the elements of this set with ς. We additionally refer to the totality of the warehouse (denoted by w) and the ship-from points as supply points $\mathcal{SP} = \mathcal{S} \cup \mathcal{F} \cup \{\mathrm{w}\}$ and use the same index with a slight abuse of notation.

Parameters

- $s^{\max}$: Maximum number of ship-from points that can be deployed
- $c^{\mathrm{f}}_{\varsigma}$: Fixed/usage cost of setting up ship-from point ς to ship online orders
- c_{wo}: Charge applied to order o if it is satisfied from the warehouse includes fixed, replenishment and inventory holding costs
- $c^{\mathrm{r}}_{p\varsigma}$: Unit cost of replenishing product p at ship-from point ς
- $c^{\mathrm{h}}_{p\varsigma}$: Unit inventory holding cost for product p at ship-from point ς
- $c^{\mathrm{t}}_{\varsigma l}$: Unit transportation cost (per order) of servicing customer location l from supply point ς
- c^{b}_{ps} : Unit backlogging cost for product p at store s
- $c^{\mathrm{p}}_{\varsigma}$: Unit picking cost at supply point ς
- τ_{ς}: Set of replenishment periods for ship-from point ς
- k_t: Available transportation capacity on day t
- $r_{\varsigma l}$: Response time for servicing a client situated in customer location l from supply point ς
- π^p: Net sales revenue of product p
- $p(\omega)$: Probability of scenario $\omega \in \Omega$
- $b_{\varsigma t}(\omega)$: Capacity of ship-from point ς for preparing online orders on day t under scenario ω
- $d^p_{st}(\omega)$: Physical demand for product p at store s in period t under scenario ω
- $n^o_p(\omega)$: Number of products of type p in online order $o \in \vec{\mathcal{O}}(\omega)$ under scenario ω
- $n^o(\omega)$: Total number of products in online order $o \in \vec{\mathcal{O}}(\omega)$ under scenario ω
- $\pi^o(\omega)$: Net sales revenue of online order $o \in \vec{\mathcal{O}}(\omega)$ under scenario ω
- $\eta^o(\omega)$: Required response time for online order $o \in \vec{\mathcal{O}}(\omega)$ under scenario ω

Decision variables

- y_{ς}: Binary variable set to 1 if ship-from point ς is deployed and 0 otherwise
- $z^o_{\varsigma}(\omega)$: Binary variable set to 1 if under scenario ω order $o \in \vec{\mathcal{O}}(\omega)$ is satisfied from supply point ς and 0 otherwise
- $x^p_{\varsigma t}(\omega)$: The amount of product p shipped from the central warehouse to ship-from point ς on day $t \in \tau_{\varsigma}$ under scenario ω
- $I^p_{\varsigma t}(\omega)$: Inventory level of product p at ship-from point ς at the beginning of period t
- $B^p_{st}(\omega)$: Backlog level of product p at store s at the beginning of period t

3.4. E-DND model

We consider two types of decisions in our model: those that are here-and-now and must be made before the realization of uncertainty, and those that are wait-and-see and are made after observing the uncertainty. The selection of the ship-from points to fulfill online orders is here-and-now, therefore deciding the deployment strategy of the distribution network ahead of the realization of uncertainty, whereas the decisions concerning the operations of this distribution network are made observing the demand and capacity realizations, and are therefore wait-and-see. These decisions include, the daily flow of inventory replenishment at ship-from points under each scenario, and the decision to allocate an order to

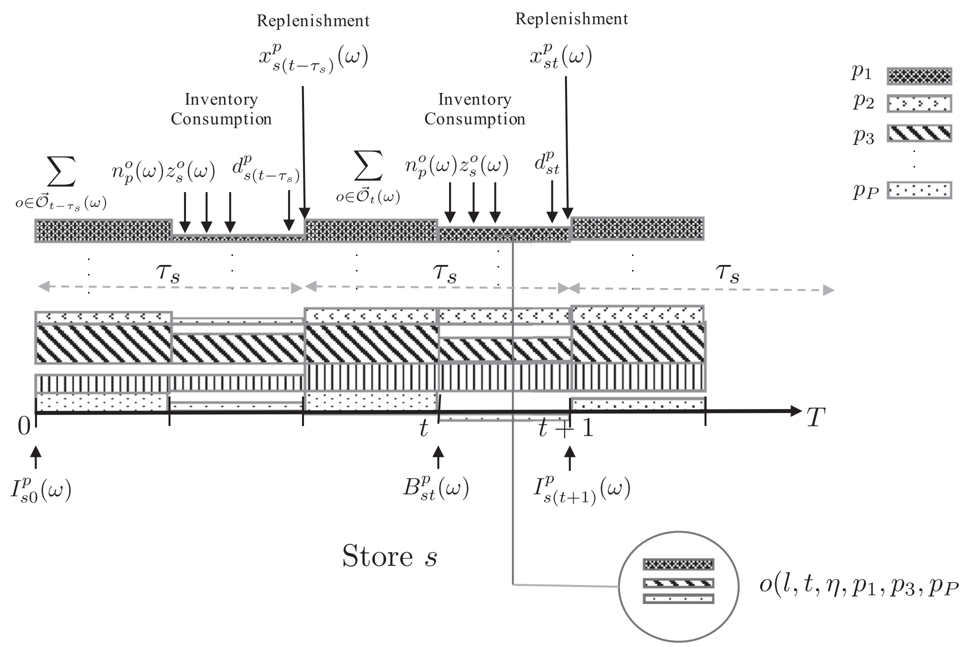


Fig. 2. Flows in and out of inventory at SS configured store s.

a supply point (be it a store, FC, or the warehouse) under each scenario. More specifically, the e-DND problem is modeled as a two-stage stochastic location-allocation model incorporating flow-based replenishment decisions and order fulfillment decisions. The first-stage decisions determine a deployment strategy by selecting either to keep fulfilling online orders from the warehouse, to operate a dedicated urban FC, and/or to assign the fulfillment of online orders to a subset of the stores located in the city. The second stage approximates the revenues and operational costs for the planning horizon given a set of online orders to be fulfilled and the simultaneous arrival of physical demand. Its inheritance of the combinatorial complexity of the facility location problem and the modeling of the OFP, which forces the inclusion of a very large number of binary variables, makes e-DND very difficult to solve.

We next describe in detail how inventory management is handled in our mathematical model. To adequately capture the inventory management system, we calculate the time-phased inventory requirements as in a distribution planning requirement (DRP) approach, which is widely used in practice and in the case of multi-echelon distribution networks, which can handle stochastic time-varying demand (Martel, 2003, Firoozi, Babai, Klibi, & Ducq (2019)). Fig. 2 provides an illustration of how the flows in (replenishment) and flows out (online order fulfillment and physical demand) are considered to compute the on-hand inventory and backlog levels for a given store. As depicted, in period t the business has to make the regular replenishment of store s and the quantity received by the end of the period will only be available to satisfy demand at the beginning of period $t+1$. The inventory level at the beginning of period t is given by $I_{s,t}(\omega)$. Then, during period t, the demand for online orders is received and an assignment of online orders to store s is undertaken based on the inventory level $I_{s,t}(\omega)$.

By the end of day t, once all the online orders assigned to store s are handled and the physical demand is revealed, the backlog level $B^p_{st}(\omega)$ is calculated for the physical demand of product p, which is reported to period $t+1$ using the inventory balance equations in order to be satisfied in upcoming days.

With the notation and the assumptions just introduced, the e-DND problem can be written as follows:

$$\max \quad -\sum_{\varsigma\in\mathcal{SF}} c^{\mathrm{f}}_{\varsigma} y_{\varsigma} + \mathbb{E}_{\Omega} Q(\boldsymbol{y},\omega) \tag{1}$$

$$\text{s.t} \quad \sum_{\varsigma\in\mathcal{SF}} y_{\varsigma} \le s^{\max} \tag{2}$$

$$\boldsymbol{y} \in \{0,1\}^{|\mathcal{SF}|} \tag{3}$$

where $Q(\boldsymbol{y},\omega)$ represents the optimal value of the order fulfillment problem for a selection of ship-from points $\boldsymbol{y}$ and an uncertainty realization $\omega\in\Omega$ and is given as

$$\begin{aligned} Q(\boldsymbol{y},\omega) &\\ = \max \Bigg\{ &\sum_{s\in\mathcal{S}}\sum_{p\in\mathcal{P}}\sum_{t\in\mathcal{T}} \pi^p (d^p_{st}(\omega) - B^p_{s(t+1)}(\omega)) \\ &- \sum_{s\in\mathcal{S}}\sum_{t\in\mathcal{T}}\sum_{p\in\mathcal{P}} c^{\mathrm{b}}_{ps} B^p_{st}(\omega) - \sum_{\varsigma\in\mathcal{SF}}\sum_{t\in\tau_{\varsigma}}\sum_{p\in\mathcal{P}} c^{\mathrm{r}}_{p\varsigma} x^p_{\varsigma t}(\omega) \\ &- \sum_{\varsigma\in\mathcal{SF}}\sum_{t\in\mathcal{T}}\sum_{p\in\mathcal{P}} c^{\mathrm{h}}_{p\varsigma} I^p_{\varsigma t}(\omega) \\ &+ \sum_{t\in\mathcal{T}}\sum_{l\in\mathcal{L}}\sum_{o\in\bar{\mathcal{O}}_{lt}(\omega)}\sum_{\varsigma\in\mathcal{SF}} (\pi^o(\omega) - c^{\mathrm{t}}_{\varsigma l} - c^{\mathrm{p}}_{\varsigma} n^o(\omega)) z^o_{\varsigma}(\omega) \\ &+ \sum_{t\in\mathcal{T}}\sum_{l\in\mathcal{L}}\sum_{o\in\bar{\mathcal{O}}_{lt}(\omega)} (\pi^o(\omega) - c_{\mathrm{W}o} - c^{\mathrm{t}}_{\mathrm{W}l} - c^p_{\mathrm{w}} n^o(\omega)) z^o_{\mathrm{w}}(\omega) \Bigg\} \end{aligned} \tag{4}$$

