CIRRELT

Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport

Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation

Sequential versus Integrated Optimization: Lot Sizing, Inventory Control and Distribution

Maryam Darvish
Leandro C. Coelho

October 2016

CIRRELT-2016-60

Document de travail également publié par la Faculté des sciences de l'administration de l'Université Laval, sous le numéro FSA-2016-015.

Bureaux de Montréal :
Université de Montréal
Pavillon André-Aisenstadt
C.P. 6128, succursale Centre-ville
Montréal (Québec)
Canada H3C 3J7
Téléphone : 514 343-7575
Télécopie : 514 343-7121

Bureaux de Québec :
Université Laval
Pavillon Palasis-Prince
2325, de la Terrasse, bureau 2642
Québec (Québec)
Canada G1V 0A6
Téléphone : 418 656-2073
Télécopie : 418 656-2624

www.cirrelt.ca

Sequential versus Integrated Optimization: Lot Sizing, Inventory Control and Distribution

Maryam Darvish*, Leandro C. Coelho

Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT) and Department of Operations and Decision Systems, 2325 de la Terrasse, Université Laval, Québec, Canada G1V 0A6

Abstract. In this paper we study an integrated lot sizing problem with distribution and delivery time windows in which the producer has the choice to locate distribution centers from where products are stored and shipped to customers. Based on a real case study, we describe, model, and solve this rich integrated lot sizing problem. In this integrated optimization application, we consider a multi-plant, multi-product and multi-period lot sizing problem. The goal is to minimize fixed and variable production costs, inventory, and distribution costs while satisfying demands within a promised delivery time window. Our work contributes to the integrated optimization literature by simultaneously addressing production, inventory, and distribution problems, and to the production economics literature by comparing and assessing the performance of sequential and integrated solution techniques. We develop an exact method and several heuristics, based on separately solving each part of the problem, as well as a general integrated matheuristic. Our results and analysis not only compare solution costs but also highlight the value of an integrated approach.

Keywords. Physical internet, dynamic lot-sizing, integrated supply chain planning, production, inventory, distribution, optimization.

Acknowledgements. This project was partly funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) under grant 2014-05764. This support is greatly acknowledged.

---

* Corresponding author: Leandro.Coelho@cirrelt.ca

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2016

1 Introduction

The ultimate goal of any production system is to fulfill the demands of its customers quickly and efficiently. Increasing competition has forced companies to seek solutions that not only save costs and improve the efficiency but also provide faster and more flexible services to customers. Companies have realized that dramatic improvements are to occur by exploiting integrated production systems. Integration and coordination of supply chain decisions such as production lot sizing, distribution, and inventory decisions yield competitive and economic advantage [15]. In this new integrated paradigm, various functions of a company are simultaneously taken into consideration and jointly optimized. Historically, a hierarchical approach has been used to decompose a difficult problem into smaller and easier to solve cases, treating each one separately from the others. Recently, partial integration of supply chain functions has been proved beneficial [2, 9, 36]. In the hierarchical production planning approach, the solution obtained from one level is passed and imposed to the next one in the hierarchy of decisions, often resulting in sub-optimal solutions [40]. Management typically done in silos results in each department of a company making its own decisions, regardless of what other departments are doing, and even without considering the global strategy. Accordingly, most research on supply chain integration and case studies confirm the positive effect of integration on business performance [1, 2, 10].

One of the most important and challenging issues in production planning is to determine production quantities and their timing, known as the lot sizing problem (LSP) [22]. In production LSP, one of the earliest efforts in simultaneous optimization [41], production and inventory costs are minimized. Recently, integrated lot sizing optimization problems have attracted both industrial and academic interest mainly due to their vast applicability in solving real-world problems, and advantageous cost effective implications compared to hierarchical planning processes [42]. Integrated LSP aims at determining the production quantities and timing, inventory quantities and their locations, shipment quantities and delivery schedule, in order to minimize the total costs and fulfill customers' demands.

Starting from the seminal paper of Wagner and Whitin [41], literature on the LSP has developed at a rapid pace, and numerous studies on the variants of the LSP are reported in the literature [7]. Axsäter [5] provides a comprehensive review on LSPs; Jans and Degraeve [19] give an overview of modeling deterministic single-level dynamic LSPs, and Robinson et al. [32] review the multi-item capacitated dynamic demand coordinated LSPs. The abundant LSP literature asserts that the complexity of lot sizing models is directly related to the considered features [18, 22]. So far, different variants of the LSP have been integrated with inventory, distribution, and routing decisions. Following this trend, the problem studied in this paper combines LSP with four distinct features: direct shipment, delivery time windows, facility location decisions, and a leasing period for distribution centers (DCs). In what follows we briefly review the relevant LSP literature. A list of these papers with their characteristics is presented in Table 1.

The transportation decision in integrated LSPs is considered as either direct shipment (full-truck loads) or vehicle routes (milk runs). A survey of formulations and solution algorithms for the combination of the LSP and the vehicle routing problem, known as the production routing problem, is presented in Adulyasak et al. [2]. In this paper, however, we consider direct shipment due to the increasing growth in the number of firms outsourcing the transportation function to third party logistics service providers [4]. Dhaenens-Flipo and Finke [13] consider a real-case consisting of multiple facilities, products, periods and stages. While successful in solving small instances, their branch-and-bound approach could not prove optimality for larger ones. The capacitated LSP in a multi-product, multi-plant, multi-period LSP is solved using a Lagrangian-based heuristic in Sambasivan and Yahya [35] and with a GRASP heuristic in Nascimento et al. [30]; in both papers, only the inter-plant transfers are considered as opposed to the direct shipments between plants and DCs, and between DCs and customers that we address here. An uncapacitated two-level LSP is solved by dynamic programming (DP) in Melo and Wolsey [27]. The same authors formulate the problem and its special cases as a mixed integer program (MIP) and solve it using a hybrid heuristic, providing good solutions in reasonable time for medium size instances [28]. Steinrücke [39] studies the multi-stage production-shipping and distribution-scheduling problem in an aluminum industry by means of a relax-and-fix decomposition method. Hwang et al. [17] model a deterministic two-stage capacitated LSP with discrete time horizon as a minimum network flow problem and identify the production and transportation quantities in a two-stage plant-retailer supply chain. Finally, Solyalı et al. [37] present polynomial-time algorithms for capacitated and uncapacitated two-level LSP with backlogging.

Due to its significant research and practical potential, much attention is devoted toward time windows in the LSP. Generally, they appear as either at the production or at the delivery side [3]. We consider a delivery time window meaning that the demand must be satisfied within the specified time slot. However, products can be produced earlier and be stored prior to the delivery time, whereas under the presence of a production time window, nothing can be produced prior to the demand realization [6] . In Lee et al. [24], a polynomial-time DP algorithm is proposed to solve the dynamic uncapacitated LSP with demand time windows, with no shortage but with and without backlogging. The LSP with time windows studied in Merzifonluoğlu and Geunes [29] differs from that of Lee et al. [24] as the demand could be accepted or rejected within the customer specified time windows. The goal is to maximize the net profit by increasing the accepted orders while decreasing the LSP related costs. They solve the uncapacitated problem by DP and with a dual-based heuristic. Wolsey [43] formulates LSPs with production and delivery time windows, but only solves the delivery time window case, employing DP algorithms and linear programming relaxation. Another paper that generalizes the classical dynamic capacitated LSP by adding delivery and/or production time windows is that of Hwang et al. [16]. Employing a polynomial-time DP algorithm, they solve a single plant, single product, multi-period capacitated LSP. Akbalik and Penz [3] compare just-in-time and time windows policies for a two level LSP with storage availability at both the plant and the customer levels. Their study shows the potential cost benefits of the time window approach compared to the just-in-time one.

