



A matheuristic for a multi-period three-echelon network design problem with temporary capacity acquisition

Isabel Correia^a, Teresa Melo^{b,*}

^a Center for Mathematics and Applications (NovaMath) and Department of Mathematics, NOVA School of Science and Technology (NOVA FCT), P 2829-516 Caparica, Portugal

^b Business School, Saarland University of Applied Sciences, D 66123 Saarbrücken, Germany

ARTICLE INFO

Keywords:

Network design
Temporary warehouse rental
Capacity expansion and reduction
Mixed integer programming
Matheuristic

ABSTRACT

This study addresses a three-echelon network design problem that determines the location and size of new warehouses, the removal of company-owned warehouses, the inventory levels of multiple products at the warehouses, and the assignment of suppliers as well as customers to warehouses over a multi-period planning horizon. A distinctive feature of our problem is that new warehouses operate with modular capacities that can be expanded or reduced over several periods, the latter not necessarily having to be consecutive. Moreover, in every period, the demand of a customer for a given product has to be satisfied by a single warehouse. The problem arises in the context of warehousing-as-a-service, a business scheme that offers flexible conditions for temporary capacity leasing. The associated fixed warehouse lease cost reflects economies of scale in the capacity size and the length of the lease contract. We develop a mixed-integer linear programming formulation and propose a matheuristic to solve this problem, which exploits the structure of the optimal solution of the linear relaxation to successively assign customers to open warehouses and fix other binary variables related to warehouse operation. Additional variable fixing rules are also developed, based on a scheme for managing inventories at warehouses and using the product quantities provided by suppliers. Numerical experiments with randomly generated large-sized instances reveal that the proposed matheuristic outperforms a general-purpose solver in 74% of the instances by identifying higher quality solutions in a substantially shorter computing time.

1. Introduction

The location of facilities such as manufacturing plants, distribution centers and retail outlets is one of the most critical decisions to make when designing a logistics network. Typically, these decisions are strategic in nature, have a long-term impact, and strongly affect other decisions made at the tactical and operational levels. Location decisions are therefore unlikely to be reversible in the short or medium term (Cordeau et al., 2021; Melo et al., 2009). This is the case when facilities are built or purchased, as these actions require a sizeable financial investment. On the other hand, if facilities are rented or operations are subcontracted, location decisions may be reversible in the medium term. Warehousing-as-a-service is an emerging trend that casts location decisions as a tactical planning problem (Costa & Melo, 2024). This new business model offers flexible, short-term storage space as well as additional logistics services (e.g., goods handling) (Ceschia et al., 2023; Lahy et al., 2022). Thus, for companies needing extra storage space on a temporary basis to deal with seasonal peaks

and variations in demand (e.g., omnichannel retailers), warehousing-as-a-service provides an opportunity to improve their efficiency and create competitive advantage. This business approach contrasts with the traditional model that entails long-term contracts, minimum volume commitments, and rigid pricing schemes to access warehouse capacity.

In this paper, we incorporate this new trend into the redesign of a three-echelon logistics network over a multi-period planning horizon. For this purpose, we assume that a company operates a number of warehouses at fixed locations and that there is a set of potential warehouses where additional storage capacity can be temporarily leased. Multiple (non-overlapping) leases can be taken out on the same warehouse for different sizes of capacity, allowing for expansion and reduction of storage space as needed. In addition, a lease contract on a warehouse can be terminated and subsequently resumed with the same or a different capacity. As shown in Fig. 1, a set of suppliers provide various products that are consolidated in the (two types of) warehouses before being distributed to a set of customer zones to meet their demand.

* Corresponding author.

E-mail addresses: isc@fct.unl.pt (I. Correia), teresa.melo@htwsaar.de (T. Melo).

<https://doi.org/10.1016/j.cie.2024.110244>

Received 31 October 2023; Received in revised form 27 March 2024; Accepted 18 May 2024

Available online 23 May 2024

0360-8352/© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

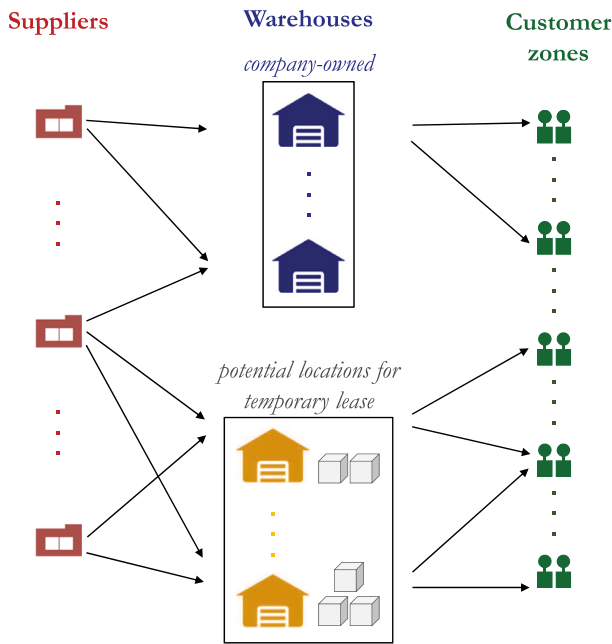


Fig. 1. Three-echelon network structure.

The aim is to specify a minimum cost schedule for the deployment of warehouses over the planning horizon. This includes discontinuing the operation of company-owned warehouses, leasing storage space at potential warehouses, and deciding how much capacity to lease in those facilities and for how long. Economies of scale are present in the fixed cost of a lease agreement, taking into account the amount of capacity leased and the term length. A fixed cost is also incurred when the operation of a company-owned warehouse is terminated. In this case, the significant capital and time investment required to purchase a new facility or negotiate a new partnership with a logistics service provider precludes the reopening of the warehouse at a later date. In addition, procurement, inventory, and distribution decisions for multiple product families also need to be made in each time period.

The main contributions of this study are threefold. First, we model a three-echelon network redesign problem in the context of tactical location planning that extends a particular case recently considered by Correia and Melo (2022). In contrast to the latter work, location decisions are combined with capacity sizing decisions for adding and removing modular units to/from leased warehouses. Furthermore, single-assignment conditions are enforced so that each customer zone is served by exactly one warehouse for a given product family in each period. As our numerical study will show, these settings result in a challenging mixed-integer linear programming (MIP) formulation for which the possibility of solving large-scale instances optimally with an off-the-shelf optimization solver in acceptable time is rather limited. Hence, the second contribution is to propose an effective matheuristic, especially for large-sized problem instances resulting from an extended planning horizon and a significant number of customers. A feasible solution is constructed using an iterative scheme that jointly exploits the information provided by the linear (LP) relaxation and the capacity utilization level of the operating warehouses. This knowledge is then used to successively fix a subset of the binary variables. Additional mechanisms are developed to fix further binary variables, taking into account the residual capacities of suppliers and the amount of inventory available in warehouses as a result of decisions taken earlier. Variable fixing yields an MIP problem of smaller size which is then solved exactly using a general-purpose solver within a pre-defined time limit. The third contribution of our study is to demonstrate the superior

quality and computational efficiency of the proposed matheuristic using randomly generated problem instances, the larger of which has 48 time periods, 400 customer zones, and more than 1.2 million binary variables.

The remainder of this paper is organized as follows. Section 2 gives an overview of the relevant literature. Section 3 describes the problem and presents an MIP formulation. The proposed matheuristic is detailed in Section 4, while Section 5 reports the results of our computational study. Finally, in Section 6, conclusions and directions for future research are provided.

2. Literature review

Facility location decisions are at the core of logistics network design problems (Cordeau et al., 2021; Melo et al., 2009). They typically involve determining the optimal number and location of facilities (e.g., manufacturing plants, warehouses) among a set of candidate sites. Naturally, these decisions are intertwined with multiple operations within the logistics network and must therefore be taken jointly with other decisions, such as allocating capacity to the selected facilities and setting the flow of products (e.g., procurement, production, distribution) across the network so that customer demand is satisfied at minimum cost or maximum profit. Regarding the planning horizon, logistics network design problems can be viewed from a static perspective (single period) or a dynamic perspective (multi-period). While the former may be appropriate when parameters (e.g., cost drivers, demand) are relatively stable, the latter allows for estimated parameter variations to be considered at the planning stage (Nickel & da Gama, 2019). In this section, we review the relevant literature, with a focus on multi-period capacitated network design problems, thereby highlighting how our study differs from previous work.

2.1. Multi-period capacitated network design

In a multi-period setting, gradual changes in the network configuration are, among others, triggered by capacity decisions. Not only is it necessary to determine the amount of capacity to be deployed in the new facilities when these are opened, but also to decide whether that capacity needs to be adjusted over time to respond to variations in costs and demand. A common approach is to assume a finite set of capacity levels at each potential location, with their acquisition costs reflecting economies of scale (Correia & Melo, 2016; Cortinhal et al., 2015; Thanh et al., 2008). In this case, the decision to open a facility is coupled with the selection of a single capacity level among the available options. Another approach is to assume that capacity is represented by modular units. A modular unit may be related to a specific type of equipment or qualified staff (Correia & Captivo, 2006). At each potential location, multiple units of the same module type can be installed provided that the sum of their capacities does not exceed a pre-specified limit. Again, the cost of deploying and operating a given number of modules at a specific location is subject to economies of scale. As shown in the recent survey by Alarcon-Gerbier and Buscher (2022), this form of capacity sizing is gaining increasing attention, with emerging applications in a variety of industries such as forestry, chemicals, healthcare, telecommunications, and military logistics, among others.

In addition to capacity acquisition via the opening of facilities, other strategies may be adopted to scale the capacity available on the network over time. This is the case when existing facilities can also be closed, resulting in the reconfiguration of a logistics network already in place. The network redesign problems studied by Correia and Melo (2016), Cortinhal et al. (2015), and Delmelle et al. (2014) are examples of this strategy, which also allows for gradually expanding the network through opening additional facilities with specific capacity levels over the planning horizon. Another way of modeling capacity adjustments is to temporarily close a facility and reopen it at a later time. Motivated

by an application in the forestry sector, Jena et al. (2015b, 2016) investigate this problem variant. A further strategy for resizing capacity is to move capacity between different facilities in order to be in close proximity to customer markets or suppliers (Allman & Zhang, 2020; Becker et al., 2019; Melo et al., 2012).

The possibility of expanding and reducing the capacity of facilities during their operating term is an approach that has received less attention, but which increases the flexibility of the logistics network to respond to demand shifts. An early contribution addressing this case was made by Antunes and Peeters (2001). In this study, the capacity of a facility can either be expanded or reduced, but not both, over the planning horizon. By contrast, Wilhelm et al. (2013) allow the configuration of a facility to be changed more than once over time through adding or removing a capacity level. A similar approach is adopted by Correia and Melo (2017, 2021) for a problem variant with customers segmented into two groups that differ in the delivery time they accept for meeting their orders. Jena et al. (2015a, 2017) propose a general model to represent different types of capacity changes. Their framework is based on tracking the number of modular units a facility holds in two consecutive time periods. If these two values differ, then a resizing of the capacity has occurred. This allows modeling not only the expansion and reduction of the capacity of a facility, but also its temporary removal and subsequent reopening. These features are also captured in our study, but in a different context, where decisions on facility location and capacity sizing are tactical rather than strategic in nature. As a result, the length of the planning horizon we consider in our numerical study is significantly longer than the one selected by Jena et al. (2015a, 2017), which greatly increases the size of our problem instances. Moreover, the logistics network in our work includes an additional layer (recall the supplier layer in Fig. 1) and customer single-assignment conditions are imposed, thereby further adding to the complexity of the problem.