$$\bar{I}^p_{st}(\omega) = \bar{I}^p_{s(t+1)}(\omega) + d^p_{st}(\omega) + \sum_{o\in\vec{\mathcal{O}}_t(\omega)} n^o_p(\omega) z^o_s(\omega) \qquad \forall s\in\mathcal{S}, t\in\mathcal{T}\setminus\tau_s, p\in\mathcal{P} \tag{5}$$

$$\bar{I}^p_{st}(\omega) + x^p_{st}(\omega) = \bar{I}^p_{s(t+1)}(\omega) + d^p_{st}(\omega) + \sum_{o\in\vec{\mathcal{O}}_t(\omega)} n^o_p(\omega) z^o_s(\omega) \qquad \forall s\in\mathcal{S}, t\in\tau_s, p\in\mathcal{P} \tag{6}$$

$$I^p_{ft}(\omega) = I^p_{f(t+1)}(\omega) + \sum_{o\in\vec{\mathcal{O}}_t(\omega)} n^o_p(\omega) z^o_f(\omega) \quad \forall f\in\mathcal{F}, t\in\mathcal{T}\setminus\tau_f, p\in\mathcal{P} \tag{7}$$

$$I^p_{ft}(\omega) + x^p_{ft}(\omega) = I^p_{f(t+1)}(\omega) + \sum_{o\in\vec{\mathcal{O}}_t(\omega)} n^o_p(\omega) z^o_f(\omega) \quad \forall f\in\mathcal{F}, t\in\tau_f, p\in\mathcal{P} \tag{8}$$

$$\sum_{\varsigma\in\mathcal{SF}|t\in\tau_\varsigma}\sum_{p\in\mathcal{P}} x^p_{\varsigma t}(\omega) \le k_t \qquad \forall t\in\mathcal{T} \tag{9}$$

$$\sum_{\varsigma\in\mathcal{SP}|r_{\varsigma l}\le\eta^o(\omega)} z^o_\varsigma(\omega) \le 1 \qquad \forall o\in\vec{\mathcal{O}}(\omega) \tag{10}$$

$$\sum_{o\in\vec{\mathcal{O}}_t(\omega)} z^o_\varsigma(\omega) \le b_{\varsigma t}(\omega) y_\varsigma \qquad \forall\varsigma\in\mathcal{SF}, t\in\mathcal{T} \tag{11}$$

$$n^o_p(\omega) z^o_\varsigma(\omega) \le I^p_{\varsigma t}(\omega) - \sum_{o'\in\vec{\mathcal{O}}_t(\omega)\prec o} n^{o'}_p(\omega) z^{o'}_\varsigma(\omega) \quad \forall\varsigma\in\mathcal{SF}, t\in\mathcal{T}, o\in\vec{\mathcal{O}}_t(\omega), p\in\mathcal{P}(o) \tag{12}$$

$$\boldsymbol{z}\in\{0,1\}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{B} \ge 0 \tag{13}$$

where we have substituted $\bar{I}^p_{st} = I^p_{st} - B^p_{st}$ for $p\in\mathcal{P}$, $s\in\mathcal{S}$, and $t\in\mathcal{T}$ for ease of exposition.

The objective function (4) maximizes the total profit based on the expected cumulative revenues from physical and online sales minus the sum of the first-stage costs and the expected second-stage costs. In this function the term $\sum_{s\in\mathcal{S}}\sum_{p\in\mathcal{P}}\sum_{t\in\mathcal{T}}\pi^p(d^p_{st}(\omega) - B^p_{s(t+1)}(\omega))$ expresses the net revenue earned from physical sales, whereas the term $\sum_{s\in\mathcal{S}}\sum_{t\in\mathcal{T}}\sum_{p\in\mathcal{P}} c^{\mathrm{b}}_{ps} B^p_{st}(\omega)$ expresses the total backlogging cost stemming from unsatisfied physical demand. The terms $\sum_{\varsigma\in\mathcal{SF}}\sum_{t\in\tau_\varsigma}\sum_{p\in\mathcal{P}} c^{\mathrm{r}}_{p\varsigma} x^p_{\varsigma t}(\omega)$ and $\sum_{\varsigma\in\mathcal{SF}}\sum_{t\in\mathcal{T}}\sum_{p\in\mathcal{P}} c^{\mathrm{h}}_{p\varsigma} I^p_{\varsigma t}(\omega)$ express the replenishment and holding costs incurred at ship-from points, respectively. In addition to these costs which are expressed per product, for orders assigned to a ship-from point the term $\sum_{t\in\mathcal{T}}\sum_{l\in\mathcal{L}}\sum_{o\in\vec{\mathcal{O}}_{lt}(\omega)}\sum_{\varsigma\in\mathcal{SF}}(\pi^o(\omega) - c^{\mathrm{t}}_{\varsigma l} - c^{\mathrm{p}}_\varsigma n^o(\omega)) z^o_\varsigma(\omega)$ is added to the objective function. It accounts for the profit earned from the order minus the transportation and picking costs incurred. In the same manner, the term $\sum_{t\in\mathcal{T}}\sum_{l\in\mathcal{L}}\sum_{o\in\vec{\mathcal{O}}_{lt}(\omega)}(\pi^o(\omega) - c_{\mathrm{Wo}} - c^{\mathrm{t}}_{\mathrm{Wl}} - c^p_{\mathrm{W}} n^o(\omega)) z^o_{\mathrm{W}}(\omega)$ accounts for online orders assigned to the warehouse. This term includes the per-order transportation and picking costs, as well as the cost c_{Wo}, which is a per-order charge comprising of the fixed, holding, and replenishment costs at the warehouse.

Constraint (2) imposes the maximum number of ship-from points that can be selected. Constraints (5)-(6) are the inventory balance constraints for store $s\in\mathcal{S}$ and each product $p\in\mathcal{P}$ in scenario ω, where (5) concerns the regular periods $(\tau\setminus\tau_s)$ and (6) concerns the replenishment periods (τ_s) for $s\in\mathcal{S}$. Similarly, constraints (6)-(7) are the inventory balance constraints for FC $f\in\mathcal{F}$ and each product $p\in\mathcal{P}$ in scenario ω, where (7) concerns the regular periods $(\tau\setminus\tau_f)$ and (8) concerns the replenishment periods (τ_f) for $f\in\mathcal{F}$. Constraints (9) ensure that the transportation capacity in each period is respected. Constraints (10) ensure that each order is satisfied at most once by one supply point, where the requested response time for the order is imposed using the set $\{\varsigma\in\mathcal{SP} \mid r_{\varsigma l}\le\eta^o(\omega)\}$. Constraints (11) ensure that the capacity of stores and FCs to prepare online orders is respected. We remark that, as we have assumed the FC capacities to be deterministic $b_{ft}(\omega) = b_{ft}$ for $\omega\in\Omega$ in (11). Finally, constraints (12) impose, for each time period t and each scenario ω, that an order o can be met from ship-from point $\varsigma\in\mathcal{SF}$ only if all products $p\in\mathcal{P}(o)$ are in stock after all preceding orders $o'\in\vec{\mathcal{O}}_t(\omega)\prec o$ assigned to this ship-from point are satisfied.

Remark 3.1. The deployment strategies we consider are not mutually exclusive. In other words, under the SS strategy, it is possible to satisfy some orders from the warehouse. Similarly, under the SF strategy, it is possible to satisfy some orders from the selected stores and the warehouse. Pure strategies can be explored using our model by controlling the set $\mathcal{SF}$.

Remark 3.2. The assignment of online orders to stores in constraints (12) is based on the inventory at the beginning of period I^p_{st}. This implicitly imposes that the online orders arriving during the day will be prioritized over the physical demand and therefore is an approximation based on a practical policy.

Remark 3.3. Although constraints (12) are successful in imposing the product composition of an order, their reflection of the first-in/first-out fulfillment policy depends on the desired matching between the applied decision period (t) and the order response time. In other words, as we consider next-day deliveries, we assume that orders accumulated within the day no longer reflect a first-in/first-out fulfillment policy and that a higher profit order will be prioritized over another order that requests the same products. We note however that if same-day deliveries need to be considered, the time granularity of the applied decision period (t) must be refined in accordance with the order arrival periodicity in order to take the first-in/first-out policy into account.

We next write the deterministic equivalent of e-DND as follows:

$$\begin{aligned} \max \quad & -\sum_{\varsigma\in\mathcal{SF}} c^{\mathrm{f}}_\varsigma y_\varsigma + \sum_{\omega\in\Omega} p(\omega) Q(\boldsymbol{y},\omega) \\ \text{s.t} \quad & (2)-(3) \\ & (5)-(13) \qquad \forall\omega\in\Omega \end{aligned}$$

4. Solution approach

To solve the e-DND to optimality, a first approach is to rely on the deterministic equivalent model along with a sampling method such as the sample average approximation (SAA). This sampling method provides the adequate sample size N to formulate and obtain good quality solutions by solving directly the e-DND model presented in the previous section. However, as N grows, the deterministic equivalent model becomes computationally intractable due to the large number of binary variables and constraints present in the model for each sample. We therefore discuss in this section the integer L-shaped method used to solve e-DND exactly.