Table 1: Related literature

Author(s)	**LSP characteristics** Number of products	Delivery time window	Direct shipment	Facility location	Leasing period	**Solution method**
Dhaenens-Flipo and Finke [13]	multi		✓			Branch-and-bound
Sambasivan and Yahya [35]	multi		✓			Lagrangean-based heuristic
Nascimento et al. [30]	multi		✓			GRASP with path-relinking
Melo and Wolsey [27]	single		✓			DP
Melo and Wolsey [28]	multi		✓			Hybrid neighborhood search heuristic
Camacho-Vallejo et al. [8]	single		✓			Relax-and-fix decomposition method
Steinrücke [39]	single		✓			Scatter search
Hwang et al. [17]	single		✓			DP
Solyalı et al. [37]	single		✓			DP
Lee et al. [24]	single	✓				DP
Merzifonluoğlu and Geunes [29]	single	✓				DP, dual-based heuristic
Hwang et al. [16]	single	✓				DP
Akbalik and Penz [3]	single	✓				DP
Wolsey [43]	multi	✓				DP
Jaruphongsa et al. [20]	single	✓	✓			DP
Jaruphongsa et al. [21]	single	✓	✓			DP
Keskin and Üster [23]	single		✓	✓		Local search algorithm
Nezhad et al. [31]	multi		✓	✓		Lagrangean-based heuristic
Liang et al. [25]	multi		✓	✓		Column generation & decomposition method
Darvish et al. [12]	single	✓	✓	✓		Branch-and-bound
This paper	**multi**	✓	✓	✓	✓	**Branch-and-bound, ALNS**

The integration of LSP with time windows and direct shipment distribution is also investigated in some papers. Motivated by third-party logistics and real-life applications, Jaruphongsa et al. [21] extend the work of Lee et al. [24] to a two-echelon dynamic lot sizing model with demand time windows, early shipment penalties, and capacity constraints. Later, they solve the same problem but add the late shipping penalties, solving it by DP [20].

Traditionally, determining the facility locations and their capacities involve strategic decisions, and all other decisions such as production quantities, inventories and transportation follow on. In modern days that customers always impose tighter delivery time windows, keeping a high service level and managing inventory require simultaneous production and dynamic facility location planning. This problem is so prevalent that flexible network integration is identified as one of the important recent trends in logistics [38]. Considering the trade-offs among conflicting costs such as fixed (rental or ownership), inventory holding, and transportation as well as proximity to the customer zones, network design needs to be reviewed periodically due to changing demand patterns. This mainly addresses decisions related to the facilities that should be selected and the customers that should be served from these facilities [26]. Again, as separate problems, they have been vastly studied, but the integration of location decisions with LSPs has not been extensively studied. Next, we outline the recent facility location literature which considers production setup costs.

Keskin and Üster [23] model a production/distribution problem consisting of retailers, DCs and potential capacitated suppliers. The results support significant advantages of the integrated approach. Nezhad et al. [31] define an uncapacitated multi-product production/distribution facility location problem in which the known demands of customers need to be met from a set of potential plants while the total of direct shipment, production, and setup costs are minimized. They propose two Lagrangian-based heuristics, and solve small size instances to optimality. Liang et al. [25] present a column generation and decomposition method for the integrated capacitated production planning and facility location problem with backlogging. We classify their work as the integrated LSP and facility location problem since they introduce a set of binary facility location variables. Inspired by a real-case problem, Darvish et al. [12] investigate a rich integrated capacitated LSP for a single-product, multi-plant, and multi-period problem. They incorporate direct shipment, time windows and facility location decisions. A branch-and-bound approach is used to solve the problem, assessing the trade-offs between costs and fast deliveries; the results show the competitive advantage of the integrated approach both in terms of total costs and service level.

This paper extends the work of Darvish et al. [12] in two distinguishing ways. First, we consider multiple products, each potentially incurring different production, holding, and transportation costs; second, we assume that some DCs are rented in each period, and this decision imposes that the DC must remain open for a fixed number of periods. The problem we investigate here is also motivated by our observations from an industrial partner. The company is facing a steady but gradual increase in demand which requires expanding the operations, and has invested abundant capital on its production and storage facilities. Production capacities currently exceed the demand of the company yet with the demand growth rate, capacity constraints seem to be fated. Currently, the production manager makes decisions on the production quantities and scheduling, which are later dispatched by the transportation manager. At this point, the company is interested in how to conduct production planning to save on costs while still maintaining a high service level.

In this paper we describe, model, and solve a multi-product, multi-plant, multi-period, three-level production, inventory, and distribution problem. To the best of our knowledge, this integrated problem with the possibility of choosing among different DCs for a fixed number of periods has not yet been studied in the literature. We first solve the problem exactly. Then, we propose three heuristics and a lower bound procedure. Two of our heuristics are sequential non-integrated decision making processes; one of them mimics the current situation of the company, and the other is used as a benchmark; both are used to assess whether the integrated decision making approach we propose is valuable.

This third heuristic is a hybrid adaptive large neighborhood search (ALNS) in which subproblems are solved exactly. In summary, the main contributions of this paper are as follows. First, we develop a mathematical model for the integrated production, inventory and distribution decisions which also includes the facility location decision and delivery times windows. Second, we present exact and heuristic methods to solve the integrated problem. Finally, we demonstrate the value of the integrated approach by comparing its costs with those obtained from traditional hierarchical decision making approach, and we evaluate the quality and performance of our heuristic solutions by comparing them against the solutions obtained from the exact approach.

The remainder of this paper is organized as follows. In Section 2 we formally describe and model the problem at hand. This is followed by a description of the techniques used to solve it sequentially in Section 3, in which we describe all three heuristic approaches. Our proposed integrated ALNS is explained in Section 4. We present the results of extensive computational experiments in Section 5, followed by our conclusions in Section 6.

2 Problem description and mathematical formulation

We now formally describe the integration of dynamic lot sizing, facility location, inventory management, and distribution with delivery time windows. We consider a set of plants available over a finite time horizon producing multiple products; setting up the production facilities incur a setup cost plus a variable cost proportional to the quantity produced. Each plant owns a warehouse where the products are stored and an inventory holding cost is due. The products are sent to DCs from each plant, to be shipped to the final customers. There is a set of potential DCs from which some are selected to be open, when a fixed cost is due and the DC remains open for a given duration of time, charging an inventory holding cost per product per period. These products are then shipped to the customers scattered geographically to satisfy their demand. There is a maximum allowed lateness for the delivery of products to customers, meaning that the demand must be met within the pre-specified delivery time window. A logistic service provider is in charge of all shipments from plants to DCs as well as from DCs to final customers. The transportation cost associated with the shipping is proportional to the distance, quantity, and type of products delivered.

Formally, the problem is defined on a graph $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ where $\mathcal{N} = \{1, \ldots, n\}$ is the node set and $\mathcal{A} = \{(i,j) : i,j \in \mathcal{N}, i \neq j\}$ is the arc set. The node set $\mathcal{N}$ is partitioned into a plant set $\mathcal{N}_p$, a DC set $\mathcal{N}_d$, and a customer set $\mathcal{N}_c$, such that $\mathcal{N} = \mathcal{N}_p \cup \mathcal{N}_d \cup \mathcal{N}_c$. Let $\mathcal{P}$ be a set containing P products, and $\mathcal{T}$ be the set of discrete periods of the planning horizon of length T. The inventory holding cost of product p at node $i \in \mathcal{N}_p \cup \mathcal{N}_d$ is denoted h_{pi}, the unit shipping cost of product p from the plant i to the DC j is c_{pij}, and the unit shipping cost of product p from the DC j to the customer k is c'_{pjk}. Let also f_i be the fixed opening cost for DC i, after which it will remain open for the next g periods, s_{pi} be the fixed setup cost per period for product p in plant i, v_{pi} be the variable production cost of product p at plant i, and d^t_{pi} be the demand of customer i for product p in period t. The demand occurring in period t can be fulfilled up to period $t + r$, as r represents the delivery time window. For ease of representation, let D be the total demand for all products from all customers in all periods, i.e., $D = \sum\limits_{t \in \mathcal{T}} \sum\limits_{p \in \mathcal{P}} \sum\limits_{i \in \mathcal{N}_c} d^t_{pi}$. In order to simultaneously minimize the sum of the production, inventory and distribution costs in such a way that all demands are met within the promised delivery time window, one needs to determine:

- the product(s) and quantities to be produced in each plant in each period,
- the DCs to be selected in each period,
- the quantity of products sent from plants to DCs in each period,
- the period in which demand of any customer is satisfied,
- the quantity of products sent from DCs to customers in each period.