In the context of location decisions being reversible in the medium or even short term (e.g., quarterly or monthly), the literature is considerably scarcer. An early contribution was made by Thanh et al. (2008) to redesign a four-echelon production and distribution network. Manufacturing plants and company-owned warehouses that are in place at the beginning of the planning horizon can be closed or have their capacity extended over time, but not both. In addition, public warehouses can be rented for periods of time that need not be consecutive. However, a minimum interval of two periods is enforced between two rental contracts at the same public warehouse, which is a rather restrictive condition that we do not consider in our problem. The four-echelon network design problem studied by Badri et al. (2013) and Bashiri et al. (2012) also includes private and public warehouses, but no fixed costs for renting storage space at public facilities are incurred. As a result, economies of scale are not present in the cost of acquiring capacity from these facilities. In the three-echelon network design problem addressed by Darvish and Coelho (2018), distribution centers can be rented with unlimited capacity for a pre-specified number of time periods. By contrast, in the problem we study, the lease term in a selected facility is an endogenous decision, as is the amount of capacity to be deployed, with the later being scalable over time. Unnu and Pazour (2022) also assume that when a so-called on-demand warehouse is rented with a certain capacity level, this facility remains open for a pre-defined number of time periods. In addition, the warehouse rental cost does not take into account economies of scale, but is determined by a fixed rate multiplied by the length of the lease. Recently, Clavijo López et al. (2024) also considered a pre-defined contract duration to operate potential cross-docking facilities owned by different carriers. Successive contracts can be in place with a given carrier, thus allowing the reopening of cross-docking facilities. Moreover, sizing the capacity of these facilities is also addressed by selecting a capacity level from a set of available options.

The problem investigated in this paper draws on previous work by Correia and Melo (2022), which was the first to integrate the temporary rental of a subset of facilities into a three-echelon network redesign

problem and where the duration of lease contracts is an endogenous decision. As will be detailed in the next section, the problem studied here captures additional features, namely it is possible to expand and reduce capacity in facilities that are leased on a temporary basis and single-assignment conditions to meet customer demand for a particular product family are enforced. It should be noted that the latter feature has not been widely studied in the literature. In addition, Correia and Melo (2022) use a general-purpose optimization solver, whereas we propose a specially-tailored matheuristic.

2.2. Solution approaches

Multi-period capacitated network design problems belong to the class of $\mathcal{NP}$ -hard optimization problems as they are generalizations of the uncapacitated facility location problem which is a well-known $\mathcal{NP}$ -hard problem (Krarup & Pruzan, 1983). Incorporating decisions beyond the location of facilities into the planning of a logistics network makes this type of problems even more challenging in terms of identifying high-quality solutions within reasonable computing time using off-the-shelf optimization solvers (Bashiri et al., 2012; Correia & Melo, 2016, 2017; Cortinhal et al., 2015; Delmelle et al., 2014; Jena et al., 2015a, 2015b; Thanh et al., 2008; Unnu & Pazour, 2022; Wilhelm et al., 2013). For this reason, a great deal of effort has been devoted to developing optimization techniques (e.g., MIP-based decomposition methods, meta- and matheuristics) capable of providing good feasible solutions to the problem of interest within acceptable time, particularly for large-scale instances. While some exact solution approaches such as branch-and-price (Allman & Zhang, 2020) and Benders Decomposition (Clavijo López et al., 2024) have been proposed for this class of location problems, most contributions resort to heuristic strategies based on, for example, Lagrangean relaxation (Badri et al., 2013; Jena et al., 2016, 2017), local search (Sauvey et al., 2020), simulated annealing (Antunes & Peeters, 2001), tabu search (Melo et al., 2012), evolutionary algorithms (Silva et al., 2021), and DC programming (Thanh et al., 2012), among others.

Matheuristics have become a popular family of successful solution approaches for handling many optimization problems (Boschetti et al., 2023), including facility location problems. For example, Darvish and Coelho (2018) combine variable neighborhood search with an exact method to solve a multi-period network design problem. Another way to construct feasible solutions is to use the information gained from the optimal solution to the LP relaxation of the original integer problem. In this case, an (iterative) rounding procedure is typically applied to the fractional variables, together with a mechanism that repairs infeasibilities (Silva et al., 2021; Thanh et al., 2010). The matheuristic that we propose in this paper also draws on information provided by the LP relaxation, but unlike the aforementioned works, it does not employ a rounding procedure. Instead, and since some binary variables take integer values in the LP relaxation (the details will be presented in Section 4), specific rules are developed to decide which variables should be fixed in the original problem. This approach is combined with additional schemes to fix further binary variables. To the best of our knowledge, our study is the first to present a matheuristic for a network redesign problem arising at the tactical level, involving location and capacity scalability decisions over a multi-period planning horizon.

3. Problem description and mathematical formulation

The problem investigated in this study is characterized as follows:

- A time horizon for planning, divided into a finite number of periods.
- A set of suppliers offering several product families and having varying capacity over the planning horizon.

- A set of capacitated warehouses belonging to a company or rented from a logistics service provider, and operating at the beginning of the planning horizon. Without loss of generality, these facilities are referred to as *company-owned warehouses*.
- A set of potential warehouses where capacity can be leased for a variable timeframe. The storage space at each location is set up by the total number of modular units that are leased, which cannot be greater than a pre-specified number. For the sake of simplicity, a single type of modular unit with a fixed size is considered at all potential locations.
- Rental of space in a potential warehouse location for multiple periods with varying capacity. Thus, over the planning horizon, the number of modules leased can be extended or reduced at the same facility. Furthermore, the lease periods need not be consecutive. Consequently, multiple lease agreements may be established at the same location, as long as they do not overlap.
- A set of customer zones with known demand for multiple product families in each period.
- Single-assignment conditions for meeting the demand of a customer zone for a given product family from a single warehouse in each time period.
- Inventories in both company-owned and leased warehouses that can be carried over from one period to the next for the duration of the warehouse's operation and which incur variable holding costs. Without loss of generality, initial and ending inventories are assumed to be zero in any storage location.
- Fixed lease costs at new warehouse locations reflecting economies of scale with respect to volume and commitment duration.
- Fixed costs for operating and discontinuing the activity of company-owned warehouses.
- Variable costs for distributing products from suppliers to warehouses and from the latter to customer zones.

The problem considers the distribution of multiple product families from suppliers to customer zones via company-owned and leased warehouses to meet demand requirements. This involves defining a schedule for discontinuing the activity of the company-owned warehouses and a schedule for leasing storage space in new locations, as well as sizing their capacity, which may be temporarily expanded or reduced. The objective is to minimize the total cost, which includes fixed costs for leasing new storage space, fixed costs for operating and removing company-owned warehouses, and variable distribution and inventory holding costs.

Next, the notation is introduced and then the mathematical formulation is presented.

Sets

T	Set of discrete time periods in the planning horizon, with $ T $ being the last period.
S	Set of suppliers.
I^e	Set of company-owned warehouses that are operating at the beginning of the planning horizon.
I^n	Set of potential warehouses where storage space can be leased on a temporary basis.
I	Set of all warehouse locations, $I = I^e \cup I^n$ and $I^e \cap I^n = \emptyset$.
J	Set of customer zones.
P	Set of product families.

Parameters related to suppliers, warehouses, and customer zones

SC_{sp}^t	Capacity of supplier $s \in S$ for product family $p \in P$ at time period $t \in T$.
Q_i	Storage capacity of company-owned warehouse $i \in I^e$.
β_p	Amount of capacity used by one unit of product family $p \in P$ stored at warehouse $i \in I$.

u_{ip}^0	Amount of product family $p \in P$ held in stock in warehouse $i \in I$ at the beginning of the planning horizon. Without loss of generality, $u_{ip}^0 = 0$ for all $i \in I$ and $p \in P$.
d_{jp}^t	Demand of customer zone $j \in J$ for product family $p \in P$ at time period $t \in T$.
M_i	Capacity of one modular unit at warehouse location $i \in I^n$.
n_i	Maximum number of modular units that can be leased in warehouse $i \in I^n$.

Cost parameters

OC_i^t	Fixed cost of operating the company-owned warehouse $i \in I^e$ at time period $t \in T$.
FC_i^t	Fixed cost of discontinuing the operation of the company-owned warehouse $i \in I^e$ at the <i>beginning</i> of time period $t \in T$.
$L_{ik}^{t,t'}$	Total fixed cost of leasing warehouse $i \in I^n$ with $k \in \{1, \dots, n_i\}$ modular units from the beginning of time period t until the end of time period t' ($t, t' \in T; t' \geq t$).
SC_{sip}^t	Cost of distributing one unit of product family $p \in P$ from supplier $s \in S$ to warehouse $i \in I$ at time period $t \in T$.
DC_{ijp}^t	Cost of distributing one unit of product family $p \in P$ from warehouse $i \in I$ to customer zone $j \in J$ at time period $t \in T$.
IC_{ip}^t	Cost of storing one unit of product family $p \in P$ in warehouse $i \in I$ at the end of time period $t \in T$.

The decisions that need to be made are represented by three sets of binary variables and two sets of (non-negative) continuous variables.

Binary variables

y_i^t	=1 if the company-owned warehouse i ceases to be operated at the beginning of time period t , and 0 otherwise ($i \in I^e, t \in T$).
$z_{ik}^{t,t'}$	=1 if the new warehouse i is leased with k modular units from the beginning of time period t until the end of time period t' , and 0 otherwise ($i \in I^n, t, t' \in T, t' \geq t, k = 1, \dots, n_i$).
x_{ijp}^t	=1 if the demand of customer zone j for product family p is served by warehouse i at time period t , 0 otherwise ($i \in I, j \in J, p \in P, t \in T$).

Continuous variables

v_{sip}^t	Quantity of product family p distributed from supplier s to warehouse i at time period t ($s \in S, i \in I, p \in P, t \in T$).
u_{ip}^t	Quantity of product family p held in stock in warehouse i at the end of time period t ($i \in I, p \in P, t \in T$).

The MIP formulation is as follows:

$$\begin{aligned}
 \text{Min} \quad & \sum_{i \in I^e} \sum_{t \in T} OC_i^t \left(1 - \sum_{t'=1}^t y_i^{t'} \right) + \sum_{i \in I^e} \sum_{t \in T} FC_i^t y_i^t \\
 & + \sum_{i \in I^n} \sum_{k=1}^{n_i} \sum_{t \in T} \sum_{t'=t}^{|T|} L_{ik}^{t,t'} z_{ik}^{t,t'} + \\
 & \sum_{s \in S} \sum_{i \in I} \sum_{p \in P} \sum_{t \in T} SC_{sip}^t v_{sip}^t + \sum_{i \in I} \sum_{j \in J} \sum_{p \in P} \sum_{t \in T} DC_{ijp}^t d_{jp}^t x_{ijp}^t + \\
 & \sum_{i \in I} \sum_{p \in P} \sum_{t \in T} IC_{ip}^t u_{ip}^t
 \end{aligned} \tag{1}$$

subject to

$$\sum_{t \in T} y_i^t \leq 1 \quad i \in I^e \tag{2}$$

$$\sum_{k=1}^{n_i} \sum_{m=1}^t \sum_{m'=t}^{|T|} z_{ik}^{m,m'} \leq 1 \quad i \in I^n, t \in T \tag{3}$$

$$\sum_{i \in I} v_{sip}^t \leq SQ_{sp}^t \quad s \in S, p \in P, t \in T \quad (4)$$

$$u_{ip}^{t-1} + \sum_{s \in S} v_{sip}^t = u_{ip}^t + \sum_{j \in J} d_{jp}^t x_{ijp}^t \quad i \in I, p \in P, t \in T \quad (5)$$

$$\sum_{p \in P} \beta_p u_{ip}^{t-1} + \sum_{s \in S} \sum_{p \in P} \beta_p v_{sip}^t \leq Q_i \left(1 - \sum_{t'=1}^t y_i^{t'} \right) \quad i \in I^e, t \in T \quad (6)$$

$$\sum_{p \in P} \beta_p u_{ip}^{t-1} + \sum_{s \in S} \sum_{p \in P} \beta_p v_{sip}^t \leq M_i \sum_{k=1}^{n_i} \sum_{m=1}^t \sum_{m'=t}^{|T|} k z_{ik}^{m,m'} \quad i \in I^n, t \in T \quad (7)$$

$$\sum_{p \in P} \beta_p u_{ip}^{t-1} \leq M_i n_i \left(1 - \sum_{k=1}^{n_i} \sum_{t'=1}^{|T|} z_{ik}^{t,t'} \right) + M_i \sum_{k=1}^{n_i} \sum_{t'=1}^{t-1} k z_{ik}^{t,t'} \quad i \in I^n, t \in T \setminus \{1\} \quad (8)$$

$$u_{ip}^{|T|} = 0 \quad i \in I, p \in P \quad (9)$$

$$\sum_{i \in I} x_{ijp}^t = 1 \quad j \in J, p \in P, t \in T \quad (10)$$

$$y_i^t \in \{0, 1\} \quad i \in I^e, t \in T \quad (11)$$

$$z_{ik}^{t,t'} \in \{0, 1\} \quad i \in I^n, k = 1, \dots, n_i, t, t' \in T, t' \geq t \quad (12)$$

$$x_{ijp}^t \in \{0, 1\} \quad i \in I, j \in J, p \in P, t \in T \quad (13)$$

$$v_{sip}^t \geq 0 \quad s \in S, i \in I, p \in P, t \in T \quad (14)$$

$$u_{ip}^t \geq 0 \quad i \in I, p \in P, t \in T \quad (15)$$

The first two components of the objective function (1) represent the fixed costs to operate and remove company-owned warehouses, respectively. The third component of (1) gives the total fixed cost of leasing storage space in new warehouses. The total cost of distributing the product families from the suppliers to the warehouses and from the latter to the customer zones are calculated in the fourth and fifth components, respectively. The last component determines the total inventory holding cost at the warehouses.