Let Ω_N be a scenario sample of size N, then the expectation $\mathbb{E}_{\Omega_N} Q(\boldsymbol{y},\omega)$ takes the form $\frac{1}{N}\sum_{\omega\in\Omega_N} Q(\boldsymbol{y},\omega)$ by the sample average approximation (SAA) method. In the remainder, we let $\mathcal{Q}(\boldsymbol{y}) = \mathbb{E}_{\Omega_N} Q(\boldsymbol{y},\omega) = \frac{1}{N}\sum_{\omega\in\Omega_N} Q(\boldsymbol{y},\omega)$, and we write

$$\max \quad -\sum_{\varsigma\in\mathcal{SF}} c^{\mathrm{f}}_\varsigma y_\varsigma + \frac{1}{N}\sum_{\omega\in\Omega_N} Q(\boldsymbol{y},\omega) \tag{14}$$

$$\text{s.t} \quad \sum_{\varsigma\in\mathcal{SF}} y_\varsigma \le s^{\max} \tag{15}$$

$$\boldsymbol{y}\in\{0,1\}^{|\mathcal{SF}|} \tag{16}$$

When $Q(\boldsymbol{y},\,\omega)$ is a linear program, the formulation (14)–(16) lends itself to what is known in the literature as Benders decomposition (Benders, 1962) (or L-shaped method; Van Slyke & Wets, 1969), where the cost $\mathcal{Q}(\boldsymbol{y})$ of a first-stage solution $\boldsymbol{y}$ is defined by *optimality cuts* and infeasible first-stage decisions $\boldsymbol{y}$ are cut off by *feasibility cuts* defined through the linear programming dual of $Q(\boldsymbol{y}, \omega)$.

However, when $Q(\boldsymbol{y}, \omega)$ is an integer program as is the case for e-DND, the optimality cuts obtained through the linear programming relaxation no longer suffice to define an optimal solution to (4)–(13). However, when $\boldsymbol{y}$ are binary we may still impose a description of $\mathcal{Q}(\boldsymbol{y})$ through specific optimality cuts. To describe these cuts let $\mathcal{Y} = \left\{\boldsymbol{y} \in \{0,1\}^{|\mathcal{SF}|} \mid \sum_{\varsigma \in \mathcal{SF}} y_\varsigma \le s^{\max}\right\}$. Assume first that a finite upper bound $U \ge \max_{y \in \mathcal{Y}} \mathcal{Q}(\boldsymbol{y})$ exists. Assume further that $\bar{\mathcal{S}} \subseteq \mathcal{SF}$, and let $\boldsymbol{y}(\bar{\mathcal{S}})$ be a first-stage solution such that $y_i(\bar{\mathcal{S}}) = 1$ if $i \in \bar{\mathcal{S}}$ and $y_i(\bar{\mathcal{S}}) = 0$ otherwise. Further, let $q_{\bar{\mathcal{S}}}$ be the value of $\mathcal{Q}(\boldsymbol{y}(\bar{\mathcal{S}}))$. Then we may write

$$\max \quad -\sum_{\varsigma \in \mathcal{SF}} c^{\mathrm{f}}_\varsigma y_\varsigma + \theta \tag{17}$$

$$\text{s.t} \quad \sum_{\varsigma \in \mathcal{SF}} y_\varsigma \le s^{\max} \tag{18}$$

$$\theta \le (q_{\bar{\mathcal{S}}} - U)\left(\sum_{i \in \bar{\mathcal{S}}} y_i - \sum_{i \notin \bar{\mathcal{S}}} y_i\right) - (q_{\bar{\mathcal{S}}} - U)(|\bar{\mathcal{S}}| - 1) + U \quad \forall \bar{\mathcal{S}} \subseteq \mathcal{SF} \tag{19}$$

$$\boldsymbol{y} \in \{0,1\}^{|\mathcal{SF}|}. \tag{20}$$

Cuts (19) form an exponential family and define the value of θ for each $\boldsymbol{y} \in \mathcal{Y}$ as $\theta \le q_{\bar{\mathcal{S}}}$ when $\boldsymbol{y} = \boldsymbol{y}(\bar{\mathcal{S}})$ and $\theta \le U$ otherwise. Clearly, constructing (17)–(20) in its compact form requires enumerating all feasible $\boldsymbol{y} \in \mathcal{Y}$ and evaluating their optimal recourse values $Q(\boldsymbol{y}, \omega)$ for $\omega \in \Omega_N$. However, a decomposition scheme where cuts (19) are first relaxed and then added to the relaxation progressively by evaluating the current candidate solution $(\boldsymbol{y}^*, \theta^*)$ as integer feasible solutions are found can be developed. This method, known as the integer L-shaped method was first proposed by Laporte and Louveaux (1993). We remark at this point that $Q(\boldsymbol{y}, \omega)$ for e-DND is always feasible, and therefore feasibility cuts are not needed for our discussion.

We next address the question of computing the bound U. We remark that U is an upper bound on the maximum profit that can be achieved by any $\boldsymbol{y} \in \mathcal{Y}$. In e-DND, the first-stage solution that can achieve the most profit considers all ship-from points to be available to fulfill online orders. Let $\boldsymbol{e}$ be an $|\mathcal{SF}|-$dimensional vector of 1s. Then U can be calculated as $\mathcal{Q}(\boldsymbol{e})$ where $Q(\boldsymbol{e}, \omega)$ can be computed individually for each scenario $\omega \in \Omega_N$.

It is clear that the cuts (19) are expensive to generate at each iteration because they require the exact evaluation of the value function $\mathcal{Q}(\boldsymbol{y}^*)$ for each candidate solution $(\boldsymbol{y}^*, \theta^*)$. Further, these cuts define only the value of the solution $\boldsymbol{y}^*$ and do not bring any information about different solutions (other than the fact that their value is bounded by U). To overcome these drawbacks Angulo et al. (2016) suggested cultivating the information provided by the LP-relaxation in the form of subgradient cuts whenever possible. The subgradient cuts are in fact Benders' cuts written in the form

$$\theta \le \nabla\mathcal{Q}(\boldsymbol{y}^*)(\boldsymbol{y} - \boldsymbol{y}^*) + \mathcal{Q}_{\mathrm{LP}}(\boldsymbol{y}^*)$$

where $\nabla\mathcal{Q}(\boldsymbol{y}^*)$ is a subgradient of $\mathcal{Q}(\boldsymbol{y})$ at $\boldsymbol{y}^*$ and $\mathcal{Q}_{\mathrm{LP}}(\boldsymbol{y}^*)$ is its LP-relaxation value. In the case of e-DND the subgradient $\nabla\mathcal{Q}(\boldsymbol{y}^*)$ corresponds to the optimal dual solution associated with the constraints (11) multiplied by $\boldsymbol{b}(\omega)$, the capacity realization of scenario $\omega \in \Omega_N$. Let $\boldsymbol{\rho}^*(\boldsymbol{y}^*, \omega)$ be the optimal dual solution associated with the constraints (11) and $\mathcal{Q}_{\mathrm{LP}}(\boldsymbol{y}^*, \omega)$ be the optimal value of the LP-relaxation of (4)-(13) at $\boldsymbol{y} = \boldsymbol{y}^*$. Then the subgradient cut associated with the solution $\boldsymbol{y}^*$ can be written as

$$\theta \le \frac{1}{N}\sum_{\omega \in \Omega_N} \boldsymbol{\rho}^*(\boldsymbol{y}^*, \omega)\boldsymbol{b}(\omega)(\boldsymbol{y} - \boldsymbol{y}^*) + \frac{1}{N}\sum_{\omega \in \Omega_N} \mathcal{Q}_{\mathrm{LP}}(\boldsymbol{y}^*, \omega). \tag{21}$$

We implement the integer L-shaped algorithm by embedding it in the structure of the branch-and-bound tree using the callback functionalities of a commercial solver. This type of implementation requires a cut generation function to be called whenever a candidate solution $(\boldsymbol{y}^*, \theta^*)$ is found. We next describe the implementation of this function with an alternating cut strategy (see Angulo et al., 2016), which is outlined in Algorithm 1. This implementation requires two lists, V and V_{LP}, the list of solutions for which $\mathcal{Q}(\boldsymbol{y})$ is known and, the list of solutions for which $\mathcal{Q}_{\mathrm{LP}}(\boldsymbol{y})$ is known, respectively. A solution $(\boldsymbol{y}^*, \theta^*)$ is explored first to see if a subgradient cut in the form of (21) can be added ($\boldsymbol{y}^*$ added to the list V_{LP}). If this is the case, the algorithm returns adding the subgradient cut to the master problem (17)-(20). If the subgradient cut is not violated (or if solution $\boldsymbol{y}^*$ is already in the list V_{LP}) then the algorithm explores the solution $\boldsymbol{y}^*$ to see if an optimality cut in the form of (19) can be added ($\boldsymbol{y}^*$ transferred from the list V_{LP} to the list V). If this is the case, the algorithm proceeds by adding the optimality cut to the master problem (17)–(20). Otherwise, the candidate solution $(\boldsymbol{y}^*, \theta^*)$ is accepted as an incumbent solution.

Algorithm 1: Optimality cut function with alternating cut strategy.