We formulate the problem with the following binary variables. Let θ_{pi}^t be equal to one if product p is produced at plant i in period t, and zero otherwise; λ_i^t be equal to one if and only if DC i is chosen in period t to be used for g consecutive periods, and ω_i^t be equal to one to indicate whether DC i is open in period t. Integer variables to represent quantities produced and shipped are defined as follows. Let $\alpha_{pij}^{tt'}$ be the quantity of product p delivered from DC i to customer j in period t to satisfy the demand of period t', with $t \geq t'$, ρ_{pi}^t represent the quantity of product p produced at plant i in period t, β_{pij}^t represent the quantity of product p delivered from plant i to DC j in period t, κ_{pi}^t as the amount of product p held in inventory at DC i at the end of period t, and μ_{pi}^t the amount of product p held in inventory at plant i in the beginning of period t. The problem is then formulated as follows:

$$\begin{aligned}\min \sum_{p\in\mathcal{P}}\sum_{i\in\mathcal{N}_p}\sum_{t\in\mathcal{T}} v_{pi}\rho_{pi}^t + \sum_{p\in\mathcal{P}}\sum_{i\in\mathcal{N}_p}\sum_{t\in\mathcal{T}} s_{pi}\theta_{pi}^t + \sum_{p\in\mathcal{P}}\sum_{i\in\mathcal{N}_d}\sum_{t\in\mathcal{T}} h_{pi}\kappa_{pi}^t + \sum_{p\in\mathcal{P}}\sum_{i\in\mathcal{N}_p}\sum_{t\in\mathcal{T}} h_{pi}\mu_{pi}^t + \\ \sum_{i\in\mathcal{N}_d}\sum_{t\in\mathcal{T}} f_i\lambda_i^t + \sum_{p\in\mathcal{P}}\sum_{i\in\mathcal{N}_p}\sum_{j\in\mathcal{N}_d}\sum_{t\in\mathcal{T}} c_{pij}\beta_{pij}^t + \sum_{p\in\mathcal{P}}\sum_{i\in\mathcal{N}_d}\sum_{j\in\mathcal{N}_c}\sum_{t\in\mathcal{T}}\sum_{t'\in\mathcal{T}} c'_{pij}\alpha_{pij}^{tt'}\end{aligned} \quad (1)$$

subject to:

$$\begin{aligned}
\rho_{pi}^t &\leq \theta_{pi}^t D \quad i \in \mathcal{N}_p, t \in \mathcal{T}, p \in \mathcal{P} && (2)\\
\mu_{pi}^t &= \mu_{pi}^{t-1} + \rho_{pi}^t - \sum_{j\in\mathcal{N}_d} \beta_{pij}^t \quad p \in \mathcal{P}, i \in \mathcal{N}_p, t \in \mathcal{T} \setminus \{1\} && (3)\\
\mu_{pi}^1 &= \rho_{pi}^1 - \sum_{j\in\mathcal{N}_d} \beta_{pij}^1 \quad p \in \mathcal{P}, i \in \mathcal{N}_p && (4)\\
\kappa_{pi}^t &= \kappa_{pi}^{t-1} + \sum_{j\in\mathcal{N}_p} \beta_{pji}^t - \sum_{j\in\mathcal{N}_c}\sum_{t'=t-r}^{t} \alpha_{pij}^{tt'} \quad p \in \mathcal{P}, i \in \mathcal{N}_d, t \in \mathcal{T} \setminus \{1\} && (5)\\
\kappa_{pi}^1 &= \sum_{j\in\mathcal{N}_p} \beta_{pji}^1 - \sum_{j\in\mathcal{N}_c} \alpha_{pij}^{11} \quad p \in \mathcal{P} \quad i \in \mathcal{N}_d && (6)\\
\sum_{p\in\mathcal{P}} \kappa_{pi}^t &\leq \omega_i^t D \quad i \in \mathcal{N}_d \quad t \in \mathcal{T} && (7)
\end{aligned}$$

$$\sum_{p\in\mathcal{P}} \kappa_{pi}^t \leq \omega_i^{t+1} D \quad i \in \mathcal{N}_d \quad t \in \mathcal{T} \setminus \{T\} \tag{8}$$

$$\sum_{\substack{t'=t-g+1 \\ t'\geq 1}}^{t} \lambda_i^{t'} \geq \omega_i^t \quad i \in \mathcal{N}_d \quad t \in \mathcal{T} \tag{9}$$

$$\sum_{t'=t}^{t'=t+g-1,\ t'\leq T} \omega_i^{t'} \geq \lambda_i^t min(g, T-t) \quad i \in \mathcal{N}_d \quad t \in \mathcal{T} \tag{10}$$

$$\sum_{t'=t}^{t'=t+g-1,\ t'\leq T} \lambda_i^{t'} \leq 1 \quad i \in \mathcal{N}_d \quad t \in \mathcal{T} \tag{11}$$

$$\sum_{i\in\mathcal{N}_d} \sum_{t=1}^{s} \sum_{t'=1}^{t} \alpha_{pij}^{tt'} \leq \sum_{t=1}^{s} d_{pj}^t \quad p \in \mathcal{P} \quad j \in \mathcal{N}_c \quad s \in \mathcal{T} \tag{12}$$

$$\alpha_{pij}^{tt'} = 0 \qquad p \in \mathcal{P} \qquad i \in \mathcal{N}_d \quad j \in \mathcal{N}_c \quad t \in \mathcal{T} \quad t' \in \{0, ..., t-r\} \cup \{t, ..., T\} \tag{13}$$

$$\sum_{i\in\mathcal{N}_d} \sum_{t=1, t<T}^{s+r} \sum_{t'=1}^{s} \alpha_{pij}^{tt'} \geq \sum_{t=1}^{s} d_{pj}^t \quad p \in \mathcal{P} \quad j \in \mathcal{N}_c \quad s \in \mathcal{T} \tag{14}$$

$$\sum_{p\in\mathcal{P}} \sum_{j\in\mathcal{N}_c} \sum_{t'=1}^{t} \alpha_{pij}^{tt'} \leq D\omega_i^t \quad i \in \mathcal{N}_d \quad t \in \mathcal{T} \tag{15}$$

$$\sum_{i\in\mathcal{N}_d} \sum_{t\in\mathcal{T}} \alpha_{pij}^{tt'} = d_{pj}^{t'} \quad p \in \mathcal{P} \quad j \in \mathcal{N}_c \quad t' \in \mathcal{T} \tag{16}$$

$$\omega_i^t, \theta_{pi}^t, \lambda_i^t \in \{0, 1\} \tag{17}$$

$$\rho_{pij}^{tt'}, \kappa_{pi}^t, \alpha_{ijp}^{tt'}, \beta_{ijp}^t \in \mathbb{Z}^*. \tag{18}$$

The objective function (1) minimizes the total cost consisting of the production setup and variable costs, inventory holding costs and transportation from plants to DCs and from the DCs to final customers. Constraints (2) guarantee that only products set up for production are produced. Constraints (3) and (4) ensure the inventory conservation at each plant. Similarly, constraints (5) and (6) are applied to DCs. Constraints (7) and (8) guarantee that the remaining inventory at the DC is transferred to the next period only if it is open. Constraints (9)–(11) ensure that once a DC is selected, it will be open for the next g consecutive periods. Constraints (12) and (13) guarantee that no demand is satisfied in advance, while constraints (14) impose r periods as the maximum allowed lateness for fulfilling the demand. Thus, the total demand up to period t must be delivered by period $t + r$. No delivery to customers can take place from a closed DC as ensured by constraints (15). Constraints (16) make sure that every single demand is delivered to the customers. Finally, constraints (17) and (18) define the domain and nature of the variables.