Constraints (2) allow the activity of each company-owned warehouse to be terminated at most once during the planning horizon. Hence, when the warehouse activity is discontinued, it cannot be restored later. Constraints (3) guarantee that multiple contracts cannot be in place at the same new warehouse location with overlapping rental periods. In addition, these constraints also ensure that at each potential warehouse location a single capacity size (i.e., a fixed number of modular units) can be selected during the lease term of a given contract. Inequalities (4) pertain to supplier capacities. Inventory balance conditions are imposed by constraints (5) for every product family and warehouse location at each time period. The capacities of company-owned warehouses and leased warehouses are enforced through constraints (6) and (7), respectively. The latter constraints also guarantee that when a lease commitment is not in place on a new warehouse in a given period (i.e., the right-hand side of the associated inequality is zero), the inventory level must be equal to zero at the end of the previous period. Constraints (8) ensure that no inventory is held in a new warehouse at the beginning of the first period in which the lease takes effect, unless another contract was in place at that location in the preceding period. If a given warehouse is leased with m modular units until the end of some period t , and a new lease begins in period $t + 1$ with m' modular units, then constraints (7) and (8) govern the maximum quantity that can be held in stock at the end of each period. If $m < m'$ (i.e., capacity expansion occurs), then the inventory level at the end of period t is at most as large as the total capacity of the m modules. In the opposite case, i.e. $m \geq m'$ (capacity reduction), the stock level at the end of period t cannot exceed the total

capacity of m' modular units. Equalities (9) enforce the inventory level to be zero at the end of the planning horizon for any product family and in any warehouse. Constraints (10) dictate that each customer zone be served by exactly one warehouse for a given product family in each period. Finally, binary and non-negativity conditions are imposed by constraints (11)–(15).

In the following, formulation (1)–(15) will be denoted by $(\mathcal{M})$. This formulation differs from that of Correia and Melo (2022) in three ways. First, the storage capacity in new warehouses is given by the number of modular units that are leased, and this number can vary over the planning horizon (i.e., increase or decrease), thus enhancing the flexibility to respond to variations in costs and demand. By contrast, in the problem studied by Correia and Melo (2022), all lease commitments in a new warehouse include the same fixed amount of capacity, thereby precluding the adjustment of capacity to respond to variations in demand. Second, since in the new problem there may be multiple consecutive lease agreements at the same warehouse for a varying number of modular units, inventory can be transferred from one contract to the next, a feature that was also not captured by Correia and Melo (2022). The third important difference concerns the fact that each product family has to be distributed by a single warehouse to a customer zone to meet its demand at a given period. The associated single-assignment constraints (10) are not present in the model of Correia and Melo (2022) and pose significant challenges due to the large number of additional binary variables (x) that have to be defined. However, many companies favor this distribution strategy because it simplifies the management of their supply chain and reduces operational costs. On the other hand, customers also often prefer to be supplied by a single facility for a particular product, as they can take advantage of lower distribution costs due to higher shipping volumes. Note that we do not enforce each customer zone to be served by a single warehouse for all product families, in order to mitigate the risk of supply disruption due to dependence on a single source. In addition to equalities (10), also constraints (3), (7), (8), (9) as well as conditions (12) and (13) do not appear in the formulation proposed by Correia and Melo (2022). The flow conservation constraints (5) also differ from those modeled by Correia and Melo (2022) due to the single-assignment requirements. Owing to these differences, formulation $(\mathcal{M})$ is structurally different from Correia and Melo (2022), resulting in a very challenging problem for which the chance of a state-of-the-art MIP solver to identify (near-)optimal solutions for large-scale instances within acceptable time is rather limited, as our computational study in Section 5 will demonstrate.

4. Matheuristic

An in-depth analysis of the characteristics of the optimal solution to the LP relaxation of formulation $(\mathcal{M})$, for a preliminary set of randomly generated problem instances, revealed that a very large number of customer assignment variables x are binary in that solution (an average of 99%). This feature motivated the development of the matheuristic that will be presented in this section. In addition to selecting a subset of the binary variables x in the LP solution that will be retained in the original MIP problem according to specific rules, several additional schemes to fix further binary variables are also developed. Fig. 2 summarizes the various steps of the proposed matheuristic, while Algorithm 1 describes the general framework.

Initially, the binary variables x , y and z are free, being gathered in the sets X_{nf} , Y_{nf} and Z_{nf} , respectively (lines 3–6 in Algorithm 1). Throughout the heuristic procedure, some (or all) binary variables will be successively fixed according to different rules, and moved to sets X_f , Y_f and Z_f . In the first phase of the heuristic (lines 8–12), variable fixing decisions build upon information drawn from the optimal solution $(\bar{x}, \bar{y}, \bar{z}, \bar{u}, \bar{v})$ to the LP relaxation of a residual problem. The latter corresponds to the original MIP problem $(\mathcal{M})$ with a subset of the

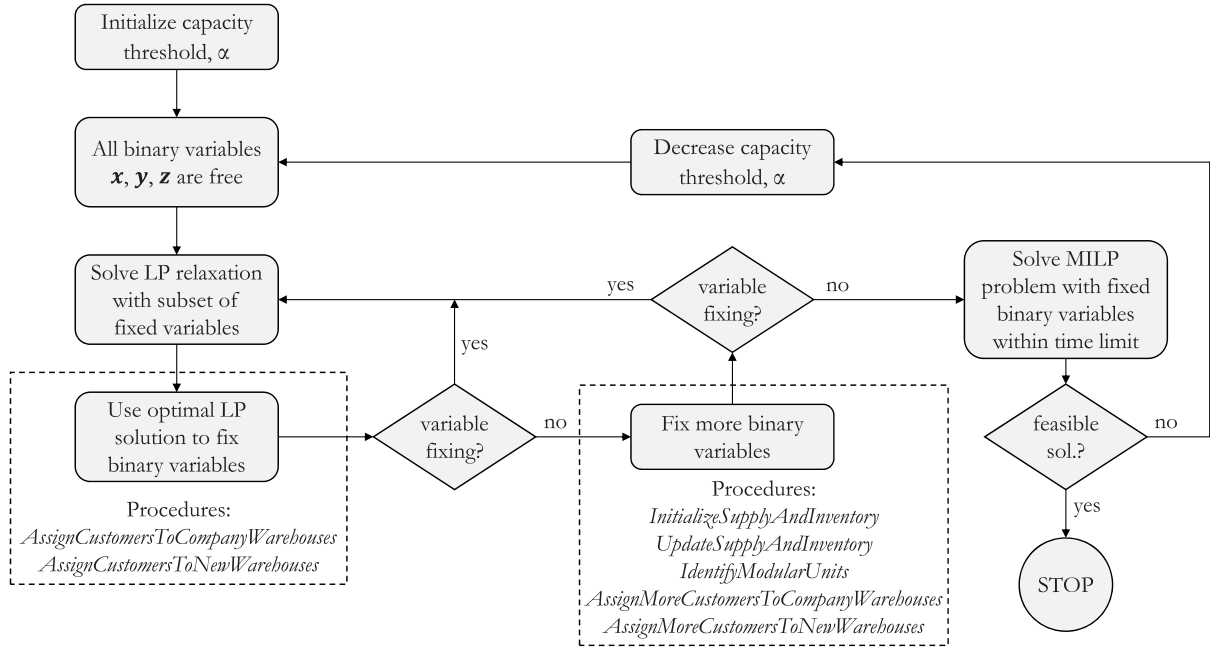


Fig. 2. Flow diagram of the proposed matheuristic.

binary variables having fixed values. Naturally, the first time this mechanism is employed, the residual problem coincides with the original problem. Procedures *AssignCustomersToCompanyWarehouses* and *AssignCustomersToNewWarehouses* (lines 10–11) include the rules governing variable fixing that are applied to the optimal solution of the last relaxed problem solved. They rely on the identification of those warehouses that are operating and whose capacity utilization rate for meeting the demand of selected customer zones and holding stock is equal to or even exceeds a pre-specified threshold α ($0 < \alpha < 1$) in a certain period. Details are provided in Section 4.1. Each time the values of (some) binary variables x , y and z are fixed, the LP relaxation of the resulting residual problem is solved optimally. This process is repeated until no more variables can be fixed.

In the next phase of the matheuristic (lines 13–23), an attempt is made to fix the values of further binary variables. For this purpose, we first estimate in each time period and for every product family the total quantity provided by suppliers that has not yet been consumed due to decisions related to set X_f . Moreover, we also estimate the inventory level in each warehouse and time period that has not yet been used to serve the customer demands associated with set X_f . These calculations are performed in Procedures *InitializeSupplyAndInventory* and *UpdateSupplyAndInventory* (lines 13–14), see Section 4.2 for a detailed description. Next, and following Procedure *AssignMoreCustomersToCompanyWarehouses* (see Section 4.3), we try to assign more customer zones to company-owned warehouses with sufficient residual capacity to meet their demand, taking into account the available supply and inventory levels. Preference is given to customer zones with lower unit distribution costs. A similar strategy is also adopted to extend the set X_f with additional assignments of customer zones to leased warehouses (see Procedure *AssignMoreCustomersToNewWarehouses* in Section 4.3). In this case, we first determine the number of modular units that should be operated in each period in a leased warehouse due to previously having fixed some variables x (Procedure *IdentifyModularUnits*, line 21). If either or both of these mechanisms succeed (lines 17 and 23), we explore the LP relaxation of the associated residual problem, repeating the cycle in lines 8–12. Otherwise, and since no more binary variables were fixed, we solve a restricted MIP problem with the choices made so far (line 26). To this end, a general-purpose MIP solver is used. If it is not

possible to identify a feasible solution to this residual MIP problem within a pre-specified time limit T_{max} , the capacity threshold α is decreased by a given amount θ . The matheuristic stops when a feasible solution to the original problem is returned or when the parameter α becomes negative, which means that no feasible solution could be identified. We emphasize that in all the computational experiments we have performed, we have always obtained a feasible solution with the initial choice of the capacity threshold (α_0 , see line 1 in Algorithm 1), and therefore it was never necessary to decrease α and re-run the matheuristic.

4.1. Using the LP relaxation for variable fixing

As stated at the beginning of Section 4, in our set of test instances, an average of 99% of the customer assignment variables x take on the value 0 or 1 in the optimal LP solution. While a natural approach would be to retain all these values in the original formulation ($\mathcal{M}$), this could lead to some warehouses being in place to serve a relatively small amount of demand in a given period, thus producing a costly solution. To overcome the potential risk of obtaining an inferior solution, we carefully select, from among the variables $\bar{x}$ with value equal to 1, those that we will transfer to the set X_f . This choice is based on the amount of capacity used by a warehouse in a given period for meeting the demand of the (fully) assigned customer zones and holding inventory, according to the components of the LP solution for that warehouse and period. If the warehouse load rate reaches or even exceeds a certain threshold α , then these customer assignments are retained. For company-owned warehouses $i \in I^e$, these steps are described in lines 4–7 of Procedure *AssignCustomersToCompanyWarehouses* for every period $t \in T$. Fixing one or several variables x_{ijp}^t at one implies that warehouse i has to be open from the beginning of the planning horizon until the period t under consideration. Hence, variables $y_i^1, \dots, y_i^t$ are also fixed at zero (lines 9–10). Furthermore, due to the single-assignment conditions (10), we also set $x_{i'jp}^t = 0$ for all warehouses $i' \in I^e \setminus \{i\}$.