Input: $(\boldsymbol{y}^*, \theta^*)$, V, V_{LP}
if $\boldsymbol{y}^* \in V$ **then**
    Accept $(\boldsymbol{y}^*, \theta^*)$ as an incumbent solution
    return
end
if $\boldsymbol{y}^* \notin V_{\mathrm{LP}}$ **then**
    Compute $\mathcal{Q}_{\mathrm{LP}}(\boldsymbol{y}^*)$ by solving the LP-relaxation of (4)-(13) with $\boldsymbol{y} = \boldsymbol{y}^*$
    $V_{\mathrm{LP}} \leftarrow V_{\mathrm{LP}} \cup \{\boldsymbol{y}^*\}$
    if $\theta > \mathcal{Q}_{\mathrm{LP}}(\boldsymbol{y}^*)$ **then**
        Add the subgradient cut (21) to the master problem (17)-(20)
        return
    end
end
Compute $\mathcal{Q}(\boldsymbol{y}^*)$ by solving (4)-(13) with $\boldsymbol{y} = \boldsymbol{y}^*$
$V \leftarrow V \cup \{\boldsymbol{y}^*\}$
if $\theta < \mathcal{Q}(\boldsymbol{y}^*)$ **then**
    Add the integer optimality cut (19) to the master problem (17)-(20)
    return
else
    Accept $(\boldsymbol{y}^*, \theta^*)$ as an incumbent solution
end

We additionally test the improved optimality cuts in our implementation of the integer L-shaped algorithm (see Birge & Louveaux, 2011). To express these cuts, define the set $N(k, \bar{\mathcal{S}})$ of k-neighbors of $\bar{\mathcal{S}}$ as the set of solutions $\left\{\boldsymbol{y} \in \mathcal{Y} \mid \delta(\boldsymbol{y}, \bar{\mathcal{S}}) = |\bar{\mathcal{S}}| - k\right\}$ where $\delta(\boldsymbol{y}, \bar{\mathcal{S}}) = \sum_{i \in \bar{\mathcal{S}}} y_i - \sum_{i \notin \bar{\mathcal{S}}} y_i$. Let $\lambda(k, \bar{\mathcal{S}}) \ge \max_{\boldsymbol{y} \in N(k, \bar{\mathcal{S}})} \mathcal{Q}(\boldsymbol{y})$ for $k = 0, \ldots, |\bar{\mathcal{S}}|$ with $\lambda(0, \bar{\mathcal{S}}) = q_{\bar{\mathcal{S}}}$. Let $\boldsymbol{y}(\bar{\mathcal{S}})$ be a first-stage solution with value $q_{\bar{\mathcal{S}}}$ such that $y_i(\bar{\mathcal{S}}) = 1$ if $i \in \bar{\mathcal{S}}$ and $y_i(\bar{\mathcal{S}}) = 0$ otherwise. Define $a = \min\{q_{\bar{\mathcal{S}}} - \lambda(1, \bar{\mathcal{S}}), (q_{\bar{\mathcal{S}}} - U)/2\}$. Then the inequalities

$$\theta \le a\left(\sum_{i \in \bar{\mathcal{S}}} y_i - \sum_{i \notin \bar{\mathcal{S}}} y_i\right) + q_{\bar{\mathcal{S}}} - a|\bar{\mathcal{S}}| \qquad \forall \bar{\mathcal{S}} \subseteq \mathcal{SF} \tag{22}$$

Please cite this article as: A.N. Arslan, W. Klibi and B. Montreuil, Distribution network deployment for omnichannel retailing, European Journal of Operational Research, https://doi.org/10.1016/j.ejor.2020.04.016

Table 1
Instance sizes.

	City1	*City2*
$\|\mathcal{S}\|$	5	9
$\|\mathcal{L}\|$	10	20
$s^{\max}$	3	5

are valid for (17)-(20) and improve the bound on all 1-neighbors of a solution from U to $\lambda(1, \bar{\mathcal{S}})$.

To conclude this section, we discuss how we calculate $\lambda(1, \bar{\mathcal{S}})$ for our computations in Section 5. We first note that the 1-neighbors of $\boldsymbol{y}(\bar{\mathcal{S}})$ are those solutions obtained by either closing a ship-from point that is open to online order fulfillment in $\boldsymbol{y}(\bar{\mathcal{S}})$ or by opening a ship-from point that is closed in $\boldsymbol{y}(\bar{\mathcal{S}})$ to online order fulfillment. Keeping in mind that $\mathcal{Q}(\boldsymbol{y}(\bar{\mathcal{S}}))$ is known, we remark that we do not need to consider those solutions that close an open ship-from point because this cannot increase the second-stage profit. Therefore, we may compute $\lambda(1, \bar{\mathcal{S}})$ by solving the deterministic equivalent problem after fixing $y_i = 1$ for $i \in \bar{\mathcal{S}}$ and adding the constraints $\sum_{i \notin \bar{\mathcal{S}}} y_i = 1$. However, this calculation is too expensive to perform in each iteration. In fact, obtaining this bound is of the same theoretical difficulty as solving e-DND. To overcome this difficulty, we propose solving instead a perfect-information relaxation by allowing $\boldsymbol{y}$ to take a different value $\boldsymbol{y}^{\omega}$ for $\omega \in \Omega_N$. Therefore, we add $\boldsymbol{y}^{\omega}$ to each scenario subproblem (4)-(13) as decision variables along with the constraints $y_i^{\omega} = 1$ for $i \in \bar{\mathcal{S}}$ and $\sum_{i \notin \bar{\mathcal{S}}} y_i^{\omega} = 1$. The resulting bound-calculating integer programs can now be solved independently for each scenario $\omega \in \Omega_N$. For the results presented in Section 5, we compute $\lambda(1, \bar{\mathcal{S}})$ by solving the continuous recourse relaxations of these bound-calculating integer programs for $\omega \in \Omega_N$ and add improved optimality cuts when $1 \leq |\bar{\mathcal{S}}| < s^{\max}$.

5. Computational experiments

In this section, we present the numerical experiments conducted in order to validate the model and the solution approach proposed in the previous sections and to analyze the performance of the introduced deployment strategies. We first provide the details of the set of problem instances generated, which are inspired by the real data of a European retailer. Most of the parameters used to generate these instances are primary data extracted from the ERP system of the company. Parameters that were not available are estimated relying on company experts.

5.1. Instance generation

To create instances of various sizes, we chose two cities in Europe that represent various scales in terms of area/density, number of physical stores, and number of customer locations. We will refer to the associated instances by *City1* and *City2*, selected such that the latter is a larger city than the former. Table 1 specifies the number of stores, customer locations, and the parameter $s^{\max}$ considered for each problem size. In each city an FC was identified and considered in accordance with the SF strategy. The number of products P in each instance was kept to 100, based on a joint analysis of ABC classification of the product portfolio by volume in terms of online sales and in store sales. The number of periods T was fixed to 30 consecutive days to represent a typical month of business.

Table 2 summarizes the parameters used to generate instances within each category, where percentages are to be taken on the sales price of each product. For each instance, the real replenishment schedule of each store was applied such that stores are visited once, twice, or three times per week. Stores that are replenished three times per week were categorized as having high physical customer traffic, and their fixed cost, physical demand, and capacity to serve online orders were adjusted accordingly. We remark that an initial inventory of 5 per product in each store and 10 per product for the FC was considered in order to reflect the business-as-usual situation at the beginning of the horizon. Additionally, we assumed that the FCs were replenished at each time period. For all of our computational tests the transportation capacity was estimated based on the average demand per period such that the available capacity is 1.2 times higher than average demand per period. In this context, the location of the warehouse allows a two-day response time as a result of its distance from each of the two cities selected. Conversely, the potential set of stores and urban FCs are selected within each city such that next-day delivery is feasible.

Moreover, to generate plausible scenarios, the following data were gathered as input to the Monte Carlo process described in Section A.2 of the Appendix A. For each store the expected daily order processing capacity b_s was generated randomly as $U_d[5, 7]$ for regular stores and $U_d[10, 14]$ for high physical customer traffic stores. The deterministic capacity for the FC, on the other hand, was set to 30 orders per day. The online order arrival rate for each customer location was generated randomly as $0.165{*}U_d[1, 10]$. The inclusion probability of each product was generated randomly as $U_c[0, 1]$ and then normalized by the sum of probabilities of all products. Regarding the number of items in each order, $k_{\min}$ was set to 1, $k_{\max}$ was set to 6, and the mean and the standard deviation of the normal distribution for the quantity requested was set to 1 and $U_c[0.7, 1.2]$ for all products, respectively. To generate the physical demand, the expected demand for each product p, was generated randomly as $\lambda_p = 1.3 * U_d[1, 6]$. This parameter was directly used for regular stores, and multiplied by 1.2 for high physical customer traffic stores. To generate the requested response time for each arriving customer three different distributions that characterize this parameter based on the response time expectation were used. These distributions are given in Table 3 for low, moderate and high response time expectation. Finally, to generate the market conditions (see Section A.1 of the Appendix A), the parameters were chosen as $m_{\min} = 0$, $m_{\max} = 1$, $\underline{s} = 0.1$, $\delta_{\bar{m}} = 0.8$, $\delta_s = 0.9$ and $\beta = 0.8$. The process was started at time $t = 0$ with initial parameters $a = 0$, $b = 1$, steady state mean $\bar{m} = 0.5$, and steady state standard deviation $\bar{s} = 0.15$.

Finally, another key dimension we considered in our experiments is the variability of the online-to-physical demand ratio (DR). To this end, we considered in our experiments three cases: low DR, moderate DR, and high DR, obtained by multiplying the online order arrival rate for each customer location by 1, 4/3, and 2, respectively. Combining all these elements yielded 18 problem instances. Each instance type is a combination of the problem size (City), the response time expectation level (RE), and the demand ratio (DR). We recall that each of the 18 instances is solved three times, once for each deployment strategy, namely, ship-from warehouse (SW), ship-from stores (SS), and ship-from fulfillment platforms (SF).

5.2. Sample size and solution approach validation

We first establish the number of sample scenarios N required to obtain a good solution with the sample average approximation method (see Shapiro, Dentcheva, & Ruszczynski, 2009). To determine the best value of N, different sample sizes were tested and the quality of the solutions obtained was evaluated using a statistical optimality gap. For our tests we used the *City1* instance and generated 25 independent samples of 50 scenarios each. After fixing the total number of scenarios to $K \in \{10, 25, 50\}$, for each sample we formed an SAA problem with K scenarios. We solved these

Table 2
Instance generation parameters.

Warehouse		Store	
Transportation cost	5$/order	Fixed cost	1000*U_d[3, 6] (1000*U_d[7, 10] for High)
Replenishment charge	15%	Picking cost	0.25 c_w^p
Fixed charge	40%	Holding cost	1.5%
Holding charge	1%	Backlogging cost	50%
Picking cost	1$/item	Replenish cost	20%
		Transportation cost	U_c[3, 5]$/order
Product		**Fulfillment Center**	
Price	U_c[15, 50]	Usage cost	20,000
		Picking cost	0.5 c_w^p
		Holding cost	1.2%
		Replenish cost	20%
		Transportation cost	3.5$/order

Table 3
Response time expectation distributions considered.