3 Sequential heuristics and lower bound

In this section we propose three sequential heuristic algorithms to solve the problem. These algorithms work by making decisions in steps, in a non-integrated fashion. Their motivation is twofold. First, we want to mimic production systems managed in silos, in which one department makes decisions irrespective of the others. Second, we want to assess how a sequential algorithm performs compared to the integrated approach proposed in this paper. These comparisons are presented in Section 5. In what follows, we present in Section 3.1 a Top-down heuristic, for the cases in which production is the most important part of the process and has priority in determining how the system works, followed by inventory allocation in DCs and finally by distribution. In Section 3.2 we describe a Bottom-up heuristic, simulating the alternative scenario in which distribution has priority, followed by DC allocation, and lastly by production. Finally, in Section 3.3 we describe an Equal power heuristic, in which all three departments would have similar but independent power; we explain how this heuristic yields a lower bound on the optimal cost.

3.1 Top-down heuristic

In the Top-down heuristic, production managers have the most power and can therefore determine how the rest of the system works. This is done by taking all constraints of the problem into account, but optimizing it only for the production costs. Once production managers have decided what to do, their decision is fixed, and inventory allocation decisions are made, namely when and which DCs to open. Note that because the previous production decisions were made considering all constraints of the problem, feasibility is ensured. This is done by solving the same problem with a new set of fixed decisions (related to production), and optimizing only DC-related costs. Once this part is determined, all these decisions are fixed and no longer change. Finally, all production and DC variables are known, and one must optimize the remaining variables of the problem, namely all distribution to final customers. By putting together all three levels of decisions, one can obtain the overall solution and easily compute the cost of the solution yielded by the Top-down heuristic. A pseudocode of this heuristic is presented in Algorithm 1.

Algorithm 1 Top-down heuristic

```
Consider all constraints of the problem formulation from Section 2.
Build an objective function with production variables θ^t_pi and ρ^t_pi.
Optimize the problem, obtain optimal values for θ^t_pi and ρ^t_pi.
Fix θ^t_pi and ρ^t_pi to their obtained values.
Add DC-related variables λ^t_i to the objective function.
Optimize the problem, obtain optimal values λ^t_i and ω^t_i.
Fix variables λ^t_i and ω^t_i to their obtained values.
Add all variables to the objective function, as it is defined in Section 2.
Optimize the problem, obtain optimal values for all variables.
Return the objective function value.
```

3.2 Bottom-up heuristic

In the Bottom-up heuristic, we suppose that the distribution managers have the most power and can therefore determine how the rest of the system works. This is done by taking all constraints of the problem into account, but optimizing it only for the distribution variables. Once distribution managers have decided what to do, their decision is fixed, and inventory allocation decisions are made, namely when and which DCs to open. Again, because the previous decisions were made considering all constraints of the problem, feasibility is guaranteed. We now solve the same problem with a new set of fixed decisions (related to distribution), and optimize only DC-related costs. When this part is determined, all these decisions are fixed and no longer change. Finally, once DCs have been selected, and all distribution and DC variables are known, we must optimize the remaining variables of the problem. By putting together all three levels of decisions, one can obtain the overall solution and easily compute the cost of the solution yielded by the Bottom-up heuristic. A pseudocode of this heuristic is presented in Algorithm 2.

Algorithm 2 Bottom-up heuristic

```
Consider all constraints of the problem formulation from Section 2.
Build an objective function with distribution variables α_pij^tt'.
Optimize the problem, obtain optimal values for α_pij^tt'.
Fix α_pij^tt' to their obtained values.
Add DC-related variables λ_i^t to the objective function.
Optimize the problem, obtain optimal values of λ_i^t and ω_i^t.
Fix variables λ_i^t and ω_i^t to their obtained values.
Add all variables to the objective function, as it is defined in Section 2.
Optimize the problem, obtain optimal values for all variables.
Return the objective function value.
```

3.3 Equal power heuristic

In the Equal power heuristic, we assume that all three decision levels have equal power, and that decisions are made in parallel: each department optimizes its own decisions, and all information is shared with the others at the same time. This will likely yield an infeasible solution, since each part is optimized individually and decisions are not made in sequence. However, the sum of the costs of all three levels indicates the optimal decision for each level, when the costs of the other levels are not considered. Having all three put together, if these three decision levels yield a feasible solution, it is obviously the optimal. Moreover, if the solution is not optimal, the sum of these costs constitutes a valid lower bound on the costs of the problem. Algorithm 3 describes the pseudocode for this heuristic.

4 Integrated solution algorithm

Most variants of LSPs are known to be hard to solve and therefore, over the past years, several techniques are proposed for their resolution [19]. Although the uncapcitated LSP is known to be easier to solve, the multi-plant version is NP-complete [34]. Our problem is reducible to the multi-plant uncapacitated LSP and the joint-replenishment problem, an extension of the uncapacitated fixed charge network flow [27] and known to be NP-hard [11]. While small-size instances of our problem can be solved to optimality in reasonable time using exact methods, we propose a matheuristic for solving large instances. Our ALNS-based algorithm uses a com-

Algorithm 3 Equal power heuristic

1: Consider all constraints of the problem formulation from Section 2.
2: Build an objective function with distribution variables $\alpha_{pij}^{tt'}$.
3: Optimize the problem, obtain optimal values for $\alpha_{pij}^{tt'}$, and optimal solution z_c.
4: Build an objective function with DC-related variables λ_i^t, β_{pij}^t, and κ_{pi}^t.
5: Optimize the problem, obtain optimal values for λ_i^t, ω_i^t, β_{pij}^t, and κ_{pi}^t, and optimal solution z_d.
6: Build an objective function with plant-related variables θ_{pi}^t, ρ_{pi}^t, and μ_{pi}^t.
7: Optimize the problem, obtain optimal values for θ_{pi}^t, ρ_{pi}^t, and μ_{pi}^t, and optimal solution z_p.
8: **if** the combination of all three decisions is feasible **then**
9: Return optimal solution and its cost $z^* = z_p + z_d + z_c$.
10: **else**
11: Return lower bound value $\underline{z} = z_p + z_d + z_c$.
12: **end if**

bination of a heuristic and an exact method, employed iteratively to find good solutions. The ALNS introduced by Ropke and Pisinger [33] has shown outstanding results in solving various supply chain management problems. This method is not only flexible but also very efficient as it explores large complex neighborhoods and avoids local optima. Hence, this type of algorithm is highly suitable for the problem at hand due to its generality and flexibility. Our proposed heuristic has three main levels. At the first one, we develop an ALNS heuristic in order to decide whether a plant or a DC should be open, and to determine which products have to be produced. Once these decisions are fixed, all the other remaining decisions on deliveries from open plants to open DCs, and from open DCs to the customers, as well as the inventory level held at plants and open DCs are computed exactly and optimally by solving an integer linear programming sub-problem, which is done efficiently by exploiting its structure via a minimum cost network flow (MCNF) algorithm. The MCNF finds a feasible flow with minimum cost on a graph in which a cost is associated to each arc [14]. Finally, if needed and to avoid local optimum solutions, we improve the obtained solution to move toward the global optimal by solving the model presented in Section 2 exactly for a very short period of time. We explain the details of each level next.