A similar mechanism is applied to locations where storage space can be leased, as described in Procedure *AssignCustomersToNewWarehouses*. In this case, it is first required to identify for each location $i \in I^n$ and period $t \in T$ the number of modular units

Algorithm 1: Matheuristic framework

Input : All instance data; initial capacity threshold α_0 ; threshold reduction factor θ ; time limit to solve residual MIP problem T_{max}

Output: Solution S

```

1  $\alpha := \alpha_0$ 
2 while feasible solution is not available and  $\alpha > 0$  do
3   Fix all binary variables  $x$ ,  $y$  and  $z$  as free variables
4   Initialize  $X_{nf}$  with all  $x_{ijp}^t$  ( $i \in I$ ,  $j \in J$ ,  $p \in P$ ,  $t \in T$ ) and  $X_f := \emptyset$ 
5   Initialize  $Y_{nf}$  with all  $y_i^t$  ( $i \in I^e$ ,  $t \in T$ ) and  $Y_f := \emptyset$ 
6   Initialize  $Z_{nf}$  with all  $z_{ik}^{t,t'}$  ( $i \in I^n$ ,  $1 \leq k \leq n_i$ ,  $t \in T$ ,  $t \leq t' \leq |T|$ )
   and  $Z_f := \emptyset$ 
7   repeat
8     repeat
9       Obtain optimal solution  $(\bar{x}, \bar{y}, \bar{z}, \bar{u}, \bar{v})$  to LP relaxation of
       residual problem with fixed variable values according to
        $X_f$ ,  $Y_f$  and  $Z_f$ 
10      AssignCustomersToCompanyWarehouses( $\alpha$ ,  $\bar{x}$ ,  $\bar{u}$ ,
        $x$ ,  $y$ ,  $X_f$ ,  $X_{nf}$ ,  $Y_f$ ,  $Y_{nf}$ )
11      AssignCustomersToNewWarehouses( $\alpha$ ,  $\bar{x}$ ,  $\bar{u}$ ,  $x$ ,  $z$ ,
        $X_f$ ,  $X_{nf}$ ,  $Z_f$ ,  $Z_{nf}$ )
12    until no variable fixing occurs
13    InitializeSupplyAndInventory( $\bar{u}$ ,  $TS$ ,  $Inv$ )
14    UpdateSupplyAndInventory( $x$ ,  $TS$ ,  $Inv$ )
15    if  $X_{nf} \neq \emptyset$  then
16      // serve additional customer demands from
      company-owned warehouses
17      AssignMoreCustomersToCompanyWarehouses( $\bar{u}$ ,
        $x$ ,  $y$ ,  $X_f$ ,  $X_{nf}$ ,  $Y_f$ ,  $Y_{nf}$ ,  $TS$ ,  $Inv$ )
18    end
19    if  $X_{nf} \neq \emptyset$  then
20      // for each new warehouse identify the
      periods during which it is leased and the
      number of modular units that are in place
21      IdentifyModularUnits( $\bar{u}$ ,  $x$ ,  $nTotalPeriods$ ,
        $nTotalModUnits$ ,  $nModUnits$ )
22      // serve additional customer demands from
      leased warehouses
23      AssignMoreCustomersToNewWarehouses( $\bar{u}$ ,  $x$ ,  $z$ ,
        $X_f$ ,  $X_{nf}$ ,  $Z_f$ ,  $Z_{nf}$ ,  $TS$ ,  $Inv$ ,  $nTotalPeriods$ ,
        $nTotalModUnits$ ,  $nModUnits$ )
24    end
25  until no variable fixing occurs
26  Solve the residual MIP problem with the binary-valued variables
  according to  $X_f$ ,  $Y_f$ ,  $Z_f$ , within the time limit  $T_{max}$ 
27  if a feasible solution is not found then
28     $\alpha := \alpha - \theta$  // reduce the capacity threshold
29  end
30 end
31 return solution  $S = (x, y, z, u, v)$ 

```

($\bar{m}$) needed to cover the total quantity processed by that warehouse (line 4). Note that this quantity corresponds to all the demand fully satisfied by the warehouse and the amount of inventory built up in this facility, according to the information drawn from the linear relaxation. After fixing some variables x at 1 (lines 8–9) following the same rule as described before, we determine a lower bound Δ on the number of modular units at location i and period t (line 11), by disregarding the information on inventory holding from the LP relaxation. Observe that the matheuristic is concerned with fixing the values of (a subset of) the binary variables, deferring all decisions about the quantities to be distributed and stored (i.e., the values of the continuous variables v and u) to a later stage, where a general-purpose MIP solver is used. This explains the calculation of Δ . Given Δ , it is clear that any lease commitment in place at period t cannot include fewer than Δ modules. As a result, a subset of variables z can be fixed at zero for all leases

Algorithm AssignCustomersToCompanyWarehouses: Assignment of customer zones to company-owned warehouses over the planning horizon using information drawn from the optimal solution to the LP relaxation of the residual MIP problem

Input : α , $\bar{x}$, $\bar{u}$, x , y , X_f , X_{nf} , Y_f , Y_{nf}

Output: Additional fixing of variables x , y ; updated sets X_f , X_{nf} , Y_f , Y_{nf}

```

1 for  $i \in I^e$  do
2   for  $t \in T$  do
3     // determine warehouse capacity utilization
     level to serve demands of assigned customer
     zones and to hold inventory according to LP
     solution
4      $WCL := \frac{\sum_{j \in J, p \in P: \bar{x}_{ijp}^t = 1} \beta_p d_{jp}^t + \sum_{p \in P} \beta_p \bar{u}_{ip}^t}{Q_i}$ 
5     if  $WCL \geq \alpha$  then
6       For all  $j \in J$  and  $p \in P$  such that  $\bar{x}_{ijp}^t = 1$  and  $x_{ijp}^t \in X_{nf}$ ,
       fix  $x_{ijp}^t := 1$  and  $x_{i'jp}^t := 0$  for all  $i' \in I^e \setminus \{i\}$ 
7       Move all fixed variables from  $X_{nf}$  to  $X_f$ 
8       // company-owned warehouse  $i$  must be open at
       least until period  $t$ 
9       for  $\tau := 1$  to  $t$  do
10        if  $y_i^\tau \in Y_{nf}$  then fix  $y_i^\tau := 0$  and move this variable
        from  $Y_{nf}$  to  $Y_f$ 
11      end
12    end
13  end
14 end
15 return  $x$ ,  $y$ ,  $X_f$ ,  $X_{nf}$ ,  $Y_f$ ,  $Y_{nf}$ 

```

with k modular units such that $1 \leq k < \Delta$ (lines 13–19). Moreover, the LP relaxation of the resulting residual problem is strengthened by adding an inequality guaranteeing that a valid lease contract at location i and period t must have at least Δ modular units (line 20).

4.2. Identification of available supply and inventory levels

Solving the LP relaxation of a residual problem that no longer allows fixing variables x (and as a result, y and z too) sets the end of the first phase. The associated optimal LP solution serves as a starting point for the next phase, which aims to fix some more binary variables driven by the assignment of additional customer zones to open warehouses. This requires first identifying in every period the quantity of each product family that can be made available by suppliers and that has not yet been used. In addition, it is also necessary to determine the amount of stock in every open warehouse at the end of each period that has not yet been consumed.

For each warehouse and time period, the flow conservation constraints (5) set the link between supply, inventory, and demand for a given product family. However, they do not specify whether demand is entirely satisfied from suppliers, from inventory on hand, or from a combination of both sources. As a result, at the beginning of the second phase we are unable to deduce the exact quantities of each product family that are available to serve customer zones not yet assigned to any warehouse. To overcome this lack of information, we estimate these quantities based on certain assumptions. To this end, we start by creating two vectors, TS and Inv in Procedure InitializeSupplyAndInventory, that aggregate the total quantity of each product family made available by suppliers and the inventory level in each warehouse (according to the optimal LP solution, see lines 2–3) over the planning horizon, respectively. These vectors are initialized without considering the demand already met in the first phase, and will be later updated taking into account this information.

Algorithm AssignCustomersToNewWarehouses: Assignment of customer zones to leased warehouses over the planning horizon using information drawn from the optimal solution to the LP relaxation of the residual MIP problem

Input : $\alpha, \bar{x}, \bar{u}, x, z, X_f, X_{nf}, Z_f, Z_{nf}$
Output: Additional fixing of variables x, z ; updated sets X_f, X_{nf}, Z_f, Z_{nf}

```

1 for  $i \in I^n$  do
2   for  $t \in T$  do
3     // identify the number of modular units
      required to hold inventory and serve demands
      assigned to warehouse  $i$  in period  $t$ 
4      $\bar{m} := \left\lceil \frac{\sum_{j \in J, p \in P: \bar{x}_{ijp}^t = 1} \beta_p d_{jp}^t + \sum_{p \in P} \beta_p \bar{u}_{ip}^t}{M_i} \right\rceil$ 
5     // determine the warehouse capacity
      utilization level
6      $WCL := \frac{\sum_{j \in J, p \in P: \bar{x}_{ijp}^t = 1} \beta_p d_{jp}^t + \sum_{p \in P} \beta_p \bar{u}_{ip}^t}{\bar{m} \times M_i}$ 
7     if  $WCL \geq \alpha$  then
8       For all  $j \in J$  and  $p \in P$  such that  $\bar{x}_{ijp}^t = 1$  and  $x_{ijp}^t \in X_{nf}$ ,
          fix  $x_{ijp}^t := 1$  and  $x_{i'jp}^t := 0$  for all  $i' \in I^n \setminus \{i\}$ 
9       Move all fixed variables from  $X_{nf}$  to  $X_f$ 
10      // identify the minimum number of modular
          units required to meet the demand of
          customers assigned to warehouse  $i$  in
          period  $t$  due to variable fixing
11       $\Delta := \left\lceil \frac{\sum_{j \in J, p \in P: \bar{x}_{ijp}^t = 1} \beta_p d_{jp}^t}{M_i} \right\rceil$ 
12      // any lease contract covering period  $t$ 
          cannot include fewer than  $\Delta$  modular units
13      for  $k := 1$  to  $\Delta - 1$  do
14        for  $\tau := 1$  to  $t$  do
15          for  $\tau' := t$  to  $|T|$  do
16            if  $z_{ik}^{\tau, \tau'} \in Z_{nf}$  then fix  $z_{ik}^{\tau, \tau'} := 0$  and move
              this variable from  $Z_{nf}$  to  $Z_f$ 
17          end
18        end
19      end
20      Add the following constraint to the LP relaxation of
          residual problem ensuring that warehouse  $i$  is leased in
          period  $t$  with at least  $\Delta$  modular units:
           $\sum_{k=\Delta}^{n_i} \sum_{\tau=1}^t \sum_{\tau'=t}^{|T|} z_{ik}^{\tau, \tau'} \geq 1$ 
21    end
22  end
23 end
24 return  $x, z, X_f, X_{nf}, Z_f, Z_{nf}$ 

```

The quantities made available by suppliers in the first and last periods are specified in lines 5 to 8 for each product family. Since company-owned warehouses do not hold stock at the beginning of the planning horizon, any quantity stored at the end of the first period must be provided by the suppliers (see line 6). Thereafter (i.e., for $t = 2, \dots, |T| - 1$), we decide when and what fraction of a warehouse's inventory is carried over from one period to the next. A reasonable approach to managing inventories is to assume that when the stock level of product p in warehouse i at the end of period t is larger than the stock level in the preceding period, this results from having transferred the entire stock from $t - 1$ to t . For illustration purposes, suppose that $Inv_{ip}^{t-1} = 40$ and $Inv_{ip}^t = 60$. We assume that all 40 units are carried over to t and the remaining 20 units are provided by suppliers. Hence, Inv_{ip}^{t-1} cannot be used to meet demand, which is indicated by dropping it to zero. Lines 14–17 deal with this case. In the opposite situation, e.g., $Inv_{ip}^{t-1} = 100$ and $Inv_{ip}^t = 70$, we assume that 70 units