	Low RE	Moderate RE	High RE
(1D, 24h)	0.1	0.25	0.6
(2D, 48h)	0.2	0.3	0.3
(3D, 72h)	0.7	0.45	0.1

Table 4
Statistical optimality gap with $K \in 10, 25, 50$.

	$K = 10$	$K = 25$	$K = 50$
statgap	9.60	5.32	2.97

SAA problems to optimality (with a 0.1% optimality tolerance) and obtained 25 values. We then used these values to construct a 95% confidence interval on the lower bound. We also recorded the optimal solutions of these SAA problems. To compute similarly a confidence interval on the upper bound we generated a sample of 500 scenarios independently. We evaluated each of the 25 solutions obtained from the lower bound process with these 500 scenarios and recorded the optimal value. We then used these values to construct a 95% confidence interval on the upper bound. Finally, an upper bound for the optimality gap was calculated as $100 \times (UB - LB)/LB$, where UB is the upper limit of the upper bound confidence interval and LB is the lower limit of the lower bound confidence interval. We present the statistical gaps obtained in Table 4. Based on these results the gap between 25 and 50 scenarios is still significant. As a result, we decided to conduct our computational tests with a sample size of 50. We note here that since the planning horizon is composed of 30 periods, about $K \times 30$ realizations are sampled from the same probability distributions when K scenarios are considered, which explains the relatively low statistical gaps with a reasonable number of scenarios.

Second, we conduct experiments to validate the tractability of the modeling approach and the efficiency of the solution approach. Here, the aim is to validate the efficiency of the integer L-shaped method (LS) proposed to solve the two-stage stochastic model. To this end, we compare it to the deterministic equivalent model solved with CPLEX (DE) and to the continuous recourse relaxation of the deterministic equivalent model solved with automatic Benders decomposition functionalities of CPLEX (R-BD). We remark that the model (17)–(20) solved with integer L-shaped method (LS) is implemented using the lazy-cut-callback functionalities of CPLEX. It is clear that the continuous recourse relaxation model is expected to be much more tractable compared to the other two. However, it does not lead to the same first-stage strategic decisions because it does not evaluate the second-stage profit correctly. We include it in our computational experiments as a reference point for the difficulty of our instances.

We solve all models using CPLEX, imposing 0.1% percent optimality gap and a one-hour time limit. We compute the value U by solving each scenario subproblem to 1% optimality after setting all first-stage decisions to 1. All scenario subproblems for the L-shaped method are solved to 0.1% percent optimality.

We first evaluate the efficiency of the three solution methods on six instances (2 problems sizes × 3 response time expectations) with number of scenarios $N = 10$. The results are presented in Table 5.

We first remark that, as expected, the continuous recourse relaxation solved using the automatic Benders decomposition functionalities of CPLEX was fairly more tractable than the integer recourse models. When comparing the exact solution methods, we see that the L-shaped model was an order of magnitude faster in the case of the *City 1* instance and that it provided the optimal solution where the deterministic equivalent model had gaps exceeding 10% in the case of the *City 2* instance. The difference between the performances of the two models can be explained by the different ways they handle the scenario subproblems. The integer L-shaped method creates one subproblem per scenario, whereas the deterministic equivalent model optimizes all scenario subproblems simultaneously.

In light of these results, as the instance size and the number of orders grow, the integer L-shaped method presents itself as an efficient exact solution method. We have observed that with increasing scenario size N the deterministic equivalent model is less and less capable of finding and improving feasible solutions within a reasonable amount of time, whereas the integer L-shaped method can be depended on to find good feasible solutions. Clearly, further research is required to tailor this method to reduce the solution time and improve the convergence for even larger instances.

5.3. Analysis of results

In this section, we evaluate the results for the entire set of instances in order to analyze and derive insights from the deployment strategies proposed: ship-from warehouse (SW), ship-from stores (SS), and ship-from fulfillment platforms (SF).

Network structure and fulfillment. First, for the instances related to *City1*, the analysis of the network structure produced is presented in Table 6, where we report the number of stores selected (with the number of high physical customer traffic stores in parentheses), the number of urban FCs deployed, and the percentage of online orders satisfied by the warehouse, the stores, and the urban FCs, respectively.

A first observation based on these results is that the introduction of the ship-from stores (SS) strategy immediately leads

Table 5
Solution performances with different methods.

		Low RE			Moderate RE			High RE		
		DE	R-BD	LS	DE	R-BD	LS	DE	R-BD	LS
City 1	Time (sec)	2874	206	333	2648	261	391	3214	257	383
	Gap (%)	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
City 2	Time (sec)	3690	2538	3435	3662	2133	4018	3665	3670	4086
	Gap (%)	8.06	0.09	0.08	10.31	0.07	0.00	13.71	1.75	0.01

Table 6
Network structure analysis for all RE and DR attributes with the *City 1* instance.

		Low RE			Moderate RE			High RE		
		WS	SS	FS	WS	SS	SF	WS	SS	SF
Low DR	#Stores	0	2(1)	2(1)	0	2(1)	2(1)	0	2(1)	2(1)
	#FCs	0	0	0	0	0	0	0	0	0
	%Store	0.0	61.4	61.4	0.0	61.6	61.7	0.0	61.8	61.8
	%FC	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	%Warehouse	67.4	30.5	30.5	43.3	23.1	23.1	9.4	6.0	6.0
Moderate DR	#Stores	0	2(1)	2(1)	0	2(1)	2(1)	0	2(1)	2(1)
	#FCs	0	0	0	0	0	0	0	0	0
	%Store	0.0	49.7	49.7	0.0	49.8	49.8	0.0	49.8	49.8
	%FC	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	%Warehouse	66.8	39.8	39.8	42.7	29.2	29.2	9.4	7.4	7.4
High DR	#Stores	0	2(1)	2(1)	0	3(1)	1(0)	0	3(2)	1(0)
	#FCs	0	0	0	0	0	1	0	0	1
	% Store	0.0	34.6	34.6	0.0	47.1	17.3	0.0	58.6	16.5
	% FC	0.0	0.0	0.0	0.0	0.0	60.1	0.0	0.0	60.8
	% Warehouse	66.9	50.8	50.8	43.0	31.7	12.7	9.6	6.5	3.8

to a shift in the fulfillment of online orders. Indeed, a significant percentage of online orders in the nine instances reported in Table 6 are now satisfied from stores instead of the warehouse when the SS strategy is considered. In the case of low RE, two stores are selected to serve online orders under the SS strategy, where one is a high physical customer traffic store. The percentage of online orders satisfied from these two stores exceeds 34% in all DR levels, and even reaches 61% with low DR. This percentage drops to 49% and 34%, in the moderate DR and high DR cases, respectively. The difference is explained by the available capacity of selected stores, as in the low DR case this capacity corresponds to a higher percentage of the number of online orders. This highlights the importance of store capacities for ship-from operations and argues for the importance of considering a capacitated location-allocation modeling approach in such business contexts. The same behavior is observed with online orders under moderate and high RE instances. Further, we notice that in all instances with a low RE attribute, the two stores selected to serve online orders under the SS strategy, prove to be an optimal network structure under the SF strategy, which means that urban FCs are not deployed. Similar results are observed when looking at the network structures for moderate and high RE, with low and moderate DR, respectively. On the other hand, in case of high DR attribute, under the SF strategy, we observe the deployment of an urban FC and a major change in the number of stores selected. For instance in the high RE-high DR instance, the network structure shifts from three stores with the SS strategy to one store and one urban FC with the SF strategy. Additionally, with the deployment of an urban FC and therefore availability of significant deterministic capacity, the warehouse percentage drops by 70%. Clearly, the percentage of online orders satisfied by the warehouse also drops as the response time expectation increases.

A second observation based on these results, is the fact that the SF strategy is used and is profitable when there is increased demand for online orders requiring a shorter response time. The main insight here is that the customer shift to next-day delivery becomes critical when the warehouse-city proximity does not offer such response capability, which pushes for the opening of a second tier of distribution. Although there may be many operational advantages to the SF strategy, it requires a significant increase in fixed and usage costs (see Table 8). In fact, in most cases the optimal solution is to fulfil orders from stores even under the SF strategy. The shift to urban FCs occurs only when both the online demand and the response time expectation is elevated. This shows that the current usage cost for the urban FCs are still important compared to the corresponding increase in profitability. Therefore, the deployment of FCs should be supported by a decision model that considers the operational characteristics and evaluates the uncertainty in future order arrivals.

The analysis of the network structure produced for instances related to *City2* is presented in Table 7. Similar to the *City1* instances, we remark that the urban FC is deployed only for elevated RE and DR values. However, contrary to the *City1* instances the shift starts earlier with the high RE and moderate DR case, because the number of orders in the *City2* instance and therefore their profitability is higher. Another related observation is the role of the warehouse in online order fulfillment, especially in the low RE and low DR case. We remark that 50% of online orders are shipped directly from the warehouse under the SS/SF strategies contrasting with the 30% observed for the *City1* instance. This shows that as the number of online orders are increased, assuming there will still be some customers requesting two-day and three-day deliveries, there will be a shift away from pure strategies and towards a network of fulfillment points to cultivate as much capacity as possible at the lowest cost. This further highlights the need for studying such networks from the point of view of omnichannel integration.

We also remark that between *City1* and *City2* instances, increasing city size and increasing demand led to an increased number of selected stores. However, urban FC deployment was reserved for extreme cases and sufficiently supplemented by one or two stores. This indicates that in most cases a network of FCs will not be necessary, as a single urban FC is sufficient.

Table 7
Network structure analysis for all RE and DR attributes with the *City 2* instance.