We start with generating a feasible initial solution by making all plants and DCs available in every period. This feasible initial solution is quickly improved by closing as many facilities as possible, but maintaining feasibility, while costs are not yet of concern. Once the initial solution s and its corresponding cost $z(s)$ are obtained, we improve the solution by applying the iterative heuristic.

At each iteration, one operator from the list described in Section 4.1 is selected according to a roulette-wheel mechanism. A weight is associated to their previous performances modulating their chances of being selected. The weight matrix which has initial value of one is then updated at every φ iterations. Operators are given a score based on their performances, initially set to zero, and the better the operator performs, the higher score it accumulates. A simulated annealing-based acceptance rule is considered here. The current solution s is accepted over the incumbent solution s' with the probability of $e^{(\frac{z(s') - z(s)}{H})}$, where H is the current temperature. The temperature is decreased every iteration by α. Once the temperature reaches the final temperature, H_{final}, it is reset to the initial temperature, H_{start}. The stopping criteria for our algorithm is met when either the maximum number of iterations $iter_{max}$ or the maximum allotted time is reached, if an optimal solution is obtained, or if the solution has not changed for the last $\frac{iter_{max}}{2}$ iterations. As long as the best cost $z(s_{best})$ is changing, the search starts at every φ iterations from the best solution. If no improvement is made for more than 2φ iterations, once the temperature has decreased to less than the final temperature, we use the best known solution as an input to the model of Section 2 and solve it for 20 seconds; if this model yields an optimal solution, the algorithm stops as the global optimum has been found; if it improves the solution, this improvement is passed to the ALNS framework and the procedure continues.

4.1 List of operators

At each iteration, one of the following heuristic operators are selected. The operators work for any type of facility, be it plants or DCs. Each operator is repeated a number of times, drawn from a semi-triangular distribution, with the highest probability toward small numbers. The operators are as follows:

- Random: this operator selects one plant, product and period (or DC and period), and flips its current decision; if the facility is closed it becomes open, and vice-versa.

- Based on shipping costs: first, for each product, shipping costs from closed plants to open DCs are compared, and the lowest one is identified. The corresponding product is then assigned to be produced at that plant in that period. Similarly, the highest shipping cost induces a product to have its production stopped at a given plant and period.

- Based on unit costs: among all currently open plants we identify the one producing at the highest unit cost; this plant and product allocation are closed; for DCs, we close the one with the lowest unit inventory cost.

- Based on demand: among all closed plants, we open those that could fulfill the products with the highest demand per period. Similarly, we identify the plant satisfying the minimum demand in a certain period, and close it.

- Based on delivery quantity to DCs: we identify the plant delivering the least (most), as well as the DC receiving the least (most) per period. Facilities with the least usage are closed; for those with the largest usage, a random DC is opened in the same period, and production for the same plant is set up for all products in the following period.

- Based on inventory level: we identify the plant with the maximum inventory, and ensure it is open in that period. If it was already open, we open it also in the next period. Besides, we close the DC with the lowest inventory level during its open g periods.

- Based on production quantity: we identify for which product, in which plant, and in which period, the maximum (minimum) production occurs; we unassign the production of that product for the plant with the smallest production in that period, and assign the plant and the product with the maximum production in its next period. We also identify the period with the highest production, and open an extra random DC.

- Based on delivery quantity to customers: we identify all DCs and periods with deliveries lower than a percentage of the total demand and close a random DC for that period (and for the next g periods). For plants, we select a random one and close it in the previously identified period.

4.2 Parameter settings and the pseudocode

We have tested different combinations of parameters and tuned them mostly by trial and error. The initial temperature H_{start} is set to $(r+1)\times 100,000$. This initial temperature is cooled down until it reaches the final temperature $H_{final} = 0.01$. The cooling rate, α, also depends on the delivery time window and ranges from 0.9986 ($r = 5$) to 0.9988 ($r = 0$). In our implementation, iteration count is one of our stopping criterion, and it is satisfied once 3,000,000 iterations are performed. We set φ to 1,000 iterations and update the scores with $\sigma_1 = 10$, $\sigma_2 = 4$, and $\sigma_3 = 3$. The pseudocode for the proposed algorithm is presented in Algorithm 4.

5 Computational Experiments

We now describe the details related to the computational experiments used to evaluate our algorithms. All computations are conducted on Intel Core i7 processors running at 3.4 GHz with up to 64 GB of RAM installed, with the Ubuntu Linux operating system. A single thread was used for up to one hour. The algorithms are coded in C++ and we use IBM Concert Technology and CPLEX 12.6.3 as the MIP solver. Section 5.1 describes how the instances are generated, detailed computational results are provided in Section 5.2, and sensitivity and managerial analysis are provided in Section 5.3.

5.1 Generation of the instances

We have generated a large data set by varying the number of products, periods, plants, DCs, and customers. Our test bed is generated as shown in Table 2, for a total of 11 setting combinations and five random instances per combination. The number of plants and DCs are determined by the number of periods: if $T = 5$, then $N_d = 8$ and $N_p = 5$, if $T = 10$, then $N_d = 15$ and $N_p = 10$, and finally if $T = 50$, then $N_d = 25$ and $N_p = 15$. For each instance we assume a delivery time window $r = 0$, 1, 2, and 5 periods. Thus, we solve 220 instances in total.

Algorithm 4 Proposed hybrid matheuristic

Initialize weights to 1, scores to 0, $H \leftarrow H_{start}$.
$s \leftarrow s_{best} \leftarrow$ initial solution.
while stopping criteria are not met **do**
    $s' \leftarrow s$
    Select an operator and apply it to s'
    Solve the remaining flow problem, obtain solution $z(s')$
    if $z(s') < z(s)$ **then**
        if $z(s') < z(s_{best})$ **then**
            $s_{best} \leftarrow s'$
            update the score for the operator used with σ_1
        else
            update the score for the operator used with σ_2
        end if
    else
        if s' is accepted by the simulated annealing criterion **then**
            update the scores for the operator used with σ_3
            $s \leftarrow s'$
        end if
    end if
    $H \leftarrow \alpha \times H$
    if iterations is a multiple of φ **then**
        update weights and reset scores of all operators
        if no improvement found in last 2φ iterations **then**
            if $H < H_{final}$ **then**
                $H \leftarrow H_{start}$
                if no improvement found for $z(s')$ **then**
                    Input s_{best} into the IP and solve it for 20 seconds
                end if
            end if
        else
            $s \leftarrow s_{best}$
        end if
    end if
end while
Return s_{best}

Table 2: Input parameter values

Name	Parameter	Values
Products	P	1, 5, 10
Periods	T	5, 10, 50
Plants	N_p	5, 10, 15
DCs	N_d	8, 15, 25
Customers	N_c	20, 50, 100, 200
Delivery time window	r	0, 1, 2, 5
DC active period	g	$\frac{T}{5}$
Demand	d_{pk}^t	[0, 2]
Plant setup cost	s_{pi}	[10, 15]
Plant variable cost	v_{pi}	[1, 10]
Fixed DC opening cost	f_j	[100, 150]
Inventory holding cost	h_{pj}	[1, 4]
Shipping cost (plants-DC)	c_{pij}	[10, 100]
Shipping cost (DC-customers)	c'_{pjk}	[10, 1000]

5.2 Results of the computational experiments

We now present the results of extensive computational experiments carried out to evaluate the performance of all algorithms, and to draw meaningful conclusions for the problem at hand. We first describe the results of the experiments with the mathematical model proposed in Section 2. This is followed by the comparison of the performance of all three sequential algorithms proposed in Section 3, and our integrated hybrid matheuristic from Section 4 with that of the other algorithms.