Algorithm InitializeSupplyAndInventory: Identification of supply and inventory levels over the planning horizon that are available to meet demand

Input : $\bar{u}$: inventory levels in the LP optimal solution
Output: TS, Inv : total supply and inventory available to cover additional demand

```

1 // initialize inventory vector with inventory levels
  from LP relaxation
2  $Inv_{ip}^0 := 0$  ( $i \in I, p \in P$ )
3  $Inv_{ip}^t := \bar{u}_{ip}^t$  ( $t \in T, i \in I, p \in P$ )
4 // determine total quantity of each product family
  made available by the suppliers in the first and
  last time periods for demand satisfaction
5 for  $p \in P$  do
6    $TS_p^1 := \sum_{s \in S} SQ_{sp}^1 - \sum_{i \in I} \bar{u}_{ip}^1$ 
7    $TS_p^{|T|} := \sum_{s \in S} SQ_{sp}^{|T|}$ 
8 end
9 // supply and inventory levels in the remaining time
  periods
10 for  $p \in P$  do
11   for  $t := 2$  to  $|T| - 1$  do
12      $TS_p^t := \sum_{s \in S} SQ_{sp}^t$ 
13     for  $i \in I$  do
14       if  $Inv_{ip}^t \geq Inv_{ip}^{t-1}$  then
15         // all inventory at the end of period  $t - 1$ 
          is carried over to period  $t$  and
          suppliers provide the additional
          quantity needed
16          $TS_p^t := TS_p^t - (Inv_{ip}^t - Inv_{ip}^{t-1})$ 
17          $Inv_{ip}^{t-1} := 0$  // no inventory in  $t - 1$  to serve
          more demand
18       else
19         // inventory at the end of period  $t$  is
          drawn from period  $t - 1$ 
20          $Inv_{ip}^{t-1} := Inv_{ip}^{t-1} - Inv_{ip}^t$  // inventory for
          covering further demand
21       end
22     end
23   end
24 end
25 return  $TS, Inv$ 

```

are transferred from $t - 1$ to t , and only 30 units can be used to satisfy demand in period t (see line 20).

In Procedure UpdateSupplyAndInventory, the quantities initially recorded in the vectors TS and Inv have to be reduced due to the customer allocation decisions made during the first phase of the matheuristic. After calculating the total demand for product family p that was already allocated to warehouse i in period t (line 5), it is decided from which source (inventory, supply, or both) it will be satisfied (lines 6–13). It is natural to consider that the quantity delivered by the suppliers is only used when the stock on hand is insufficient to cover the demand.

4.3. Further strategies for variable fixing

After having estimated the remaining supply and inventory levels in each time period, we attempt to meet the demand of customer zones not yet allocated. Procedure AssignMoreCustomersToCompanyWarehouses describes the steps underlying this process for the company-owned warehouses. For each warehouse $i \in I^e$ and period $t \in T$, we determine the total capacity used until the end of the planning horizon (line 4) that results from the customer assignments fixed in the first phase and the inventory built up (i.e., using the information from

Algorithm UpdateSupplyAndInventory: Adjustment of supply and inventory levels over the planning horizon, taking into account the demand allocations already made by having fixed variables x

Input : x : variables already fixed; TS : suppliers' availability; Inv : inventory levels at warehouses

Output: Updated vectors TS, Inv

```

1 for  $i \in I$  do
2   for  $p \in P$  do
3     for  $t \in T$  do
4       // determine total demand of product family
5       //  $p$  already served by warehouse  $i$  in period  $t$ 
6        $D := \sum_{j \in J: x_{ijp}^t = 1} d_{jp}^t$  // demand associated with
7       // fixed variables  $x$ 
8       if  $D \geq Inv_{ip}^{t-1}$  then
9         // demand is partially covered by stock
10        // on hand and supply
11         $TS_p^t := TS_p^t - (D - Inv_{ip}^{t-1})$  // remaining
12        // supply
13         $Inv_{ip}^{t-1} := 0$  // inventory is completely used
14        // up
15      else
16        // demand is fully covered by stock on
17        // hand
18         $Inv_{ip}^{t-1} := Inv_{ip}^{t-1} - D$  // remaining inventory
19      end
20    end
21  end
22 end
23 return  $TS, Inv$ 

```

the last LP relaxation solved). If this quantity is zero, we assume that the warehouse is not operating in period t , since we choose not to make any additional customer assignments to this facility. In the opposite case, the warehouse must be open over the entire planning horizon. Note that even if in one or several intermediate periods the warehouse does not serve any demand or hold stock, it must remain open due to constraints (2). The residual warehouse capacity is calculated in line 7 and used in Procedure AssignCustomers that will be detailed later. Furthermore, a positive capacity utilization rate prevents the warehouse activity from being discontinued before period t . Therefore, the associated y variables are set to zero (line 11) unless this has already occurred in the first phase.

Regarding the potential locations $i \in I^n$, Procedure IdentifyModularUnits starts by finding the total number of periods during which a warehouse is leased and the number of modular units needed in each period. Recall that some customer zones may have already been assigned to this type of facility in the first phase (see line 11 of Algorithm 1). This information is then used in Procedure AssignMoreCustomersToNewWarehouses to decide on further customer allocations. Since the fixed costs of leased warehouses reflect economies of scale, we rank these facilities, giving preference to those that have a higher mean number of modular units per period (line 1). Next, we examine each of the facilities in this ordered list. Suppose, for example, that according to the decisions made in the first phase, a certain warehouse is leased in periods 2, 3 and 5 with 4, 1 and 3 modular units, respectively. In period 4, the warehouse is not used. Due to economies of scale, it will certainly be more advantageous to have a lease agreement that runs continuously from period 2 to period 5. Furthermore, since the cost per unit of capacity decreases as the number of modular units grows, we can increase the current capacity level, thereby enhancing the opportunity to meet the demand from additional customer zones. In our example, the warehouse is leased with an

Algorithm AssignMoreCustomersToCompanyWarehouses: Serve further customer demands from company-owned warehouses over the planning horizon

Input : $\bar{u}, x, y, X_f, X_{nf}, Y_f, Y_{nf}, TS, Inv$

Output: Additional fixing of variables x, y ; updated sets X_f, X_{nf}, Y_f, Y_{nf} ; updated vectors TS and Inv due to demand satisfaction

```

1 for  $i \in I^e$  do
2   for  $t := |T|$  to 1 do
3     // determine total capacity used at warehouse  $i$ 
4     // from period  $t$  until the end of the planning
5     // horizon due to variable fixing
6      $WCU := \sum_{\tau=t}^{|T|} \sum_{j \in J, p \in P: x_{ijp}^\tau = 1} \beta_p d_{jp}^\tau + \sum_{\tau=t}^{|T|} \sum_{p \in P} \beta_p \bar{u}_{ip}^\tau$ 
7     if  $WCU > 0$  then
8       // determine residual capacity of warehouse
9       //  $i$  in period  $t$ 
10       $RQ := Q_i - \sum_{j \in J, p \in P: x_{ijp}^t = 1} \beta_p d_{jp}^t - \sum_{p \in P} \beta_p \bar{u}_{ip}^t$ 
11      AssignCustomers( $i, t, RQ, x, X_f, X_{nf}, TS, Inv$ )
12      // company-owned warehouse  $i$  must be open at
13      // least until period  $t$ 
14      for  $\tau := 1$  to  $t$  do
15        if  $y_i^\tau \in Y_{nf}$  then fix  $y_i^\tau := 0$  and move this variable
16        // from  $Y_{nf}$  to  $Y_f$ 
17      end
18    end
19  end
20 end
21 return  $x, y, X_f, X_{nf}, Y_f, Y_{nf}, TS, Inv$ 

```

Algorithm IdentifyModularUnits: For each new warehouse identify the total number of time periods during which the warehouse is leased and the total number of modular units in use

Input : $\bar{u}, x$

Output: $nTotalPeriods, nTotalModUnits, nModUnits$: vectors with total number of periods and modular units in leased warehouses

```

1  $nTotalPeriods_i := 0$  and  $nTotalModUnits_i := 0$  for all  $i \in I^n$ 
2  $nModUnits_i^t := 0$  for all  $i \in I^n, t \in T$ 
3 for  $i \in I^n$  do
4   for  $t \in T$  do
5     // identify the number of modules required to
6     // hold inventory and serve demands assigned to
7     // warehouse  $i$  in period  $t$  due to fixing
8     // variables  $x$ 
9      $nModUnits_i^t := \left\lceil \frac{\sum_{j \in J, p \in P: x_{ijp}^t = 1} \beta_p d_{jp}^t + \sum_{p \in P} \beta_p \bar{u}_{ip}^t}{M_i} \right\rceil$ 
10   end
11    $nTotalModUnits_i := \sum_{t \in T} nModUnits_i^t$ 
12    $nTotalPeriods_i :=$  sum of all periods  $t \in T$  for which
13   //  $nModUnits_i^t > 0$ 
14 end
15 return  $nTotalPeriods, nTotalModUnits, nModUnits$ 

```

average of 3 modules per period ($\lceil (4 + 1 + 3)/3 \rceil$). We compare this mean value to the actual number of modules in each period, and select the largest one. Thus, we assume that the warehouse is running on 4 modules in period 2 and on 3 modules in each of the periods 3, 4 and 5. Line 6 describes this step.

After having decided the number of modular units the warehouse has in period t , its residual capacity is determined (line 8) and Procedure AssignCustomers is invoked with this information. We emphasize that the associated storage capacity is not final, and it can be

Algorithm AssignMoreCustomersToNewWarehouses: Serve further customer demands from leased warehouses over the planning horizon

Input : $\bar{u}, x, z, X_f, X_{nf}, Z_f, Z_{nf}, TS, Inv, nTotalPeriods, nTotalModUnits, nModUnits$

Output: Additional fixing of variables x, z ; updated sets X_f, X_{nf}, Z_f, Z_{nf} ; updated vectors TS and Inv due to demand satisfaction

```

1 Create list  $I$  with all the warehouses  $i \in I^n$  that are leased (i.e.
   $nTotalModUnits_i > 0$ ); sort these facilities by non-increasing values
  of  $nTotalModUnits_i / nTotalPeriods_i$ , i.e., by mean number of
  modular units in use per period
2 repeat
3   Extract the first warehouse from the sorted list  $I$ , say  $i'$ 
4   for  $t \in T$  do
5     // consider the largest number of modular units
      in period  $t$ 
6      $nModUnits_{i'} := \max \left\{ nModUnits_{i'}, \left\lceil \frac{nTotalModUnits_{i'}}{nTotalPeriods_{i'}} \right\rceil \right\}$ 
7     // determine residual capacity of warehouse  $i'$ 
      in period  $t$ 
8      $RQ := nModUnits_{i'} \times M_{i'} - \sum_{j \in J, p \in P: x_{i'jp}^t = 1} \beta_p d_{jp}^t - \sum_{p \in P} \beta_p u_{i'p}^t$ 
9     AssignCustomers( $i', t, RQ, x, X_f, X_{nf}, TS, Inv$ )
10    // identify the minimum number of modular
      units required to meet the demand of
      customers assigned to warehouse  $i'$  in period  $t$ 
      due to variable fixing
11     $\Delta := \left\lceil \frac{\sum_{j \in J, p \in P: x_{i'jp}^t = 1} \beta_p d_{jp}^t}{M_{i'}} \right\rceil$ 
12    // warehouse  $i'$  cannot be leased with less than  $\Delta$ 
      modules in period  $t$ 
13    for  $k := 1$  to  $\Delta - 1$  do
14      for  $\tau := 1$  to  $t$  do
15        for  $\tau' := t$  to  $|T|$  do
16          if  $z_{i'k}^{\tau, \tau'} \in Z_{nf}$  then fix  $z_{i'k}^{\tau, \tau'} := 0$  and move this
            variable from  $Z_{nf}$  to  $Z_f$ 
17        end
18      end
19    end
20    Add the following constraint to the LP relaxation of residual
      problem ensuring that warehouse  $i'$  is leased in period  $t$ 
      with at least  $\Delta$  modular units:  $\sum_{k=\Delta}^{n_{i'}} \sum_{\tau=1}^t \sum_{\tau'=t}^{|T|} z_{i'k}^{\tau, \tau'} \geq 1$ 
21  end
22 until  $I = \emptyset$ 
23 return  $x, z, X_f, X_{nf}, Z_f, Z_{nf}, TS, Inv$ 

```

changed at a later stage of the matheuristic (recall the cycle in lines 7–25 in Algorithm 1). Next, steps similar to those of Procedure AssignCustomersToNewWarehouses are performed with respect to fixing the values of z variables and strengthening the LP relaxation bound of the resulting residual problem (recall Section 4.1).