		Low RE			Moderate RE			High RE		
		WS	SS	FS	WS	SS	SF	WS	SS	SF
Low DR	#Stores	0	2 (0)	2 (0)	0	3 (0)	3 (0)	0	3 (0)	3 (0)
	#FCs	0	0	0	0	0	0	0	0	0
	%Store	0.0	37.2	37.2	0.0	57.4	57.4	0.0	57.4	57.4
	%FC	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	%Warehouse	66.9	49.3	49.3	42.8	25.5	25.5	9.4	6.7	6.7
Moderate DR	#Stores	0	3 (0)	3 (0)	0	4 (0)	4 (0)	0	4 (0)	1 (0)
	#FCs	0	0	0	0	0	0	0	0	1
	%Store	0.0	44.7	44.7	0.0	55.1	55.1	0.0	59.3	11.0
	%FC	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	67.2
	%Warehouse	67.1	43.8	43.8	43.1	27.3	27.3	9.7	6.8	3.8
High DR	#Stores	0	4 (0)	2 (0)	0	5 (0)	2 (0)	0	5 (0)	3(0)
	#FCs	0	0	1	0	0	1	0	0	1
	% Store	0.0	40.7	15.4	0.0	47.9	18.7	0.0	47.9	25.8
	% FC	0.0	0.0	44.9	0.0	0.0	44.9	0.0	0.0	44.8
	% Warehouse	67.0	46.7	30.2	43.2	31.1	21.2	9.7	8.1	5.1

Table 8
Cost analysis for all RE and DR attributes with the *City 1* instance.

		Low RE		Moderate RE		High RE	
		SS	FS	SS	SF	SS	SF
Low DR	Revenue (%)	39.4	39.5	109.1	109.1	741.6	741.5
	Fixed	11,000	11,000	11,000	11,000	11,000	11,000
	Replenishment (%)	40.6	40.6	40.4	40.4	41.5	41.5
	% Lost (%)	−75.0	−75.2	−73.1	−73.2	−64.45	−64.4
Moderate DR	Revenue (%)	39.0	39.0	104.4	104.4	679.2	679.0
	Fixed	11,000	11,000	11,000	11,000	11,000	11,000
	Replenishment (%)	46.6	46.6	46.5	46.5	48.3	48.2
	% Lost (%)	−68.2	−68.3	−63.4	−63.4	−52.8	−52.7
High DR	Revenue (%)	35.6	35.6	102.5	116.4	722.3	824.9
	Fixed	11,000	11,000	16,000	23,000	21,000	23,000
	Replenishment (%)	56.5	56.5	67.1	90.4	82.2	90.9
	% Lost (%)	−55.8	−55.7	−62.7	−82.5	−61.4	−79.1

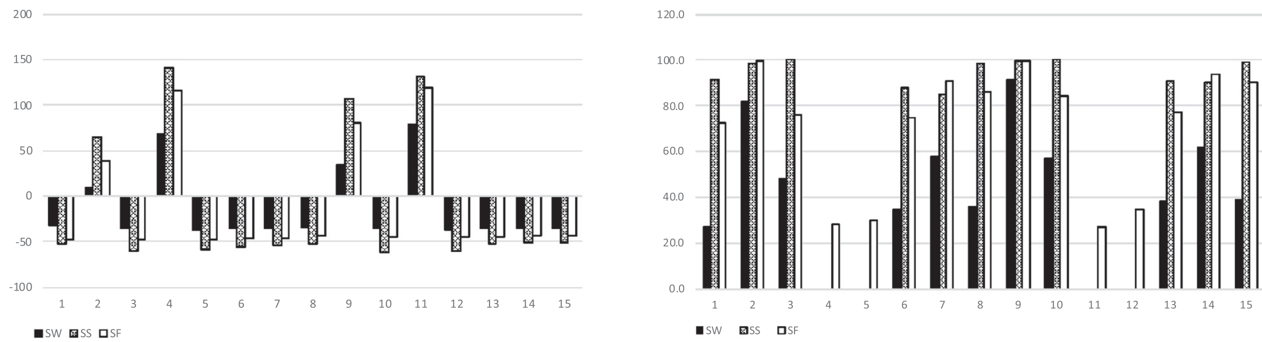
Performance measures. The summary of various performance and cost measures is given in Table 8, where the revenue term is only for online sales and the terms expressed in percentage are to be taken as the percentage difference between each strategy and the baseline ship-from warehouse (SW) strategy.

We observe that with the integration of the SS strategy the percentage of lost online sales is decreased by more than 50%. In the cases in which the SF strategy led to the deployment of an urban FC this percentage further decreases by around 50% compared to the SS strategy. This decrease comes at the cost of increased fixed, usage, and replenishment costs, as can be observed in Table 8. However, the accompanying increase in the revenue, which in the high DR case is 8.2 times that of the SW strategy, justifies the elevated investment cost.

We remark, however, that the percentage of lost online sales is still quite significant even for the SF strategy. For instance, in the high RE and high DR case, 19% of online sales was lost. The deployment strategies we propose do not directly lead to the minimization of lost sales, as online order satisfaction is not forced in our mathematical model and the objective function maximizes the profit. We delve further into the reasons behind this elevated percentage in Section 5.4.

Inventory flows and transportation capacity utilization. Fig. 3 presents a graphical analysis of the store inventory movements and transportation capacity utilization, based on data for two consecutive weeks of the planning horizon, under the three deployment strategies for the *City1* instance with high RE and high DR. In this instance, one store was selected both by the optimal SS and SF solutions. In Fig. 3(a), we show the flows in and out of inventory for this store. As can be seen from this plot, flow quantities are smallest for the SW strategy and largest for the SS strategy. In the SF strategy, because the fulfillment of online orders is shared with the FC, inventory flows are more homogeneous. In Fig. 3(b), we present the percentage transportation capacity utilization under the three deployment strategies. We remark that under the SS strategy the transportation capacity is saturated multiple times during the planning horizon, as the delivery schedules of multiple stores overlap and the online demand increases the replenishment quantities. On the other hand, with the SF strategy we see a smoothing effect due to the fact that the urban FC can be replenished each period. This result is encouraging as it shows the operational benefits of the SF strategy, namely, a more homogeneous utilization of store inventories and transportation capacities while increasing the profitability.

Store selection analysis. To conclude this section, we present a discussion on the role of high physical customer traffic stores in online order fulfillment. To this end, we note the sharp contrast between the *City1* instances in which these stores are prioritized and *City2* instances in which they are avoided. A closer examination of our instances led to the observation that this was due to the overlapping replenishment schedules of the stores coupled with the transportation capacity. In the case of the *City2* instance, this implies that the available transportation capacity is saturated and therefore additional replenishment quantities required for online order fulfillment cannot be transported. This observation is critical as it highlights how omnichannel integration can perturb the current distribution dynamics. In the next section, we perform sensitivity analysis tests to see how the network structure changes when the store replenishment schedules are relaxed.



(a) (b)

Fig. 3. Inventory, replenishment, and transportation capacity analysis under the three deployment strategies.

Table 9
Solutions for the *City 1* instance with increasing values of the capacity multiplier.

	1		1.2		1.5		2	
	SS	SF	SS	SF	SS	SF	SS	SF
#Stores	3 (2)	1 (0)	3 (2)	1 (0)	3 (1)	0	2 (1)	2 (1)
#FCs	0	1	0	1	0	1	0	0
%Lost	34.9	18.9	28.4	9.6	29.1	7.6	28.8	28.8
Revenue (%)			5.0	3.9	3.6	3.7	4.7	-6.9
Fixed cost	21,000	23,000	21,000	23,000	16,000	20,000	11,000	11,000

5.4. Sensitivity analysis

In this section, we present results exploring the sensitivity of solutions to different parameters of the model. To do so, we have selected the *City 1* instance with high RE-high DR attributes unless otherwise specified.

Fulfillment capacity. On this instance, we first introduce a multiplier to scale the capacity of stores and the urban FC and solve the same instance (with $s^{\max} = 3$) with different values of this multiplier for the SS and SF strategies. The results are presented in Table 9 where the revenue term is for online orders only and reflects the percentage difference with respect to the previous parameter setting. As can be observed from this table an increase in capacity does not directly lead to a decrease in the percentage of lost sales. The model rather chooses to decrease fixed and usage costs, as it is possible to obtain more capacity with less investment as the capacity multiplier increases. This may even lead to a decrease in the revenue as seen for the SF strategy when the multiplier is 2. Clearly, the savings in terms of the fixed and usage costs justifies this decrease. Another interesting observation is that when the store capacities are doubled the SF strategy does not deploy an urban FC. It turns out that, in this case, the capacities of the two selected stores are sufficient so that the investment cost of the urban FC is not justified. This indicates that in big cities where there are multiple stores with high online order processing capacity, one does not necessarily need an urban FC.

Maximum number of stores selected. We next change the value of parameter $s^{\max}$, which had been initially set to 3. We then solve the selected instance with each value of the parameter for the SS deployment strategy. The results are presented in Table 10, where the revenue term is from online sales only and reflects the percentage difference with respect to the previous parameter setting. Interestingly, after the value $s^{\max} = 3$, further increasing this parameter does not change the solution. This implies that lost sales of nearly 35% are not really caused by a limitation on this parameter. Another critical aspect of the problem is observed when comparing the number of high physical customer traffic stores selected with $s^{\max} = 2$ and $s^{\max} = 3$. Accordingly, an optimal solution is not necessarily to open the maximum capacity stores but rather to have the best ratio between the fixed cost of opening stores and their capacity.

Replenishment schedule. We repeat our tests with the $s^{\max}$ parameter with an updated replenishment schedule (URS). Accordingly, we assume that all stores can be replenished during each period. The results are presented in Table 11. Based on these results, a first observation is the changing role of the high physical customer traffic stores in order fulfillment. It turns out that these stores were previously favored not because of their capacity but because of their convenient replenishment schedule, implying that this is an important dimension of the problem to consider. Another interesting result is that for the $s^{\max} = 3$ case increasing the replenishment frequency of the stores did not decrease the percentage of lost sales. In fact, the savings in the fixed cost justified the drop in the profit from satisfying more orders. This also explains the final lost sales of 32.4% as in this case the profit earned from these orders does not justify the additional fixed cost of selecting another store.