Average computational results using the branch-and-bound algorithm are presented in Table 3. For each instance size we report the average of the gaps (G) with respect to the lower bound (calculated as $100 \times (\frac{(Upper\ Bound - Lower\ Bound)}{Upper\ Bound})$, the number of cases solved to optimality (O), and the average running time (T) in seconds for each pre-specified time window r. As can be seen in Table 3, only very small instances could be solved to optimality, mostly those with fewer than five products or periods. The parameter controlling the number of periods seems to have a strong effect on the performance of the exact method. Indeed, it has a huge effect on the size of the problem as measured by the number of variables and constraints. Moreover, the length of the delivery time window also affects the number of instances solved to optimality, the average gap and the running time.

Table 3: Results from exact algorithm

Instances	$r = 0$		$r = 1$		$r = 2$		$r = 5$	
P-T-N_c-N_d-N_p	$G(\%)^{(O)}$	$T(s)$	$G(\%)^{(O)}$	$T(s)$	$G(\%)^{(O)}$	$T(s)$	$G(\%)^{(O)}$	$T(s)$
1-5-20-8-5	$0.00^{(5)}$	2	$0.00^{(5)}$	2	$0.00^{(5)}$	1	$0.00^{(5)}$	0
1-10-100-15-10	$2.33^{(1)}$	3,113	$16.74^{(0)}$	3,606	$20.61^{(0)}$	3,603	$21.60^{(0)}$	3,603
1-10-50-15-10	$0.00^{(5)}$	610	$9.83^{(0)}$	3,616	$15.32^{(0)}$	3,606	$14.09^{(0)}$	3,601
1-50-100-25-15	$23.46^{(0)}$	3,640	$45.80^{(0)}$	3,626	$53.45^{(0)}$	3,624	$70.33^{(0)}$	3,634
5-5-20-8-5	$0.00^{(5)}$	342	$0.13^{(4)}$	1,321	$0.79^{(3)}$	1,703	$0.00^{(5)}$	197
5-10-100-15-10	$12.64^{(0)}$	3,612	$22.11^{(0)}$	3,601	$24.76^{(0)}$	3,601	$28.12^{(0)}$	3,601
5-10-50-15-10	$10.36^{(0)}$	3,610	$18.32^{(0)}$	3,602	$18.62^{(0)}$	3,602	$20.41^{(0)}$	3,602
10-5-20-8-5	$2.04^{(0)}$	3,602	$3.52^{(0)}$	3,604	$4.07^{(0)}$	3,604	$0.00^{(5)}$	1,275
10-10-100-15-10	$12.74^{(0)}$	3,602	$21.10^{(0)}$	3,602	$25.24^{(0)}$	3,610	$25.58^{(0)}$	3,601
10-10-50-15-10	$6.70^{(0)}$	3,605	$17.57^{(0)}$	3,603	$22.85^{(0)}$	3,602	$17.16^{(0)}$	3,611
Average	$7.03^{(0.32)}$	2,574	$15.51^{(0.18)}$	3,018	$18.57^{(0.16)}$	3,056	$28.72^{(0.30)}$	3,610

To evaluate the performance of our sequential hybrid heuristics and gain insight into managerial decisions related to the problem at hand, we present their results in Tables 4–7, one table per value of the delivery time window r. In these tables we first present the results from the Equal power heuristic compared to the lower bound from the exact algorithm, where the first two columns of each table indicate the gaps and the processing time. The improvements with respect to the solution obtained from the exact algorithm by the Top-down, Bottom-up and ALNS heuristics are presented in the subsequent columns.

5.3 Sensitivity analysis and managerial insights

We now perform sensitivity analysis to derive important managerial insights. From Table 3 we observe that the more flexible the delivery time windows become, the harder to solve the problem gets. Also, as the number of products, periods, and customers increase, the problem becomes harder to be solved to optimality. Small instances with $P = 1$, $T = 5$, and $N_c = 20$ are easily solved to optimality, however instances with only one product but $T > 5$ could not be solved to optimality under the presence of the delivery time windows.

Table 4: Heuristics results for $r = 0$

Instances	Equal power[1]		Top-down[2]		Bottom-up[2]		ALNS[2]	
P-T-N_c-N_d-N_p	I (%)	T(s)	I (%)	T(s)	I (%)	T(s)	I (%)	T(s)
1-5-20-8-5	−35.64	1	−10.03	0	−315.27	0	0.00	1
1-10-100-15-10	−65.27	9	−47.86	3	−155.53	2	0.10	3,606
1-10-50-15-10	−62.38	4	−33.05	2	−238.78	2	0.00	2,048
1-50-100-25-15	−60.98	471	−169.26	558	−98.68	245	8.14	3,602
5-5-20-8-5	−54.57	1	−18.42	1	−107.79	0	0.00	1,150
5-10-100-15-10	−53.78	116	−134.78	34	−32.34	13	−1.25	3,602
5-10-50-15-10	−59.04	267	−83.03	12	−62.34	5	−0.74	3,604
10-5-20-8-5	−52.33	3	−33.09	1	−48.09	1	−0.16	3,462
10-10-100-15-10	−47.67	905	−192.67	53	−16.52	38	−3.69	3,605
10-10-50-15-10	−55.67	738	−135.40	51	−25.93	21	−0.66	3,610
Average	−54.73	252	−85.76	72	−110.13	33	0.27	3,382

[1] Improvement over the lower bound; [2] Improvement over the upper bound

Table 5: Heuristics results for $r = 1$

Instances	Equal power[1]		Top-down[2]		Bottom-up[2]		ALNS[2]	
P-T-N_c-N_d-N_p	I (%)	T(s)	I (%)	T(s)	I (%)	T(s)	I (%)	T(s)
1-5-20-8-5	−41.38	0	−19.34	0	−423.38	0	0.00	3
1-10-100-15-10	−58.69	69	−69.06	17	−184.87	1	2.21	3,608
1-10-50-15-10	−57.62	19	−48.79	8	−291.73	3	0.71	3,474
1-50-100-25-15	−53.48	2,037	−194.29	1,391	−117.97	313	27.03	3,601
5-5-20-8-5	−51.99	3	−39.73	2	−149.61	0	0.00	2,627
5-10-100-15-10	−45.88	1,322	−178.31	1,012	−49.76	16	2.75	3,419
5-10-50-15-10	−51.59	1,191	−124.92	289	−87.15	6	4.10	3,604
10-5-20-8-5	−46.35	19	−56.30	3	−86.45	2	0.08	3,608
10-10-100-15-10	−39.18	1,408	−240.47	1,292	−39.79	38	1.58	3,601
10-10-50-15-10	−46.25	1,298	−169.62	1,054	−39.78	25	3.27	3,604
Average	−49.24	737	−114.08	507	−147.05	40	5.95	3,559

[1] Improvement over the lower bound; [2] Improvement over the upper bound

Table 6: Heuristics results for $r = 2$

Instances	Equal power[1]		Top-down[2]		Bottom-up[2]		ALNS[2]	
P-T-N_c-N_d-N_p	I (%)	T(s)	I (%)	T(s)	I (%)	T(s)	I (%)	T(s)
1-5-20-8-5	−45.20	1	−20.41	0	−459.31	0	0.00	4
1-10-100-15-10	−56.92	118	−88.23	20	−179.76	3	2.87	3,608
1-10-50-15-10	−57.01	30	−63.64	9	−267.67	2	0.57	3,371
1-50-100-25-15	−45.89	2,195	−241.74	1,549	−152.76	317	31.44	3,601
5-5-20-8-5	−50.55	3	−46.06	2	−177.29	1	0.04	3,149
5-10-100-15-10	−41.17	1,326	−203.93	1,244	−66.24	24	1.80	3,602
5-10-50-15-10	−48.32	1,243	−144.53	431	−100.53	10	4.87	3,603
10-5-20-8-5	−43.65	5	−75.38	3	−105.40	6	0.75	3,606
10-10-100-15-10	−33.53	1,427	−268.28	1,291	−38.92	66	4.74	3,601
10-10-50-15-10	−39.81	1,306	−202.70	1,034	−57.80	24	5.86	3,601
Average	−46.21	765	−135.49	558	−160.57	45	7.45	3,570