Finally, Procedure AssignCustomers governs the allocation of unsatisfied demand to a given warehouse $i^* \in I$ with positive residual capacity in period t . Among all customer zones j and product families p eligible to be served by the warehouse in that period (line 3), the pair (j, p) with the lowest unit distribution cost from facility i^* is identified (line 5). Observe that to achieve eligibility, it is not enough that the demand associated with this pair is not greater than the residual capacity in the warehouse. In addition, this demand must also not exceed the available supply in period t plus the inventory held at the end of period $t - 1$ in the warehouse. If the demand of the selected customer zone is not more than the available stock, the latter is used to fully serve this demand, and the inventory level and the residual warehouse capacity are adjusted accordingly (lines 11–12). Otherwise, the demand is satisfied jointly by the amount of supply and stock available. In this

Algorithm AssignCustomers: Assignment of additional customer zones to a given warehouse i^* in time period t

Input : i^*, t : selected warehouse and period; RQ : residual capacity of warehouse i^* in period t ; x, X_f, X_{nf}, TS, Inv

Output: Additional fixing of variables x ; updated sets X_f, X_{nf} ; updated vectors TS and Inv due to demand satisfaction

```

1 while  $RQ > 0$  and  $X_{nf} \neq \emptyset$  do
2   // identify unassigned customer zones and product
      families in period  $t$  for which there is
      sufficient capacity and supply to serve their
      demands by warehouse  $i^*$  in that period
3   Create set  $A$  with pairs  $(j, p)$  such that  $x_{i^*jp}^t \in X_{nf}, \beta_p d_{jp}^t \leq RQ$ 
      and  $d_{jp}^t \leq TS_{jp}^t + Inv_{i^*p}^{t-1}$ 
4   // identify the pair with lowest unit distribution
      cost from warehouse  $i^*$ 
5    $(j^*, p^*) := \arg \min_{(j,p) \in A} \{DC_{i^*jp}^t\}$ 
6   if  $d_{j^*p^*}^t \leq Inv_{i^*p^*}^{t-1}$  then
7     // demand of  $(j^*, p^*)$  is fully served from stock on
      hand in warehouse  $i^*$  at the beginning of
      period  $t$ 
8     Fix  $x_{i^*j^*p^*}^t := 1$  and  $x_{i^*jp^*}^t := 0$  for all  $i' \in I \setminus \{i^*\}$ 
9     Move all fixed variables from  $X_{nf}$  to  $X_f$ 
10    // update residual capacity and inventory
      level of the warehouse
11     $RQ = RQ - \beta_{p^*} d_{j^*p^*}^t$ 
12     $Inv_{i^*p^*}^{t-1} := Inv_{i^*p^*}^{t-1} - d_{j^*p^*}^t$ 
13  else
14    // demand is partly served by stock on hand and
      supply since  $d_{j^*p^*}^t \leq TS_{j^*p^*}^t + Inv_{i^*p^*}^{t-1}$ 
15    Fix  $x_{i^*j^*p^*}^t := 1$  and  $x_{i^*jp^*}^t := 0$  for all  $i' \in I \setminus \{i^*\}$ 
16    Move all fixed variables from  $X_{nf}$  to  $X_f$ 
17    // update residual capacity, supply and
      inventory level of warehouse
18     $RQ = RQ - \beta_{p^*} d_{j^*p^*}^t$ 
19     $TS_{j^*p^*}^t := TS_{j^*p^*}^t - (d_{j^*p^*}^t - Inv_{i^*p^*}^{t-1})$ 
20     $Inv_{i^*p^*}^{t-1} := 0$ 
21  end
22 end
23 return  $x, X_f, X_{nf}, TS, Inv$ 

```

case, the supply quantity and the residual capacity are decreased and the inventory level drops to zero (lines 18–20). Moreover, a successful demand allocation results in fixing all the variables x related to this customer zone, product family, warehouse, and period, as stated in lines 8 and 15.

5. Computational study

In this section, we evaluate the computational efficiency of the proposed matheuristic and the quality of the solutions obtained for a set of problem instances. In addition, we compare the heuristic solutions against the solutions identified by CPLEX within a given time limit. In Section 5.1, we describe the settings used to generate the problem instances and present the numerical results in Section 5.2.

5.1. Problem instances

Since the problem studied by Correia and Melo (2022) has the greatest number of features in common with our problem, we adapted the scheme proposed by these authors to randomly generate a total of 160 instances according to the parameters specified in Table 1. The interested reader is referred to the supplementary material, which contains detailed information on the data generation procedure.

Table 1

Parameter settings.

Description	Parameter	Value(s)
No. of periods	$ T $	12, 24, 36, 48
No. of suppliers	$ S $	3
No. of warehouses	$ I $	10, 20
No. of company-owned/potential warehouses	$ I^c / I^n $	5/5, 10/10
No. of customer zones	$ J $	2 $ I $, 5 $ I $, 10 $ I $, 20 $ I $
No. of product families	$ P $	3
Max. no. of modular units in a leased warehouse	n_i ($i \in I^n$)	5
Unit capacity consumption factor	β_p ($p \in P$)	1

Table 2

Size of problem instances.

$ T $	# binary variables			# continuous variables	# constraints
	x	y	z		
12	min	7 200	60	1 950	1 458
	avg	83 250	90	2 925	6 048
	max	288 000	120	3 900	15 768
24	min	14 400	120	7 500	2 886
	avg	166 500	180	11 250	4 320
	max	576 000	240	15 000	5 760
36	min	21 600	180	16 650	4 314
	avg	249 750	270	24 975	6 480
	max	864 000	360	33 300	8 640
48	min	28 800	240	29 400	5 742
	avg	333 000	360	44 100	8 640
	max	1 152 000	480	58 800	11 520
All	avg	208 125	225	20 813	5 400

Assuming that a time period represents one month, the values of $|T|$ shown in Table 1 refer to timeframes of 1, 2, 3, and 4 years. We emphasize that the length of the planning horizon in our instances is significantly longer than what is often reported in the literature dedicated to multi-period facility location problems (see e.g., Becker et al., 2019; Correia & Melo, 2017; Cortinhal et al., 2015; Jena et al., 2015a, 2017; Sauvey et al., 2020), thus greatly impacting the size of the instances in our testbed.

For each combination of the parameter values in Table 1, five instances were randomly generated. Observe that the larger instances have 48 time periods, 10 company-owned warehouses, 10 potential locations for leasing storage capacity, and 400 customer zones, resulting in formulation ($\mathcal{M}$) having over 1.2 million binary variables, as shown in Table 2. The customer assignment variables x account for 95% of this number. The table also gives the total number of continuous variables and the total number of constraints. It can be seen that our computational study includes medium- and large-sized problem instances.

In each instance, three product families with distinct demand and supply patterns are considered, two of which exhibiting seasonality profiles. Product family 1 undergoes irregular demand variations over the planning horizon that range from -5% to $+5\%$ between two consecutive periods. By contrast, product family 2 is subject to advertising campaigns (e.g., Black Friday, Cyber Monday, Christmas promotions), which increase its demand in the last three periods of each 12-period block. For example, for $|T| = 24$, demand peaks occur in periods 10–12 and 22–24. Many products in the consumer electronics industry exhibit this type of pattern. As for product family 3, its demand faces significant growth in the middle of each set of 12 consecutive periods. This pattern represents products with higher sales in the summer season (e.g., air coolers and air conditioners). For instance, in a 3-year planning horizon ($|T| = 36$), peak demand occurs in periods 6–8, 18–20, and 30–32. Detailed information on the random generation of the three demand profiles is presented in the supplementary material.

Table 3

Quality of the solutions identified by CPLEX within a 5 h time limit.

$ T $	# inst. solved ^a / # inst. no sol. found		MIP gap (%)	UB gap (%)
12	40/0	min	0.68	1.96
		avg	4.87	6.29
		max	26.43	36.19
24	40/0	min	1.93	3.03
		avg	5.83	6.90
		max	26.66	36.55
36	40/0	min	2.82	3.21
		avg	9.15	10.97
		max	33.62	50.84
48	39/1	min	3.72	3.96
		avg	10.07	11.90
		max	20.66	26.34
All	159/1	avg	7.46	9.00

^a No instance was solved to guaranteed optimality.

The suppliers' capacities are also generated randomly, and are associated with the individual demand profiles of the product families. Accordingly, for family 1, the capacity of each supplier is 25 to 50% higher than the average demand in each time period. For families 2 and 3, the supplier capacity increases in the three periods prior to the start of each peak demand season. This scheme reflects common practices for seasonal commodities, whereby production occurs ahead of time to respond to the projected growth in demand. The supplementary material details how the suppliers' capacities are generated for each family. We note that the occurrence of significant fluctuations in demand in our problem instances induces variations in storage capacity over time, thus triggering the need to increase and decrease the number of modular units leased in new warehouses. In addition, the mismatch between supply and demand may also prompt the need to hold inventory at some facilities in individual periods so that demand can be met in later periods. Additional parameters (i.e., capacities of the company-owned warehouses, capacity of a modular unit, and cost parameters) are also described in the supplementary material. We highlight that the generation of the fixed costs for warehouse leasing ($L_{ik}^{t,f}$) accounts for economies of scale that favor leasing a facility with a larger number of modular units for a longer rental term. Economies of scale are also present in the fixed costs of operating company-owned warehouses (OC_i^t).

5.2. Numerical results

Formulation ($\mathcal{M}$) was coded in C++ and embedded in IBM ILOG Concert Technology with IBM ILOG CPLEX 20.1. A limit of 5 h of CPU time was set for each solver run. The matheuristic was implemented in C++, and CPLEX was also employed to solve the LP relaxations of the various residual problems as well as the last residual MIP problem (recall lines 9 and 26 of Algorithm 1). For solving the latter problem, a time limit of 20 min was imposed. Moreover, CPLEX was always used with default settings under a deterministic parallel mode. All experiments were performed on a dual AMD EPYC™ 7702 workstation, with each processor having 64 cores, 256 MB cache, 2 to 3.35 GHz, and 512 GB/3200 RAM.

Table 3 summarizes the quality of the solutions returned by CPLEX according to the length of the planning horizon in the problem instances (column 1). Column 2 indicates the number of instances for which CPLEX found a feasible solution (# inst. solved) and the number of instances for which CPLEX could not identify a single feasible solution within the specified time limit (# inst. no sol. found). We remark that our data generation scheme ensures that all instances have feasible solutions. The optimality gap as reported by CPLEX upon termination is given in column 4 (MIP gap). This gap is determined by $(z^{UB} - z^{LB})/z^{UB} \times 100$, with z^{UB} denoting the objective value of

the best feasible solution returned by CPLEX for an individual instance, and z^{LB} representing the best lower bound identified by CPLEX for that instance. The ‘UB gap’ in column 5 measures the quality of z^{UB} with respect to the LP bound, i.e., $(z^{UB} - z^{LP})/z^{LP} \times 100$, where z^{LP} is the LP bound. Table 3 reports the minimum, average, and maximum values of the performance measures MIP gap and UB gap. The last row (All) of this table provides information over all problem instances.