Usage cost of urban fulfillment centers. Finally, we test the sensitivity of our results to the usage cost of setting up an urban FC that we denote by c_f^{f}. To do so, we use the *City 1* instance with high, moderate, and low RE-high DR attributes. We then test the ship-from fulfillment platforms strategy (SF) with c_f^{f} fixed to 5000, 10,000, 15,000, 20,000 and 25,000. The results are presented in Table 12 where we present the network structure in each case with the number of high physical customer traffic stores indicated in parenthesis. As can be observed from these results the usage cost is a highly determinant factor in deployment of urban FCs. When this cost is sufficiently low, an urban FC is deployed even in the low RE case, whereas when this cost is as high as 25,000, the urban FC is not deployed even in the high RE case. We remark, however, that, the reason for deploying an urban FC is not simply the low usage cost. Indeed, when the usage cost increases further than 10,000, selecting a store is always cheaper in the instances we consider. Therefore, the operational advantages, such as better

Table 10
Solutions for the *City 1* instance with increasing values of $s^{\max}$.

	$s^{\max}=0$	$s^{\max}=1$	$s^{\max}=2$	$s^{\max}=3$	$s^{\max}=4$	$s^{\max}=5$
#Stores	0	1 (1)	2 (1)	3 (2)	3 (2)	3 (2)
%Store	0.0	25.8	39.0	58.6	58.6	58.6
%Warehouse	9.6	9.3	8.5	6.5	6.5	6.5
%Lost	90.4	64.9	52.5	34.9	34.9	34.9
Fixed cost	0	10,000	13,000	21,000	21,000	21,000
Revenue (%)		445.1	25.0	20.7	0.0	0.0

Table 11
Solutions for the *City 1* instance with increasing values of $s^{\max}$ and URS.

	$s^{\max}=0$	$s^{\max}=1$	$s^{\max}=2$	$s^{\max}=3$	$s^{\max}=4$	$s^{\max}=5$
#Stores	0	1 (1)	2 (1)	3 (0)	4 (1)	4 (1)
%Store	0.0	25.8	39.0	42.3	61.8	61.8
%Warehouse	9.6	9.2	8.4	8.1	5.8	5.8
%Lost	90.4	65.0	52.6	49.5	32.4	32.4
Fixed cost	0	10,000	13,000	14,000	22,000	22,000
Revenue (%)		444.4	25.0	4.9	18.2	0.0

Table 12
Solutions for the *City 1* instance with increasing usage cost of urban FCs.

		Low RE	Moderate RE	High RE
$c_f^{\mathrm{f}} \le 15,000$	#Stores	1 (0)	1 (0)	1 (0)
	#FCs	1	1	1
$c_f^{\mathrm{f}} = 20,000$	#Stores	2 (1)	1 (0)	1 (0)
	#FCs	0	1	1
$c_f^{\mathrm{f}} = 25,000$	#Stores	2 (1)	3 (1)	3 (2)
	#FCs	0	0	0

utilization of capacities highlighted in Table 3(b) and availability of a deterministic fulfillment capacity, justify the selection of urban FCs. These additional results further highlight the operational interest of the SF strategy, relying on urban FCs not as a substitute to stores but rather as a complement. However, these results also highlight the current economic limitation of the SF strategy, which could be resolved with the proliferation of on-demand storage services in urban areas.

6. Managerial implications and conclusion

This article defines and formulates an emerging distribution problem for retail businesses aiming to integrate the online channel within a dedicated in-store sales network, namely, the e-DND. This problem involves a deployment level decision on the operation of new fulfillment stores/centers and on the assignment of the online order fulfillment mission to those fulfillment platforms. To improve the quality of the deployment decisions, the proposed model anticipates the revenues and costs of the operational level, which includes replenishment, inventory holding, delivery, and order fulfillment decisions. This hierarchical decision problem is formulated as a two-stage stochastic program with mixed-integer recourse. Due to the inherent uncertainty of the ordering process and the complexity of modeling the order fulfillment problem, the resulting MIP is extremely large and cannot be solved to optimality with commercial solvers for realistic problems. For this reason, an exact solution approach that combines scenario sampling and the integer L-shaped method is proposed. Two sets of computational experiments were presented to test the effectiveness of the solution approach proposed and compare the different deployment strategies considered. Our results show that the integer L-shaped method produces good-quality solutions in a reasonable time compared to alternative approaches with the CPLEX solver. They also show that the integer L-shaped method is capable of solving realistically sized instances. However, we emphasize that the integer L-shaped method could be fine-tuned, for instance, with additional cuts, in order to improve its performance for larger-sized instances and an increased number of scenario samples.

When comparing different deployment strategies, our results showed the improved performance of the proposed ship-from-store (SS) and ship-from urban fulfillment platforms (SF) strategies as compared to the baseline. The ship-from warehouse (SW) strategy showed limitations in terms of potential demand satisfaction, mainly with the increase of the proportion of orders requiring a shorter response time. To this end, the ship-from stores strategy (SS) provided an increase in profitability, which is mainly due to the increased proportion of satisfied online orders. Our results highlighted the importance of considering store capacities for ship-from stores operations. Additional sensitivity analysis highlighted how omnichannel integration can perturb the current distribution dynamics. In this regard, we believe that for effective omnichannel integration one must reconsider in the decision model the replenishment and distribution dynamics, as well as the transportation network capacities. In the ship-from fulfillment platforms strategy (SF) we also observed an increase in profitability, although this strategy led in several instances to the same solution as the SS strategy. Our results were encouraging as they showed operational benefits from the SF strategy, namely, a more homogeneous utilization of store inventories and transportation capacities while increasing the profitability. However, they revealed the sensitivity of solutions to the usage costs incurred by operating FCs, the main managerial insight being that this strategy seems to perform better in cases of high levels of online orders and saturation of store capacities in large cities with the cost structure considered. This further highlights the need for studying various settings of such networks of urban fulfillment platforms from the point of view of omnichannel integration.

The work in this article shows that the integration of the online channel into existing distribution networks is a field in need of further research. For instance, although this article bases its analysis on distributions that depend on parameter estimations coming from real data, there is a need for further incorporation of data science in the scenario generation process. Further, deployment strategies based on emerging practices such as click-and-collect and ship-to-stores could be investigated. A benchmark operating dedicated e-commerce distribution centers could be proposed. Our work emphasizes the importance of modeling omnichannel distribution systems at the tactical and operational levels by tackling the joint replenishment, inventory, and delivery problems. One could additionally explore distribution with alternative delivery modes, as well as to pick-up lockers. We assumed in this work that the deliveries are done to aggregated customer locations but further work could include more detailed constraints in connection with city logistics. Another line of research could also look at the role of fulfillment platforms as a hub point for physical and online demand. As urban logistics will shift to reduce the truck traffic from city centers, fulfillment platforms could take on the role of centralizing the inward flow to cities. Store replenishment and distribution of online orders could then be managed with environmentally friendly smaller vehicles. In this case, a refined assessment of the environmental impact of the omnichannel retailing strategies could be further investigated.

Appendix A

A1. Market condition generation

Let t be a time period in the future (e.g., a period being a day and the planning horizon being a month), and let $\mathbb{P}(m_t = m)$ be the probability of the market condition having value m in time period t, given that we do not know the current market condition. We assume that $\mathbb{P}(m_t = m)$ has a symmetric, bell shaped, truncated probability distribution for $m \in (a, b)$, with mean $\bar{m}$, standard deviation s and lower and upper bounds a and b, respectively. Furthermore, we assume that the market condition for the next period follows a truncated normal distribution where $\bar{m}$, σ, a and b are determined by the current market condition. Let $F(m_t)$ be the market condition for period $t+1$ given the current condition m_t. Moreover, let $\delta_{\bar{m}}$ be a multiplier that determines how fast the mean approaches the steady state mean, $\bar{m}$, and δ_s be a multiplier that determines how fast the standard deviation increases with the distance from steady state mean. Based on the observed market condition m_t, we update the mean using

$$\mu = \begin{cases} m_t + \min\{\delta_{\bar{m}}(\bar{m} - m_t), \bar{m} - m_t\} & \text{if } m_t < \bar{m} \\ m_t - \min\{\delta_{\bar{m}}(m_t - \bar{m}), m_t - \bar{m}\} & \text{if } m_t \geq \bar{m}. \end{cases}$$

Further, we write $s = \underline{s} + \delta_s(m_t - \bar{m})^2$, $a = \max\{m_{\min}, m_t - \beta\}$ and $b = \min\{m_{\max}, m_t + \beta\}$, where $\underline{s}$ is the lowest possible standard deviation for a period, the lowest and the highest market condition are represented by $m_{\min}$ and $m_{\max}$, respectively, and the truncation of the distribution is adjusted by β. Finally, we have that $F(m_t) \sim N(\bar{m}, s^2)$ and $m_t \in (a, b)$.

A2. Scenario generation

In this section, we present the details of the Monte Carlo procedure used to generate online orders, in-store demands, and capacities for a given scenario based on descriptive models discussed in Section 3.4.

Each scenario corresponds to a realization of all the random variables shaping the uncertain future environment. More specifically, the scenario is generated directly from the cumulative distribution functions of the random variables of multi-item online order, demand and capacity processes. The Monte Carlo procedure presented in Algorithm 2 can be used to generate a scenario $\omega \in \Omega$ over the planning horizon T. Repeating this Monte Carlo sampling procedure N times yields the required sample of scenarios N.

For the online order process, assuming that the online orders are independent of each other, we generate independent pseudo-random numbers u uniformly distributed on the interval [0, 1] (denoted $U_c[0, 1]$) and we compute the inverse of the distributions. In this procedure, a continuous variable is used to denote order-arrival times. Order arrivals are generated in the interval $[0, \mathcal{T}]$ and mapped onto the corresponding periods $t \in \mathcal{T}$. For the market condition process, assuming that the market conditions are independent of each other, we generate similarly pseudo-random numbers u uniformly distributed on the interval [0, 1] and compute the inverse of the updated normal distribution. As a starting market condition m_0 we take the steady-state expected market condition $\bar{\mu}$. Based on this parameter, we then update the arrival rates of the physical orders at each store s for each product p as well as the capacity of each store s in each period t. The physical demand for each product p at each store s in period t is then obtained by calculating the inverse of the Poisson distribution with the updated arrival using again the pseudo-random numbers $u \sim U_c[0, 1]$.