[1] Improvement over the lower bound; [2] Improvement over the upper bound

Table 7: Heuristics results for $r = 5$

Instances	Equal power[1]		Top-down[2]		Bottom-up[2]		ALNS[2]	
P-T-N_c-N_d-N_p	I (%)	T(s)	I (%)	T(s)	I (%)	T(s)	I (%)	T(s)
1-5-20-8-5	−51.28	0	−41.46	0	−514.97	0	0.00	2
1-10-100-15-10	−53.00	23	−121.87	19	−198.35	4	4.65	3,607
1-10-50-15-10	−55.42	11	−100.72	6	−255.19	3	1.13	3,608
1-50-100-25-15	−35.55	2,635	−298.84	1,769	−203.54	412	49.66	3,601
5-5-20-8-5	−46.81	1	−76.86	0	−199.40	1	0.00	3,205
5-10-100-15-10	−34.66	1,449	−254.33	470	−88.65	30	3.47	3,601
5-10-50-15-10	−41.92	1,097	−203.82	104	−135.73	12	9.31	3,601
10-5-20-8-5	−37.90	3	−113.14	1	−126.80	1	0.00	3,605
10-10-100-15-10	−31.10	1,594	−286.26	1,151	−45.14	98	0.89	3,601
10-10-50-15-10	−35.16	1,205	−238.52	459	−80.05	43	7.14	3,601
Average	−42.28	802	−173.58	398	−184.78	60	7.62	3,203

[1] Improvement over the lower bound; [2] Improvement over the upper bound

This difficulty in solving the problem when delivery time windows exist shows two interesting aspects of the business problem. The first one is related to the potential cost saving if one is to properly exploit the added flexibility of time windows. This is evident since all solutions without time windows are still valid to the cases in which they are considered. However, in order to take advantage of such flexibility, using a tailored method seems necessary as we have shown that simply modeling the problem into a commercial solver or using any sequential method do not yield any good solutions, and in fact their quality degrades as the size of the problem and the added flexibility increases.

The results presented in Tables 4–7 reveal that for large instances our ALNS-based heuristic outperforms the exact algorithm. The highest average improvement (7.45%) is obtained for $r = 5$. For all the small instances that could be solved to optimality, our approach also obtains the optimal solution. Regarding the processing time, the exact algorithm is slightly better, which is mainly due to the fact that the iterative heuristic reaches the time limit to search the solution area aiming to improve the solution obtained and is truncated at 3600 seconds. However, as can be seen in Table 8, our algorithm takes on average less than 20 minutes to find its best solution, which is often better than the ones from the exact algorithm.

Table 8: Average time for ALNS to obtain the best solution

Time window	$r = 0$	$r = 1$	$r = 2$	$r = 5$	Average
Average time (S)	1,160	1,159	1,022	909	1,063

Regarding our integrated method, Tables 4–7 also show that taking an integrative approach toward production, inventory, and distribution can lead to enormous reduction in costs. For all time-windows, the average results obtained from the integrated approach is at least 86% better than the hierarchical ones. It is interesting to note that on average the costs obtained by Top-down heuristic are lower than the ones from the Bottom-up approach; however, the Bottom-up heuristic is much faster. As presented in Tables 4–7, the Bottom-up heuristic generates better results in less time than the Top-down when $P > 1$ and $N_c > 20$.

As expected, applying the Equal power heuristic, or the management in silos, where each department of the company is focused only on its own decisions regardless of the others, results in not even one instance with a feasible solution. The sum of cost at plants including setup, variable, and inventory costs, costs at DCs including, rental fees, deliveries from plants to DCs, and the inventory holding costs and shipment to customer costs are obtained. The calculated cost is then compared to the lower bound obtained from the exact algorithm. On average this infeasible cost from the Equal power heuristic is 48.11% worse than the lower bound, which forgoes any hopes that this approach would yield a good solution.

Finally, the exact method could not obtain any solution for our large instances with $P = 1$, $T = 50$, $N_c = 200$, $N_d = 25$ and $N_p = 15$. We ran these large instances using both our proposed heuristic in order to obtain valid and good solutions, and the Equal power heuristic to obtain a lower bound. The differences between the lower bound of the Equal power and the best solutions of the proposed ALNS algorithm are reported in Table 9. From results presented on Tables 4–7 we already know that as the instance becomes larger, the gap between the lower bound obtained from Equal power heuristic and the lower bound from the exact algorithm decreases. The same instance with half the number of customer ($N_c = 100$) has the average of 48.98% gap over all 4 time windows. Therefore, for the larger instance with $N_c = 200$, one could extrapolate and estimate an expected gap lower than 48.98%. Knowing that the average gap between our proposed approach and the lower bound from the Equal power method is 60.71%, one could also anticipate an approximate average gap of 11.73% between our proposed approach and a lower bound from the exact method if it could be achieved.

Table 9: Gaps for the large instance with $N_c = 200$

Time-window	$r = 0$	$r = 1$	$r = 2$	$r = 5$	Average
Average gap (%)	64.01	62.32	60.23	56.26	60.71

6 Conclusions

This paper investigates a variant of integrated LSP, in which multiple products are produced over a discrete time horizon, stored and shipped to final customers. The paper contributes to the integrated optimization literature as it combines distinct features of delivery time windows, distribution with direct shipment, and location decisions from potential DCs to store products for a fixed leasing period. A state of the art commercial solver is able to find optimum solutions for small size instances of our problem, however, it does not prove optimality in reasonable running time for larger instances. To achieve better solutions in acceptable computation time, we have developed an ALNS heuristic. Several instances are generated and the solutions are compared to the optimal ones (if any) obtained by the exact method on average improving the solutions of the exact method by on average 5.32% and up to 49.66%, generally in only third of the running time. We have also evaluated how a typical management in silos would perform, by deriving and implementing sequential solution methods. Our results also confirm the cost benefits of the integrated approach toward decision making comparing to the hierarchical ones. A Top-down method (production first and distribution last) performs on average 132.55% worse than the integrated approach, while a Bottom-up method (distribution first, production last) performs on average 155.95% worse than our method. Between these two methods, the Bottom-up works better for instances with larger planning horizons and more products and customers, while Top-down is preferred when there is only one product and fewer than 20 customers.

We have undoubtedly proved the benefits of an integrated management and solution method, as opposed to hierarchical and/or silo management. Moreover, we have shown that for complex and rich integrated problems such as this variant of the lot sizing observed in practice, neither a hierarchical solution approach nor modeling and solving it by a commercial solver yield good solutions in reasonable time. We have developed a flexible and very powerful method capable of effectively handling all aspects of the problem in an efficient manner.

References

[1] Y. Acar and S. N. Atadeniz. Comparison of integrated and local planning approaches for the supply network of a globally-dispersed enterprise. *International Journal of Production Economics*, 167:204–219, 2015.

[2] Y. Adulyasak, J.-F. Cordeau, and R. Jans. The production routing problem: A review of formulations and solution algorithms. *Computers & Operations Research*, 55:141–152, 2015.