Although CPLEX identifies feasible solutions for all but one problem instance within the 5-hour time limit, it could not prove the optimality of any of them. This suggests that we are facing a challenging problem, even for medium-sized instances with planning horizons with 12 periods. This feature is also confirmed by the optimality gaps reported by CPLEX. A great deal of spread around the average MIP gap of 7.46% is observed, regardless of the length of the planning horizon. As expected, the quality of the solutions deteriorates as the planning horizon increases. While only 3 instances exhibit an optimality gap greater than 10% for $|T| = 12$, this number increases to 17 for $|T| = 48$. A few instances even have rather large deviations up to about 34%. Since the LP relaxation provides a very good lower bound (as will be shown in Table 6), the inferior quality of many solutions is also reflected in the ‘UB gap’, which can be as high as 51%. In short, an off-the-shelf optimization software such as CPLEX is not a suitable choice for obtaining optimal or near-optimal solutions to this challenging problem in a reasonable amount of time. Furthermore, we also found that strengthening formulation ($\mathcal{M}$) with additional inequalities linking the two sets of binary variables x and y , as well as relating variables x to binary variables z in the following way

$$x_{ijp}^t \leq 1 - \sum_{i'=1}^t y_{i'}^t \quad i \in I^e, j \in J, p \in P, t \in T$$

$$x_{ijp}^t \leq \sum_{k=1}^{n_i} \sum_{m=1}^t \sum_{m'=t}^{|T|} z_{ik}^{m,m'} \quad i \in I^n, j \in J, p \in P, t \in T$$

is not advantageous, either in terms of the quality of the feasible solutions returned by CPLEX or the computational burden they involve.

In contrast to CPLEX, the matheuristic identifies a feasible solution to every tested instance and for all values of the capacity threshold α selected. To evaluate the quality of the heuristic solutions, a comparison was carried out with the best solutions returned by CPLEX, pertaining to 159 instances. Table 4 presents the results obtained for different performance measures. For each value of $|T|$, columns 2–5 indicate the number of heuristic solutions with better (# superior) and worse (# inferior) quality compared to CPLEX for $\alpha \in \{0.6, 0.7, 0.8, 0.9\}$. Given a problem instance, let $z^{H(\alpha)}$ denote the objective function value of the heuristic solution obtained for a particular choice of α , and let z^{UB} be the best objective value achieved by CPLEX for that instance. For those heuristic solutions with superior quality compared to CPLEX (i.e., $z^{H(\alpha)} < z$), the associated improvement is calculated as follows: $(z^{UB} - z^{H(\alpha)})/z^{UB} \times 100$. Conversely, for those heuristic solutions with lower quality, their performance is measured by $(z^{H(\alpha)} - z^{UB})/z^{UB} \times 100$. For each value of the capacity threshold, columns 7–10 show the minimum, average, and maximum performance deviations, with the best values highlighted in boldface. The last row of the table summarizes the results obtained for the 159 instances.

In general, increasing the capacity threshold α favors obtaining more solutions by the matheuristic with higher quality than those found by CPLEX (e.g., 59.7% for $\alpha = 0.6$ against 68.6% for $\alpha = 0.9$). This feature is explained by the fact that when a relatively low capacity utilization rate is taken at a warehouse (e.g., $\alpha = 0.6$), this translates into considering more warehouses as being open. As a result, a larger number of binary variables x , y , and z are fixed, thus yielding in turn a residual MIP problem with smaller size. In this restricted MIP problem, which is then solved with CPLEX, the network configuration is already largely defined. The opposite effect is observed when the capacity threshold is increased, for example, to 0.9. In this case, fewer warehouses meet the minimum capacity utilization level

and, consequently, the size of the last residual problem is larger since fewer binary variables were fixed. This means that there is still some leeway for deciding on the final network configuration, thus enabling a feasible solution with higher quality to be obtained. However, this may be offset by the fact that additional computing time is required to solve this larger residual problem to optimality, and the 20-minute time limit that we impose may be insufficient to achieve this goal. Fig. 3 substantiates these observations by displaying the average percentage of binary variables x , y , and z that are fixed by the matheuristic. Furthermore, this feature is also reflected in the number of iterations performed in the first phase of the matheuristic (recall lines 8–12 in Algorithm 1) that range, on average, from 7 ($\alpha = 0.6$) to 12.7 ($\alpha = 0.9$). This is also observed for the average number of iterations carried out in the second phase (recall lines 7–25 in Algorithm 1), that vary from 2.7 ($\alpha = 0.6$) to 3.7 iterations ($\alpha = 0.9$).

According to Table 4, the performance achieved by our matheuristic competes very favorably with CPLEX regardless of the value selected for α . Among the heuristic solutions with higher quality than the CPLEX solutions, the average improvement ranges from 6.22% to 6.92%. For some instances, the matheuristic even finds substantially superior solutions, whatever the length of the planning horizon (and thus the problem size). The latter yield a total cost reduction of up to 32% compared to CPLEX (see $|T| = 36$). On the other hand, when the matheuristic returns more expensive solutions than the solutions obtained by CPLEX, the deterioration in quality is less than 5%, with the exception of ten instances (over all values of α). Although there is no warehouse capacity threshold that consistently performs the best, in practice this is not a drawback since computing time is not a critical issue in this case as shown by Table 5 (columns 7–10), even for the largest problem instances with 48 time periods. Thus, a decision maker could re-run the matheuristic for different values of the parameter α before deciding on the preferred three-echelon network configuration to be retained. In addition, when re-optimization is required to perform ‘what-if’ analyzes and a problem has to be solved repeatedly with different data sets, the reduced computational burden required by the matheuristic makes it possible to handle very large instances quite successfully.

For each length of the planning horizon, Table 5 also presents the number of instances for which the corresponding residual MIP problem could be solved either to optimality or approximately within the 20-minute time limit (columns 2–5). The impact of the selected warehouse capacity threshold is in line with the analysis of the results in Table 4 and Fig. 3. Clearly, an increment in the value of α reduces the number of binary variables that can be fixed by the matheuristic, and as a result, the size of the restricted MIP problem increases. This feature is particularly noticeable for $\alpha = 0.9$, where 30.6% (49/160) of the instances could not be solved to proven optimality within 20 min. However, the best feasible solution found for each of these residual problems is near-optimal. More specifically, the optimality gaps as reported by CPLEX do not exceed 0.08% for the three instances occurring with $\alpha = 0.7$, and vary between 0.02% and 0.34% for $\alpha = 0.8$ (12 instances). For $\alpha = 0.9$, the optimality gap ranges from 0.03% to 0.94%, with an average of 0.3% over 49 instances. Interestingly, for 48.1% (77/160) of the problem instances, the best feasible heuristic solution is obtained with the highest capacity threshold. By contrast, only 7 instances (4.4%) run with $\alpha = 0.6$ yielded the best heuristic solutions. Moreover, considering all the values selected for the parameter α , we observe that there are 118 instances (74%) where there is at least one heuristic solution of higher quality than the one returned by CPLEX (this value also includes the instance with $|T| = 48$ for which CPLEX could not identify any feasible solution within 5 h). This means that CPLEX outperforms the matheuristic for only 42 instances (26%), half of which are among the smallest (i.e., $|T| = 12$).

Next, we analyze the quality of the lower bound provided by the LP relaxation of formulation ($\mathcal{M}$). To ensure a fair evaluation, we

Table 4
Performance of the matheuristic compared to CPLEX (results reported for 159 instances).

T	# superior/inferior solutions over CPLEX					% improvement/deterioration over CPLEX			
	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.9$		$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.9$
12	14/26	16/24	16/24	18/22	min	0.00/0.37	0.57/0.02	0.06/0.02	0.06/0.00
					avg	6.56 /1.69	6.21/1.07	6.14/ 0.99	5.63/1.08
					max	25.06/3.74	25.11/3.48	25.21/3.26	25.14/5.81
24	21/19	25/15	27/13	26/14	min	0.03/0.12	0.06/0.07	0.27/0.45	0.26/0.27
					avg	4.55 /1.61	4.27/ 1.46	4.12/1.57	4.22/1.97
					max	24.18/5.27	24.98/5.27	24.97/5.42	25.43/6.10
36	30/10	31/9	32/8	34/6	min	0.18/0.07	0.21/0.17	0.27/0.00	0.03/0.29
					avg	7.24/2.15	7.35 / 1.35	7.23/1.41	6.67/2.61
					max	30.03/9.05	30.66/3.29	31.80/6.13	32.11/4.31
48	30/9	32/7	30/9	31/8	min	0.09/0.30	0.06/0.27	0.83/0.11	0.25/0.05
					avg	8.42/3.25	8.23/2.11	8.87 /2.65	7.73/ 1.72
					max	18.49/8.27	18.29/5.16	18.60/4.91	17.85/3.33
All	95/64	104/55	105/54	109/50	avg	6.92 /1.96	6.70/ 1.35	6.73/1.47	6.22/1.62

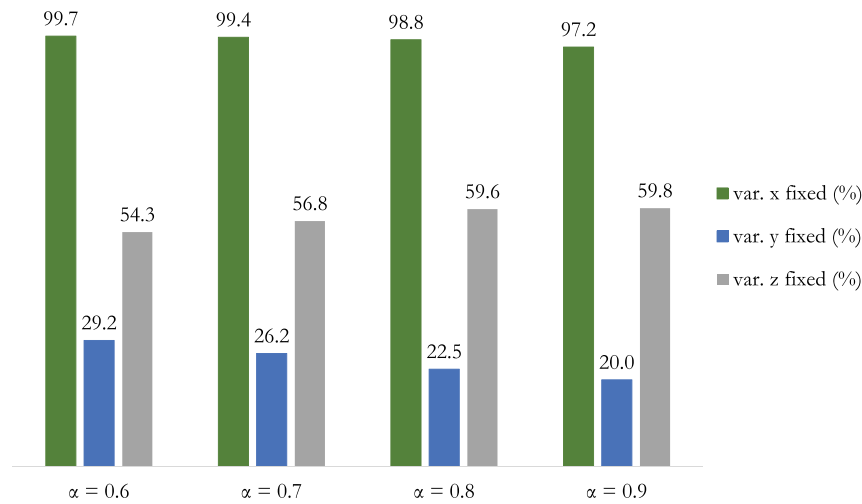


Fig. 3. Average percentage of binary variables fixed by the matheuristic.

Table 5
CPU time performance of the matheuristic.

T	# residual MIP problems solved optimally/not optimally within time limit					CPU (s)			
	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.9$		$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.9$
12	40/0	39/1	38/2	36/4	min	3.6	4.0	5.7	6.9
					avg	48.5	77.8	142.5	315.7
					max	479.9	1283.8	1500.1	1597.0
24	40/0	40/0	38/2	29/11	min	11.1	16.5	15.7	30.9
					avg	108.2	180.6	339.6	755.9
					max	345.8	1554.8	2083.5	2112.5
36	40/0	40/0	38/2	22/18	min	25.9	21.1	23.7	107.4
					avg	200.5	280.1	496.9	1164.4
					max	624.0	852.8	2025.3	3537.5
48	40/0	38/2	34/6	24/16	min	34.8	39.8	52.1	117.4
					avg	368.6	501.5	890.6	1577.4
					max	1294.9	2050.9	2920.0	5224.1
All	160/0	157/3	148/12	111/49	avg	181.4	260.0	467.4	953.3

choose to compare the LP bound with the best feasible solution available, whether this was obtained by CPLEX or by the matheuristic. Accordingly, the following ratio is calculated for every problem instance: $(z^{UB*} - z^{LP})/z^{UB*} \times 100$, with z^{LP} denoting the LP bound and z^{UB*} representing the objective value of the best feasible solution available for the individual instance, i.e. $z^{UB*} = \min\{z^{UB}, \min_{\alpha \in \{0.6, 0.7, 0.8, 0.9\}} \{z^{H(\alpha)}\}\}$. For the single instance for which CPLEX was unable to identify a feasible solution within the 5-hour time limit, $z^{UB} = +\infty$. **Table 6** gives the minimum, average, and maximum deviations measured as a percentage. We do not include in this table the computing times required to solve the LP relaxation to optimality because they were

Table 6
Comparison of the LP bound with the objective value of the best feasible solution available (%).

	T				All instances
	12	24	36	48	
min	1.79	1.80	1.72	1.59	1.59
avg	2.88	3.26	3.41	3.40	3.24
max	5.07	7.02	7.60	5.28	7.60

negligible in view of the size of the instances: on average, only 45.1 s. were needed, while the largest instance used 312.5 s.