Algorithm 2: Monte Carlo procedure for generating a single scenario over the planning horizon.

```
Output: Ō_lt(ω), l ∈ L, t ∈ T, b_st(ω), s ∈ S, t ∈ T, d^p_st(ω), p ∈ P
        s ∈ S, t ∈ T,
for l ∈ L do
    η = 0
    while η ≤ T do
        Generate u ~ U_c[0, 1] and compute the next order
        arrival time η = η + Exp_l^{-1}(u)
        Add order o_l with arrival period τ = ⌈η⌉ to the
        chronological list O_l(ω)
    end
    for o ∈ O_l(ω) do
        Generate u ~ U_c[0, 1] and compute the requested
        response time θ_o(ω) = F^{-1}(u)
        Generate k_o = U_d(k_min, k_max) of the number of
        products in an order
        i = 0
        for i < k_o do
            Generate u ~ U_c[0, 1] and obtain
            p = min{ p' ∈ {[1], ..., [P]} : Σ_{i=1}^{p'} α_[i] − u ≥ 0 }
            Add product p to the order product list P(o)
            Generate u ~ U_c[0, 1] and compute the demand
            n^o_p(ω) = Φ_p^{-1}(u)
        end
    end
end
Aggregate the orders chronologically to obtain Ō_l(ω) for l ∈ L
for t ∈ T do
    Generate u ~ U_c[0, 1] and compute the market condition
    m_t(ω) = Φ_t^{-1}(u), m_t ∈ [a, b]
    Update m̄, s, a and b based on m_t
end
for p ∈ P, s ∈ S and t ∈ T do
    Update the demand arrival rate λ^p_s(m_t(ω)/μ̄)
    Generate u ~ U_c[0, 1] and compute d^p_st(ω) = Poisson_ps^{-1}(u)
end
for s ∈ S and t ∈ T do
    Generate store capacity b_st(ω) = b_s(μ̄/m_t(ω))
end
```

References

Acimovic, J., & Graves, S. C. (2014). Making better fulfillment decisions on the fly in an online retail environment. *Manufacturing & Service Operations Management, 17*(1), 34–51.

Agatz, N. A., Fleischmann, M., & Van Nunen, J. A. (2008). E-fulfillment and multi-channel distribution–a review. *European Journal of Operational Research, 187*(2), 339–356.

Alptekinoğlu, A., & Tang, C. S. (2005). A model for analyzing multi-channel distribution systems. *European Journal of Operational Research, 163*(3), 802–824.

Angulo, G., Ahmed, S., & Dey, S. S. (2016). Improving the integer l-shaped method. *INFORMS Journal on Computing, 28*(3), 483–499.

Bektaş, T., Crainic, T. G., & Van Woensel, T. (2017). From managing urban freight to smart city logistics networks. In *Network design and optimization for smart cities* (pp. 143–188). World Scientific.
Bell, D. R., Gallino, S., & Moreno, A. (2014). How to win in an omnichannel world. *MIT Sloan Management Review, 56*(1), 45.
Ben Mohamed, I., Klibi, W., & Vanderbeck, F. (2020). Designing a two-echelon distribution network under demand uncertainty. *European Journal of Operational Research, 280*(1), 102–123.
Benders, J. F. (1962). Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik, 4*(1), 238–252.
Bendoly, E., Blocher, D., Bretthauer, K. M., & Venkataramanan, M. (2007). Service and cost benefits through clicks-and-mortar integration: Implications for the centralization/decentralization debate. *European Journal of Operational Research, 180*(1), 426–442.
Birge, J. R., & Louveaux, F. (2011). Introduction to stochastic programming (Springer series in operations research and financial engineering), Springer Science & Business Media.
Bretthauer, K. M., Mahar, S., & Venakataramanan, M. (2010). Inventory and distribution strategies for retail/e-tail organizations. *Computers & Industrial Engineering, 58*(1), 119–132.
Chopra, S. (2018). The evolution of omni-channel retailing and its impact on supply chains. *Transportation Research Procedia, 30*, 4–13.
Crainic, T. G., & Montreuil, B. (2016). Physical internet enabled hyperconnected city logistics. *Transportation Research Procedia, 12*, 383–398.
Cuda, R., Guastaroba, G., & Speranza, M. G. (2015). A survey on two-echelon routing problems. *Computers & Operations Research, 55*, 185–199.
De Koster, R. B. (2003). Distribution strategies for online retailers. *IEEE Transactions on Engineering Management, 50*(4), 448–457.
Deborah W. Sam's club is closing one-tenth of its stores and converting 12 to e-commerce hubs: Proof that store closures have become respectable? https://www.linkedin.com/pulse/sams-club-closing-one-tenth-its-stores-converting-12-hubs-weinswig/, 2018. Accessed: 2018-08-01.
Firoozi, M., Babai, M. Z., Klibi, W., & Ducq, Y. (2019). Distribution planning for multi-echelon networks considering multiple sourcing and lateral transshipments. *International Journal of Production Research*, 1–19.
Gallino, S., & Moreno, A. (2014). Integration of online and offline channels in retail: The impact of sharing reliable inventory availability information. *Management Science, 60*(6), 1434–1451.
Gao, F., & Su, X. (2016). Omnichannel retail operations with buy-online-and-pick-up-in-store. *Management Science, 63*(8), 2478–2492.
Govindarajan, A., Sinha, A., & Uichanco, J. (2018). Joint inventory and fulfillment decisions for omnichannel retail networks. *Ross School of Business Paper*, (1341).
Harsha, P., Subramanian, S., & Uichanco, J. (2019). Dynamic pricing of omnichannel inventories: Honorable mention2017 m&som practice-based research competition. *Manufacturing & Service Operations Management, 21*(1), 47–65.
Hübner, A., Holzapfel, A., & Kuhn, H. (2016). Distribution systems in omni-channel retailing. *Business Research, 9*(2), 255–296.
Hübner, A., Wollenburg, J., & Holzapfel, A. (2016). Retail logistics in the transition from multi-channel to omni-channel. *International Journal of Physical Distribution & Logistics Management, 46*(6/7), 562–583.
Ishfaq, R., & Bajwa, N. (2019). Profitability of online order fulfillment in multi-channel retailing. *European Journal of Operational Research, 272*(3), 1028–1040.
Ishfaq, R., Defee, C. C., Gibson, B. J., & Raja, U. (2016). Realignment of the physical distribution process in omni-channel fulfillment. *International Journal of Physical Distribution & Logistics Management, 46*(6/7), 543–561.
Joerss, M., Schroder, J., Neuhau, F., Klink, C., & Mann, F. (2016). Parcel delivery: The future of last mile. *Travel, Transport and Logistics*. McKinsey&Company.
Kim, N., Montreuil, B., Klibi, W., & Kholgade, N. (2020). Hyperconnected urban fulfillment and transportation for last mile delivery of large items. *Working Paper*. Atlanta, GA: Georgia Institute of Technology.
Lang, G., & Bressolles, G. (2013). Economic performance and customer expectation in e-fulfillment systems: a multi-channel retailer perspective. In *Supply chain forum: An international journal: 14* (pp. 16–26). Taylor & Francis.
Laporte, G., & Louveaux, F. V. (1993). The integer l-shaped method for stochastic integer programs with complete recourse. *Operations research letters, 13*(3), 133–142.
Li, Z., Lu, Q., & Talebian, M. (2015). Online versus bricks-and-mortar retailing: A comparison of price, assortment and delivery time. *International Journal of Production Research, 53*(13), 3823–3835.
Mancini, S., Gonzalez-Feliu, J., & Crainic, T. G. (2014). Planning and optimization methods for advanced urban logistics systems at tactical level. In *Sustainable urban logistics: Concepts, methods and information systems* (pp. 145–164). Springer.
Martel, A. (2003). Policies for multi-echelon supply: DRP systems with probabilistic time-varying demands. *INFOR: Information Systems and Operational Research, 41*(1), 71–91.
Martel, A., & Klibi, W. (2016). *Designing value-creating supply chain networks*. Springer.
Melacini, M., Perotti, S., Rasini, M., & Tappia, E. (2018). E-fulfilment and distribution in omni-channel retailing: A systematic literature review. *International Journal of Physical Distribution & Logistics Management, 48*(4), 391–414.
Montreuil, B. (2016). "Omnichannel Business-to-Consumer Logistics and Supply Chains: Towards Hyperconnected Networks and Facilities". 14th IMHRC Proceedings (Karlsruhe, Germany 2016). 19.
Paula, J., Agatza, N., Splietb, R., de Kostera, R. (2017). Customer consolidated routing problem–an omni-channel retail study.
Savelsbergh, M., & Van Woensel, T. (2016). 50th anniversary invited article city logistics: Challenges and opportunities. *Transportation Science, 50*(2), 579–590.
Shapiro, A., Dentcheva, D., & Ruszczynski, A. (2009). *Lectures on stochastic programming: modeling and theory*. Philadelphia, USA: The Society for Industrial and Applied Mathematics and the Mathematical Programming Society.
Tompkins, B., & Ferrell, C. (2016). *E-commerce supply chain: Survey results*. Tompkins Supply Chain Consortium.
Torabi, S., Hassini, E., & Jeihoonian, M. (2015). Fulfillment source allocation, inventory transshipment, and customer order transfer in e-tailing. *Transportation Research Part E: Logistics and Transportation Review, 79*, 128–144.
Van Slyke, R. M., & Wets, R. (1969). L-shaped linear programs with applications to optimal control and stochastic programming. *SIAM Journal on Applied Mathematics, 17*(4), 638–663.