[3] A. Akbalik and B. Penz. Comparison of just-in-time and time window delivery policies for a single-item capacitated lot sizing problem. *International Journal of Production Research*, 49(9):2567–2585, 2011.

[4] P. Amorim, H-O. Günther, and B. Almada-Lobo. Multi-objective integrated production and distribution planning of perishable products. *International Journal of Production Economics*, 138(1):89–101, 2012.

[5] S. Axsäter. Multi-echelon systems: Lot sizing. In *Inventory Control*, pages 171–189. Springer, Cham, 2015.

[6] N. Brahimi, S. Dauzère-Pérès, and N. M. Najid. Capacitated multi-item lot-sizing problems with time windows. *Operations Research*, 54(5):951–967, 2006.

[7] L. Buschkühl, F. Sahling, S. Helber, and H. Tempelmeier. Dynamic capacitated lot-sizing problems: a classification and review of solution approaches. *OR Spectrum*, 32(2):231–261, 2010.

[8] J-F Camacho-Vallejo, R. Muñoz-Sánchez, and J. L. González-Velarde. A heuristic algorithm for a supply chain' s production-distribution planning. *Computers & Operations Research*, 61:110–121, 2015.

[9] L. C. Coelho and G. Laporte. Improved solutions for inventory-routing problems through valid inequalities and input ordering. *International Journal of Production Economics*, 155: 391– 397, 2014.

[10] L. C. Coelho, J-F. Cordeau, and G. Laporte. Thirty years of inventory routing. *Transportation Science*, 48(1):1–19, 2014.

[11] J. O. Cunha and Rafael A. Melo. On reformulations for the one-warehouse multi-retailer problem. *Annals of Operations Research*, 238(1-2):99–122, 2016.

[12] M. Darvish, H. Larrain, and L. C. Coelho. A dynamic multi-plant lot-sizing and distribution problem. *International Journal of Production Research*, forthcoming, 2016.

[13] C. Dhaenens-Flipo and G. Finke. An integrated model for an industrial production–distribution problem. *IIE Transactions*, 33(9):705–715, 2001.

[14] A. V. Goldberg. An efficient implementation of a scaling minimum-cost flow algorithm. *Journal of Algorithms*, 22(1):1–29, 1997.

[15] F. Hein and C. Almeder. Quantitative insights into the integrated supply vehicle routing and production planning problem. *International Journal of Production Economics*, 177: 66–76, 2016.

[16] H-C. Hwang, W. Jaruphongsa, S. Çetinkaya, and C-Y. Lee. Capacitated dynamic lot-sizing problem with delivery/production time windows. *Operations Research Letters*, 38 (5):408–413, 2010.

[17] H-C. Hwang, H-S. Ahn, and P. Kaminsky. Algorithms for the two-stage production-capacitated lot-sizing problem. *Journal of Global Optimization*, 65(4):777–799, 2016.

[18] R. Jans and Z. Degraeve. Meta-heuristics for dynamic lot sizing: A review and comparison of solution approaches. *European Journal of Operational Research*, 177(3):1855–1875, 2007.

[19] R. Jans and Z. Degraeve. Modeling industrial lot sizing problems: a review. *International Journal of Production Research*, 46(6):1619–1643, 2008.

[20] W. Jaruphongsa, S. Çetinkaya, and C-Y. Lee. A two-echelon inventory optimization model with demand time window considerations. *Journal of Global Optimization*, 30(4):347–366, 2004.

[21] W. Jaruphongsa, S. Çetinkaya, and C-Y. Lee. Warehouse space capacity and delivery time window considerations in dynamic lot-sizing for a simple supply chain. *International Journal of Production Economics*, 92(2):169–180, 2004.

[22] B. Karimi, S. M. T. F. Ghomi, and J. M. Wilson. The capacitated lot sizing problem: a review of models and algorithms. *Omega*, 31(5):365–378, 2003.

[23] B. B. Keskin and H. Üster. Production/distribution system design with inventory considerations. *Naval Research Logistics*, 59(2):172–195, 2012.

[24] C-Y. Lee, S. Çetinkaya, and A. P. M. Wagelmans. A dynamic lot-sizing model with demand time windows. *Management Science*, 47(10):1384–1395, 2001.

[25] Z. Liang, Y. He, T. Wu, and C. Zhang. An informative column generation and decomposition method for a production planning and facility location problem. *International Journal of Production Economics*, 170:88–96, 2015.

[26] M. T. Melo, S. Nickel, and F. Saldanha da Gama. Facility location and supply chain management–a review. *European Journal of Operational Research*, 196(2):401–412, 2009.

[27] R .A. Melo and L. A. Wolsey. Uncapacitated two-level lot-sizing. *Operations Research Letters*, 38(4):241–245, 2010.

[28] R. A. Melo and L. A. Wolsey. MIP formulations and heuristics for two-level production-transportation problems. *Computers & Operations Research*, 39(11):2776–2786, 2012.

[29] Y. Merzifonluoğlu and J. Geunes. Uncapacitated production and location planning models with demand fulfillment flexibility. *International Journal of Production Economics*, 102(2): 199–216, 2006.

[30] M. C. V. Nascimento, M. G. C. Resende, and F. M. B. Toledo. GRASP heuristic with path-relinking for the multi-plant capacitated lot sizing problem. *European Journal of Operational Research*, 200(3):747–754, 2010.

[31] A. M. Nezhad, H. Manzour, and S. Salhi. Lagrangian relaxation heuristics for the uncapacitated single-source multi-product facility location problem. *International Journal of Production Economics*, 145(2):713–723, 2013.

[32] P. L. Robinson, A. Narayanan, and F. Sahin. Coordinated deterministic dynamic demand lot-sizing problem: A review of models and algorithms. *Omega*, 37(1):3–15, 2009.

[33] S. Ropke and D. Pisinger. An adaptive large neighborghood search heuristic for the pickup and delivery problem with time windows. *Transportation Science*, 40(4):455–472, 2006.

[34] M. Sambasivan and C. P. Schmidt. A heuristic procedure for solving multi-plant, mutli-item, mutli-period capacitated lot-sizing problems. *Asia-Pacific Journal of Operational Research*, 19(1):87, 2002.

[35] M. Sambasivan and S. Yahya. A Lagrangean-based heuristic for multi-plant, multi-item, multi-period capacitated lot-sizing problems with inter-plant transfers. *Computers & Operations Research*, 32(3):537–555, 2005.

[36] O. Sanei Bajgiran, M. Kazemi Zanjani, and M. Nourelfath. The value of integrated tactical planning optimization in the lumber supply chain. *International Journal of Production Economics*, 171:22–33, 2016.

[37] O. Solyalı, M. Denizel, and H. Süral. Effective network formulations for lot sizing with backlogging in two-level serial supply chains. *IIE Transactions*, 48(2):146–157, 2016.

[38] M. G. Speranza. Trends in transportation and logistics. *European Journal of Operational Research*, forthcoming, 2016.

[39] M. Steinrücke. Integrated production, distribution and scheduling in the aluminium industry: a continuous-time MILP model and decomposition method. *International Journal of Production Research*, 53(19):5912–5930, 2015.

[40] T. Vogel, B. Almada-Lobo, and C. Almeder. Integrated versus hierarchical approach to aggregate production planning and master production scheduling. *OR Spectrum*, forthcoming, 2016.

[41] H. M. Wagner and T. M. Whitin. Dynamic version of the economic lot size model. *Management Science*, 5(1):89–96, 1958.

[42] W. Wei, L. Guimarães, P. Amorim, and B. Almada-Lobo. Tactical production and distribution planning with dependency issues on the production process. *Omega*, forthcoming, 2016.

[43] L. A. Wolsey. Lot-sizing with production and delivery time windows. *Mathematical Programming*, 107(3):471–489, 2006.