Table 7

Comparison of the objective value of the solutions identified by the matheuristic with the LP bound.

T		Deviation of heuristic solution to LP bound (%)			
		$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.9$
12	min	2.04	1.92	1.86	1.82
	avg	4.49	3.82	3.79	3.72
	max	8.38	7.28	8.30	11.26
24	min	2.11	2.09	1.85	1.83
	avg	4.90	4.36	4.21	4.43
	max	10.46	10.46	12.78	11.20
36	min	2.31	1.98	1.80	1.75
	avg	4.98	4.44	4.32	4.57
	max	14.28	10.59	11.21	13.33
48	min	1.98	1.85	1.76	1.62
	avg	5.19	4.35	4.48	4.99
	max	13.37	9.97	11.74	14.45
All	avg	4.89	4.24	4.20	4.43

Formulation ($\mathcal{M}$) consistently produces quite good LP bounds for all instances. This feature is very useful, as it allows the quality of the feasible solutions identified by the matheuristic to be evaluated, including the instance with $|T| = 48$ that CPLEX could not solve within the time limit. In this case, the performance of the matheuristic can only be measured by comparison to the LP bound of that instance. Table 7 presents the minimum, average, and maximum gaps between the objective value of the heuristic solution for a given capacity threshold ($z^{H(\alpha)}$), and the LP bound (z^{LP}), i.e., $(z^{H(\alpha)} - z^{LP})/z^{LP} \times 100$. The best average values are shown in boldface. The results report to all 160 instances. The average deviations presented in this table reinforce the high quality performance of the matheuristic that had already been observed through other metrics, especially when the fixing of the binary variables x , y , and z is associated with a minimum warehouse capacity utilization rate α closer to 1. Gaps of more than 10% were obtained for a reduced number of instances, more precisely, 7 instances with $\alpha = 0.6$, 2 instances with $\alpha = 0.7$, 4 instances with $\alpha = 0.8$, and 6 instances with $\alpha = 0.9$.

In summary, our numerical results indicate that the proposed matheuristic consistently provides good to very good feasible solutions in significantly shorter computing time than CPLEX.

6. Conclusions

We have introduced, modeled, and solved a multi-period network redesign problem involving warehouse location and capacity sizing, distribution, and inventory holding decisions in a multi-product three-echelon network. In contrast to classical facility location problems, flexible conditions are in place regarding the amount of storage space that can be leased in new warehouses and the timeframe of each lease contract. This setting stems from warehousing-as-a-service, which allows storage capacity to be scaled over the planning horizon (i.e., capacity can be expanded or reduced, and even discontinued for multiple periods at a given location). In addition, customer single-assignment constraints have also been considered which greatly affect the size of the proposed MIP formulation, especially when the planning horizon is extensive and the number of customer zones is large, as is usually the case in practice.

We have developed the first matheuristic for this new problem which exploits the structure of the optimal solution to the LP relaxation. Thus, when the capacity utilization rate of the operating warehouses in the LP solution reaches a given threshold, the customer zones that are fully allocated to them have their variables set to one in the original MIP problem. This strategy also drives the fixing of several other binary variables. Each time a subset of binary variables is fixed, the LP relaxation of the associated residual problem is solved. This procedure is repeated until no more variables can be fixed. When that happens, the second phase of the matheuristic is started. To this end,

additional variable fixing rules are applied together with a scheme for managing inventories at the warehouses and using the product amounts made available by the suppliers. Numerical experiments with randomly generated problem instances, some of which with a very large size, reveal that the proposed matheuristic is effective both in terms of solution quality and computational efficiency. Regardless of the length of the planning horizon and the number of customer zones, the matheuristic always identifies a feasible solution to every problem instance, unlike a state-of-the-art MIP solver such as CPLEX. Moreover, it requires a significantly shorter computing time. Considering all the feasible solutions obtained with different choices for the capacity threshold, the matheuristic outperforms CPLEX in 74% of the instances, with an average improvement of 6.4%. On the other hand, when the matheuristic underperforms, the average deterioration is less than 1% compared to CPLEX. Another important finding of our study is that the proposed MIP formulation provides a tight LP bound. This feature, together with the very good performance of the matheuristic, has a practical relevance, since it enables solving large-scale problems with various parameter combinations (e.g., different demand estimates), and thus understanding how the configuration of the three-echelon network and the associated total cost are affected.

A future research venue would be to integrate uncertainty into the problem, acknowledging that at the time of planning, demand, supplier and rental warehouse capacities, and costs are unknown with certainty. Naturally, this would further increase the complexity of an already challenging problem. In this case, the study carried out in this work would also be relevant for the development of a mathematical model and a specially-tailored solution procedure.

CRedit authorship contribution statement

Isabel Correia: Conceptualization, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. **Teresa Melo:** Validation, Writing – original draft, Writing – review & editing, Conceptualization, Investigation, Methodology.

Data availability

Data will be made available on request.

Acknowledgments

This work was partially supported by Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the projects UIDB/00297/2020 (<https://doi.org/10.54499/UIDB/00297/2020>) and UIDP/00297/2020 (<https://doi.org/10.54499/UIDP/00297/2020>) (Center for Mathematics and Applications). This support is gratefully acknowledged. We also thank the anonymous reviewers for their valuable comments.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.cie.2024.110244>.

References

- Alarcon-Gerbier, E., & Buscher, U. (2022). Modular and mobile facility location problems: A systematic review. *Computers & Industrial Engineering*, 173, Article 108734.
- Allman, A., & Zhang, Q. (2020). Dynamic location of modular manufacturing facilities with relocation of individual modules. *European Journal of Operational Research*, 286, 494–507.
- Antunes, A., & Peeters, D. (2001). On solving complex multi-period location models using simulated annealing. *European Journal of Operational Research*, 130, 190–201.
- Badri, H., Bashiri, M., & Hejazi, T. H. (2013). Integrated strategic and tactical planning in a supply chain network design with a heuristic solution method. *Computers & Operations Research*, 40, 1143–1154.

- Bashiri, M., Badri, H., & Talebi, J. (2012). A new approach to tactical and strategic planning in production-distribution networks. *Applied Mathematical Modelling*, 36, 1703–1717.
- Becker, T., Lier, S., & Werners, B. (2019). Value of modular production concepts in future chemical industry production networks. *European Journal of Operational Research*, 276, 957–970.
- Boschetti, M. A., Letchford, A. N., & Maniezzo, V. (2023). Matheuristics: Survey and synthesis. *International Transactions in Operational Research*, 30, 2840–2866.
- Ceschia, S., Gansterer, M., Mancini, S., & Meneghetti, A. (2023). The on-demand warehousing problem. *International Journal of Production Research*, 61, 3152–3170.
- Clavijo López, C., Crama, Y., Pironet, T., & Semet, F. (2024). Multi-period distribution networks with purchase commitment contracts. *European Journal of Operational Research*, 312, 556–572.
- Cordeau, J. F., Klibi, W., & Nickel, S. (2021). Logistics network design. In T. G. Crainic, M. Gendreau, & B. Gendron (Eds.), *Network Design with Applications to Transportation and Logistics* (pp. 599–625). Cham: Springer.
- Correia, I., & Captivo, M. E. (2006). Bounds for the single source modular capacitated plant location problem. *Computers & Operations Research*, 33, 2991–3003.
- Correia, I., & Melo, T. (2016). Multi-period capacitated facility location under delayed demand satisfaction. *European Journal of Operational Research*, 255, 729–746.
- Correia, I., & Melo, T. (2017). A multi-period facility location problem with modular capacity adjustments and flexible demand fulfillment. *Computers & Industrial Engineering*, 255, 729–746.
- Correia, I., & Melo, T. (2021). Integrated facility location and capacity planning under uncertainty. *Computational & Applied Mathematics*, 40, 175.
- Correia, I., & Melo, T. (2022). Distribution network redesign under flexible conditions for short-term location planning. *Computers & Industrial Engineering*, 174, Article 108747.
- Cortinhal, M. J., Lopes, M. J., & Melo, M. T. (2015). Dynamic design and re-design of multi-echelon, multi-product logistics networks with outsourcing opportunities: A computational study. *Computers & Industrial Engineering*, 90, 118–131.
- Costa, Y., & Melo, T. (2024). Facility location modeling in supply chain network design: Current state and emerging trends. In J. Sarkis (Ed.), *The Palgrave Handbook of Supply Chain Management* (pp. 809–845). Springer International Publishing, Cham: Palgrave Macmillan, http://dx.doi.org/10.1007/978-3-031-19884-7_101.
- Darvish, M., & Coelho, L. C. (2018). Sequential versus integrated optimization: Production, location, inventory control, and distribution. *European Journal of Operational Research*, 268, 203–214.
- Delmelle, E. M., Thill, J.-C., Peeters, D., & Thomas, I. (2014). A multi-period capacitated school location problem with modular equipment and closest assignment considerations. *Journal of Geographical Systems*, 16, 263–286.
- Jena, S. D., Cordeau, J.-F., & Gendron, B. (2015a). Dynamic facility location with generalized modular capacities. *Transportation Science*, 49, 484–499.
- Jena, S. D., Cordeau, J.-F., & Gendron, B. (2015b). Modeling and solving a logging camp location problem. *Annals of Operations Research*, 232, 151–177.
- Jena, S. D., Cordeau, J.-F., & Gendron, B. (2016). Solving a dynamic facility location problem with partial closing and reopening. *Computers & Operations Research*, 67, 143–154.
- Jena, S. D., Cordeau, J.-F., & Gendron, B. (2017). Lagrangian heuristics for large-scale dynamic facility location with generalized modular capacities. *INFORMS Journal on Computing*, 29, 388–404.
- Krarup, J., & Pruzan, P. M. (1983). The simple plant location problem: Survey and synthesis. *European Journal of Operational Research*, 12, 36–81.
- Lahy, A., Huckle, K., Sleeman, J., & Wilson, M. (2022). An introduction to flexible, on-demand warehousing: E-space. In Y. Wang, & S. Pettit (Eds.), *Digital Supply Chain Transformation: Emerging Technologies for Sustainable Growth* (pp. 81–98). Cardiff: Cardiff University Press, chapter 5.
- Melo, M. T., Nickel, S., & da Gama, F. Saldanha (2009). Facility location and supply chain management: a review. *European Journal of Operational Research*, 196, 401–412.
- Melo, M. T., Nickel, S., & da Gama, F. Saldanha (2012). A tabu search heuristic for redesigning a multi-echelon supply chain network over a planning horizon. *International Journal of Production Economics*, 136, 218–230.
- Nickel, S., & da Gama, F. Saldanha (2019). Multi-period facility location. In G. Laporte, S. Nickel, & F. Saldanha da Gama (Eds.), *Location Science* (2nd ed.). (pp. 303–326). Heidelberg: Springer, chapter 11.
- Sauvey, C., Melo, T., & Correia, I. (2020). Heuristics for a multi-period facility location problem with delayed demand satisfaction. *Computers & Industrial Engineering*, 139, Article 106171.
- Silva, A., Aloise, D., Coelho, L. C., & Rocha, C. (2021). Heuristics for the dynamic facility location problem with modular capacities. *European Journal of Operational Research*, 290, 435–452.
- Thanh, P. N., Bostel, N., & Péton, O. (2008). A dynamic model for facility location in the design of complex supply chains. *International Journal of Production Economics*, 113, 678–693.
- Thanh, P. N., Bostel, N., & Péton, O. (2012). A DC programming heuristic applied to the logistics network design problem. *International Journal of Production Economics*, 135, 94–105.
- Thanh, P. N., Péton, O., & Bostel, N. (2010). A linear relaxation-based heuristic approach for logistics network design. *Computers & Industrial Engineering*, 59, 964–975.
- Unnu, K., & Pazour, J. (2022). Evaluating on-demand warehousing via dynamic facility location models. *IIE Transactions*, 54, 988–1003.
- Wilhelm, W., Han, X., & Lee, C. (2013). Computational comparison of two formulations for dynamic supply chain reconfiguration with capacity expansion and contraction. *Computers & Operations Research*, 40, 2340–2356.